Written Testimony of David Burt
Child Online Protection Act Commission
July 20, 2000

I) Introduction

Thank you for allowing me this opportunity to submit testimony to the Commission on Child Online Protection. In my testimony I will discuss the current state of Internet content management (ICM) technologies, how ICM technology works, the evidence gathered to date regarding the effectiveness of ICM technology, and a proposal for further study.

II) History and current state of Internet content management technology

Internet content management technology, sometimes referred to as "filtering software" or "blocking software", first appeared commercially in 1994. ICM software appeared in response to the increasing availability of graphical Internet access and the accompanying pornographic web sites. The early versions of ICM software relied heavily on artificial intelligence (AI) to block access to pornographic or otherwise objectionable web sites. When a user attempted to access a web site that contained certain words or phrases, such as "XXX" or "sex", the screen would display a message informing the user that the filter was blocking access to the web site. Artificial intelligence is in fact quite good at identifying pornographic web sites, since pornographic web sites usually use a specific set of words such as "adult", "teen", "XXX", "porn", etc. to describe themselves. Some critiques of ICM software to this day leave the reader with the impression that ICM has never progressed beyond this early state.

It quickly became apparent that artificial intelligence software alone was not an acceptable solution to the challenges of Internet content management. AI technology has difficulty distinguishing between a news story about the Internet pornography business or an anti-pornography web site and a real pornography site. While some ICM vendors still offer products that rely on AI, the most widely used ICM products today either do not use AI or offer AI as a "fail safe" option the more cautious user may choose to enable.

Artificial intelligence is still heavily used as an intermediary step by the larger ICM vendors, including the one I work for, N2H2. Like other ICM vendors, N2H2 has found that sites identified by AI must then be subjected to human review to determine the content. Indeed, many Internet users are now discovering that automated search engines are a poor substitute for human review. [1]

Instead of relying on AI, N2H2 and our largest competitors rely on what is usually called "URL blocking" or "address blocking". URL blocking involves the compilation of lists of web site URLs (Uniform Resource Locator) that have been determined by a human reviewer to belong to a content category. Early versions of URL blocking software typically offered users a small numbers of the most obvious categories of sites users would find objectionable, such as "pornography", "hate speech", or "bomb making", or simply bundled all such objectionable material into a single category.

As the popularity of these URL blocking software programs spread, customers began to ask vendors to supply more categories and finer "granularity" in category selections. Schools didn't want students using web-based chat or e-mail. Corporations didn't want

employees visiting sport sites or engaging in on-line trading. Libraries wanted to block pornography but not artistic nudity or sex education materials.

This market-driven push for greater flexibility and granularity led to the evolution from "filtering software" to Internet content management technologies. Today's ICM vendors offer customers an abundance of choices. N2H2 currently offers 34 categories with six "allow exceptions", allowing for hundreds of possible combinations. [2] WebSense offers 65 categories, [3] I-Gear 24 categories, [4] SmartFilter 31 categories, [5] X-Stop 28 categories, [6] Cyber Patrol 12 categories, [7] and SurfWatch 21 categories. [8]

A decision by an organization to purchase ICM software offers literally thousands of possible options, enabling diverse users such as Internet service providers, schools, business, libraries, government agencies, and individuals to choose a solution that meets very specific needs. By empowering choice, ICM technology liberates organizations from "one-size-fits-all" Internet access.

The widespread acceptance of ICM technology offers compelling testimony to the success of the ICM approach. According to a recent International Data Corporation survey, 82 percent of companies with more than 1,000 employees plan to purchase ICM software over the next 12 to 24 months. [9] A May 1999 report by Quality Education Data estimates that increased usage of ICM software in K-12 schools will increase to 71.5% in the 1999-2000 school year over the current 52.5% of U.S. school districts that used ICM in the 1998-1999 school year. [10] Hundreds of Internet service providers, including industry leader America Online, offer consumers the choice of filtered Internet service. The compatibility of ICM technology with good service was underscored recently when the Gwinnett County (GA) Library System, a public library that filters all Internet access, was given the prestigious "Library of the Year" award by Library Journal. [11]

Critics of ICM technology sometimes invoke the fear that individuals using ICM will somehow suffer because they will be denied access to vital information. Typical examples that are given of potential harms caused by ICM are students who will be placed at a competitive disadvantage because they will be unable to master the Internet, teens who will become pregnant or contract a venereal disease because they will be denied access to sexual health information, and gay teens who will suffer from depression or even commit suicide because they were denied access to gay web sites.

Such hyperbole has yet to be shown to match reality. Despite the fact that literally millions of students have relied on ICM enabled Internet access for years, ICM critics present no studies or statistics to suggest that these students are any less computer literate, well-educated, or emotionally well-adjusted than peers who use unfiltered Internet access. Further, ICM critics fail to even cite a single anecdote of any teen ever becoming depressed, contracting a venereal disease, becoming pregnant, committing suicide, or even receiving a bad grade on a paper because of ICM software. Millions of Americans depend on Internet access using ICM technology as their primary means of accessing the Internet. Today's ICM technology is woven into the fabric of mainstream Internet access.

III) How Internet content management technology works.

As explained in the previous section, ICM technology involves the use of "block lists" of human-reviewed web sites which administrators can choose to enable or disable. Most vendors of ICM lists select the content of these lists based on carefully defined, objective, and openly published standards.

Probably the most objective and granular ICM lists involve material of a sexual nature. N2H2 has six categories devoted to sexual material, "Adults only", "Lingerie", "Nudity", "Porn", "Sex", and "Swimsuits". Additionally, N2H2 has four "Allow exception categories" related to sexual material: "Education", for sexually explicit material that is of an educational nature, "History", for material of historic value, such as the Starr Report, "Medical", for material such as photographs of breast reduction surgery, and "Text", for pornographic or sexual material that only contains text.

Websense offers five sex-related categories:

> *Adult content. Sites featuring full or partial nudity reflecting or establishing a sexually oriented context, but not sexual activity (3.3);sexual paraphernalia; erotica and other literature featuring, or discussions of, sexual matters falling short of pornographic; sex-oriented businesses such as clubs, nightclubs, escort services, password/verification sites. Includes sites supporting online purchase of such goods and services.*
>
> *Nudity. Sites offering depictions of nude or seminude human forms, singly or in groups, not overtly sexual in intent or effect.*
>
> *Sex. Sites depicting or graphically describing sexual acts or activity, including exhibitionism.*
>
> *Sex Education. Sites offering information on sex and sexuality, with no pornographic intent.*
>
> *Lingerie and Swimsuit. Sites offering views of models in suggestive but not lewd costume; suggestive female breast nudity. Also classic "cheesecake" art and photography.* **12**

I-Gear offers seven sex-related categories:

> *Sex/Acts*
> *Sites depicting or implying sex acts, including pictures of masturbation not categorized under sexual education. Includes sites selling sexual or adult products.*
> *Sex/Attire*
> *Sites featuring pictures that include alluring or revealing attire, lingerie and swimsuit shopping areas, or supermodel photo collections but do not involve nudity.*
> *Sex/Personals*
> *Sites dedicated to personal ads, dating, escort services, or mail-order marriages.*
> *Sex/Nudity*
> *Sites with pictures of exposed breasts or genitalia that do not include or imply sex acts. Includes sites with nudity that is artistic in nature or intended to be artistic, including photograph galleries, paintings that may be displayed in museums, and other readily identifiable art forms. Includes nudist and naturist sites that contain pictures of nude individuals.*
> *Sex Education [Super Category] SexEd/Basic*
> *Sites providing information at the elementary level about puberty and reproduction. Includes clinical names for reproductive organs (e.g., penis).*

SexEd/Advanced
Sites providing medical discussions of sexually transmitted diseases such as syphilis, gonorrhea, and HIV/AIDS. May include medical pictures of a graphic nature. Sites providing information of an educational nature on pregnancy and family planning, including abortion and adoption issues. Sites providing information on sexual assault, including support sites for victims of rape, child molestation, and sexual abuse. Sites providing information and instructions on the use of birth control devices. May include some explicit pictures or illustrations intended for instructional purposes only. May include slang names for reproductive organs, or clinical discussions of reproduction.
SexEd/Sexuality
Sites dealing with topics in human sexuality. Includes sexual technique, sexual orientation, cross-dressing, transvestites, transgenders, multiple-partner relationships, and other related issues. **13**

N2H2 and other ICM vendors have developed a number of techniques for identifying web sites to add to our lists. The most common technique is the use of "robots": automated programs the search the web for web sites that contain certain words and phrases included in domain names, meta tags, or page text. N2H2 has 70 servers devoted to searching the web for candidate sites, along with multiple T3 and T1 lines to provide adequate bandwidth. This initial "catch" of candidate URLs is then matched against our existing database, and subjected to more complex AI algorithms. These automated processes continuously feed a list of sites to N2H2's review department.

ICM vendors also employ other methods to identify content to be rated. ICM vendors make use of content already indexed in the various search engines to identify candidate URLs using "search parasites." N2H2 makes use of a technique called "spidering", where a "robot" program retrieves URLs linked to pornography sites, particularly "pornography search engines" such as Persian Kitty and Naughty.com. Another technique N2H2 uses is performing "whois" searches of domain name registries for new domain name registrations that contain words commonly association with pornography sites such as "xxx" or "adult". Finally, N2H2 monitors Usenet newsgroups and e-mail lists devoted to announcing new pornography sites.

Further, nearly all of the sites ICM companies are trying to find are also trying to be found by users. Many sites, particularly commercial pornography sites, go to great lengths to be found by users, and thus are easily found by ICM companies. Even the more elusive sites, such as child pornography and illegal software pages, want to be found by their end users. This is one of the reasons that filtering the Internet is possible. Content placed on the Internet without anyway for anyone to find it really doesn't pose much of a threat to anyone.

The N2H2 review department consists of approximately 120 full-time and part-time reviewers. The N2H2 review department has a full-time equivalent (FTE) complement of 60 employees, employed 40 hours per week. N2H2 employs reviewers fluent in 15 languages, to keep up with the increasing internationalization of the Internet. These 60 FTE review staff spend 2400 person hours each week reviewing approximately 75,000 URLs, which are added to our database of millions of URLs that that N2H2 has reviewed since 1995. This translates into about two minutes spent reviewing each URL. About one in 4 URLs identified by AI as candidates for adding to our category lists are actually

added. Therefore, each week about 20,000 new URLs are added to our category lists that are currently at 4.7 million URLs. Each URL effects 1 or many web pages. One method of calculating the number of webpages tagged for filtering shows over 15 million indexed webpages.

With the size of the World Wide Web estimated at 1.5 billion pages, [14] and new web sites appearing at a rate of 4,400 per day, [15] the task of keeping up with new web sites seems daunting. However, ICM vendors are not interested in reviewing *every new web page*, nor is their any need to do so. ICM vendors need only concern themselves with *new web sites featuring content that needs to be rated, or significant changes in the content of already-rated web sites*. The studies of the current size and growth of the web do not tell us what fraction of "new web pages" corresponds with "new web sites featuring content that needs to be rated, or significant changes in the content of already-rated web sites". While the Lawrence-Giles study found that 1.5% of web pages were pornographic, they did not find what portion of new web pages were new pornography sites. Therefore, it does not follow that statistics of the rate of web growth can be used to claim that keeping up with the growth of new web sites with content that needs to be rated is unlikely or impossible. Based on N2H2's internal sampling and customer feedback, N2H2 feels confident that we have adequate resources to keep up.

The criteria used to rate URLs are both public and well defined, but the actual lists of URLs are not made public by nearly all ICM vendors. There are two obvious reasons for this. First, as described earlier, a great deal of human labor is involved in creating these lists. Creating N2H2's list of 4 million+ reviewed URLs required hundreds of thousands of person hours, at a cost of quite literally, millions of dollars. Very few companies would willingly give away such expensive and valuable proprietary data. Second, it would be irresponsible to publish a gigantic list of pornographic web sites, as this information might well land in the hands of children. This point was illustrated graphically last month when Burger King restaurants in the United Kingdom gave away a CD-ROM to children that contained a filter with a published list of over 2,000 pornographic web sites. After complaints from parents and child safety groups, Burger King recalled the CD-ROM. [16]

If a user or webmaster is concerned that a particular site might be wrongly included on an ICM vendor's list, nearly every ICM vendor has e-mail links where such a request can be made. The makers of Cyber Patrol, SurfWatch, and WebSense provide on their web sites a search function where anyone can check to see if a URL is currently being blocked. [17] N2H2 takes this concept one step further by providing a link at the bottom of every web page in our ResourceBar, where an end user who encounters a site they feel is wrongly blocked can instantly send feedback to N2H2's review staff. The end user has the choice of submitting the request for review anonymously, or providing their e-mail address in order to get a response.

IV) Evidence of the effectiveness of ICM technologies

Research on the effectiveness of ICM technologies has been highly politicized. Nearly all of the research has been conducted by individuals with a strong bias for or against ICM

(and I of course include myself here). Nearly all of the research, and I include some of my own work here, involves samples that are far too small. As my co-panelist Christopher Hunter rightly points out:

> *The majority of reports of Internet content filters being both underinclusive (failing to block the worst pornography, hate speech, violence, etc.), and overinclusive (blocking non-sexual, non-violent content), have come from journalists and anti-censorship groups who have used largely unscientific methods to arrive at the conclusion that filters are deeply flawed*. **18**

Studies using small samples

"Faulty Filters: How Content Filters Block Access to Kid-Friendly Information on the Internet", December 1997. Electronic Privacy Information Center. **19**
This study was conducted by the Internet free speech organization EPIC. EPIC states on its web site that "content filtering has been shown to pose its own significant threats to free expression on the Internet." **20** EPIC makes the striking claim that in EPIC's testing users were "denied access to 99 percent of material that would otherwise be available without the filters." However, EPIC did not actually test Internet filters to arrive at this figure, they tested an experimental filtered search engine using AltaVista in conjunction with the ICM product Net Shepherd. At the time of the study, AltaVista limited search results to 200 URLs, hence the "99%" blocked results. Further, EPIC did not use a fixed sample: researchers simply attempted to perform searches.

"The Internet Filtering Assessment Project", Karen Schneider, 1997. **21**
"The Internet Filtering Assessment Project" is the work of a critic of ICM software, Karen Schneider. Schneider used a team of 40 volunteers to test filters and found that "Over 35% of the time, the filters blocked some information they needed to answer a question." **22** Like the EPIC study, Schneider used a loose and open-ended searching method to determine if filters wrongly blocked sites. Schneider herself accurately described her work in her summary:

> *TIFAP was not a scientific study; it lacked controls, the actual conditions could not be verified, and, due to limited volunteers and resources, we could not consistently test all products the same way. The survey instruments are as amateurish as you would expect from people who do not design surveys.* **23**

Schneider tested mostly word-blocking filters by attempting to perform searches for information designed to trip word-blockers, such as "nursery rhymes, (pussycat, pussycat)". **24**

Censorware Project Reports
The Censorware Project, a group that describes its mission as "dedicated to exposing the phenomenon of *censorware*: software which is designed to prevent *another person* from sending or receiving information, usually on the web. A gag or blindfold is the physical equivalent of what such software does." **25** From 1998 to the present, the Censorware Project has issued a series of reports detailing URLs wrongly blocked by ICM vendors. With the exception of an analysis of Utah logs (see next section, "Studies using large samples"), the reports issued by the Censorware project do not attempt to set the

occurrence of “misblocks” in any context.

The Censorware Project lists four reports exposing misblocked web sites by ICM products. [26]. Cyber Patrol is charged with 67 misblocks, [27] Websense with 12 misblocks, [28] X-Stop with 50 misblocks, [29] and Bess with 34 misblocks. [30]

Unfortunately, no context for this information is provided, such as whether these small numbers of misblocks constitute all sites wrongly blocked by each filter or how big an impact these blocks would have on typical Internet traffic. Therefore, the only conclusions that can be drawn from this information is that these ICM products have been shown to block a small number of URLs incorrectly.

“Filtering the Future? Software, Filters, Porn, PICS and the Internet Content Conundrum” Christopher Hunter, 1999. [31]

“Filtering the Future”, a master’s thesis by Christopher Hunter claimed that Internet filters "improperly block 21% of benign content". [32] The sample used in the study was a non-random sample of 200 sites. The study tested for blocking of "sex", "profanity", "nudity", and "violence", with ICM products configured to block all categories, including "gambling" and "alcohol", against a sample of "purposefully selected" sites, including gambling and alcohol sites, which were then counted as "wrongly blocked". Mr. Hunter later stated:

> *I readily admit that I need a better sample and that my results shouldn't necessarily be generalized to the entire universe of web pages.* [33]

“A Guide to Filtering Software”, David Burt, Parts I and II, 1999. [34]

In 1999 I was asked to write two articles for “Dr. Laura Perspective Magazine” reviewing ICM products. My intention was not to conduct a scientific survey but to offer more of a “thumbnail sketch” of product reviews.

I reviewed 14 ICM client products and “clean ISPs”. For this review I selected 250 web sites, 100 randomly selected pornography sites, 75 purposefully selected sites promoting drugs, hate, and bomb-making, and 75 purposefully selected “innocent sites” related to gay rights, feminism, breast cancer, and news stories about hate speech and online pornography. The various products were between 85% and 99% effective at blocking pornography, and less effective at blocking other undesirable sites. Most of the products blocked none of the “innocent sites”, while several, particularly the AI-based products did block innocent sites.

Studies using large samples

Considering the vast size of the Internet, and the fact that ICM products are only targeting a small portion of the Internet, it quickly becomes obvious that the only way to accurately test ICM products is to test against large samples of URLs. Fortunately, two such tests have been conducted, “Censored Internet Access in Utah Public Schools”, a study of SmartFilter by Michael Sims of the Censorware Project, and “Dangerous Access, 2000 Edition”, a study of Bess and Cyber Patrol, by David Burt. Even though Mr. Sims and

myself are on opposite sides of the debate over the effectiveness of ICM software, the bottom-line findings we both arrived at over ICM error rates were remarkably similar.

"Censored Access in Utah Public Schools", by Michael Sims, 1999. [35]
In 1998, anti-filtering activist Michael Sims obtained one month's worth of Internet log files from the Utah Education Network, which provides Internet access for nearly all of Utah's public schools. The Utah schools use an ICM product, Smart Filter. In March of 1999, Sims issued a report analyzing the filtered log files. The logs recorded 53,103,387 total files accessed, of which 205,737 were blocked, 193,272 under the Smart Filter "sex" category. When Sims removed banner ads and image files, achieving a rough approximation of "page views", Sims records the numbers as 15,434,442 pages accessed, of which 95,059 were blocked, 86,957 under the Smart Filter "sex" category. Sims reported about 300 pages wrongly blocked. On June 28, 1999 the Censorware Project wrote a follow-up report that listed the total number of wrongly blocked pages at 5,601, but did not list all the actual pages. [36] The 5,601 wrongly blocked pages Sims found out of 15,434,442 pages accessed results in an error rate of .036%.

"Dangerous Access, 2000 Edition", by David Burt, 2000. [37]
As part of a report discussing the spread of Internet pornography I analyzed the filtered log files of two public libraries earlier this year. I found that Cyber Patrol used at the Tacoma (WA) Public Library wrongly blocked 1,853 pages out of 2,510,460 pages accessed, or .073%, and that Bess used at the Public Library of Cincinnati and Hamilton County wrongly blocked 732 pages out of 3,717,383 pages accessed, or .019%.

The advantages of log analysis studies versus studies involving small, purposefully selected samples are both considerable and obvious. First, a researcher with a possible bias is not creating the sample of URLs used, they are being taken directly from a real-world sample. Second, the size of these samples makes it much more likely that they will accurately reflect real-world conditions. Third, the rate of overall blocking by the ICM product is not being determined by the researcher, but rather is part of the original sample.

Even with these advantages, a researcher evaluating log files must still make decisions about which blocks have been applied incorrectly. Mr. Sims and myself used somewhat different criteria for evaluating "wrongly blocked" web sites. I included most sexually explicit material as being correctly within the parameters of the filtering categories used by Cyber Patrol and Bess. Mr. Sims, on the other hand, counted as wrongly blocked many sexually-themed web sites such as www.playboy.com, commenting that "Besides the photographs, Playboy of course has many interviews and well-written articles." [38]

In spite of these differences in attitude, it is well worth noting that both log analyses came to very similar conclusions about the level of inappropriate blocking. Sims found that Smart Filter wrongly blocked .036% of the time, and 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, and it is my belief that further study will verify this.

V) Suggestion for further study

My own interpretation of what the evidence gathered to date suggests is that the best ICM products accurately block over 90% of pornographic web sites, and erroneously block less than .1% of non-pornographic web sites.

However, 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. I first proposed such testing in December of 1998, when I testified before the National Commission on Library and Information Science:

> *Because of this lack of reliable data, I'd like to suggest that this commission take the lead in producing better data. I think that conducting a study that could tell us what we need to know would be pretty straightforward. Such a study would involve writing a special computer program that would run on Internet workstations in several public libraries that either filter for all patrons, or just for all minor patrons. First, the program would record the address of every website that every patron visited. Second, the program would record the address of every website someone tried to access, but was blocked by the filter. Third, the program would record if the filter were overridden in any of the cases where a patron encountered an inappropriate block. With this method we could actually get a reasonable idea of: 1) What exactly are patrons being prevented from viewing in libraries that filter, 2) How often are patrons prevented from viewing web sites they want to access, and 3) When a patron encounters an inappropriately blocked website, how likely are they to ask to see it.* **39**

Unfortunately, NCLIS did not express any interest in facilitating such a study. I find it heartening now to hear others, such as my co-panelist Mr. Hunter, also expressing the need for more rigorous studies on ICM effectiveness. Since I testified before NCLIS, my thoughts on how to conduct an ICM study have evolved.

The purpose of such a study should be twofold: 1) to determine how effective filters are at blocking pornographic web sites; 2) to determine the extent of "overblocking" of innocent web sites on Internet access. To this end two sets of data would be needed: a large sampling of pornographic web sites, and a large sampling of "typical" web traffic.

I would propose that the study be conducted by a reputable research facility well versed in software testing methodologies, using standard laboratory control procedures. The ICM vendors themselves could fund the study.

There are a number of ways to obtain the required data. The participating vendors themselves could each supply several thousand pornographic URLs to form a combined list that would be tested against all products. Alternatively, the pornographic URLs could be obtained through search engines and pornographic directory sites such as Naughty.com. The larger the sample the better, and I think a minimum of 25,000 unique pornographic URLs would be required.

The "typical" Internet traffic could be obtained from the log files of a university, library, or Internet Service Provider, then reduced to only unique web page files. I think a minimum of 250,000 unique pages would be required.

A lab could set up a server for each ICM product, with each product configured to block only pornography, then simultaneously run scripts containing the test data against each product. Once the testing was complete the results could be measured to determine 1) the percentage of pornographic URLs blocked by each product; 2) the percentage of typical web traffic blocked by each product.

More difficult is determining the amount of "wrongly blocked" URLs. Each URL from the "typical" web traffic data that was blocked would have to be examined and judged to be "rightly blocked" or "wrongly blocked". Considering that 1% to 3% of the "typical" web traffic would likely be blocked, this would involve thousands of URLs. N2H2's experience has been that it requires on average 2 minutes to review a URL. If the testing generated 10,000 blocked URLs, this would require 333 person hours to examine. Additionally, there would likely be some difference among the reviewers as to what was wrongly blocked, so ideally two different reviewers should review each URL.

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. Lisa Guernsey, "The Search Engine as Cyborg", *The New York Times*, June 29, 2000.
2. N2H2, *Human Review Filtering Solution*, http://www.n2h2.com/solutions/filtering.html.
3. WebSense, *Websense 4: Database Categorization Criteria*, http://www.websense.com/products/categories/version4.cfm
4. URLabs, *URLabs content category definitions*, http://www.symantec.com/urlabs/public/support/faq/categories.html
5. Secure Computing, *Frequently Asked Questions*, http://www.securecomputing.com/index.cfm?sKey=275
6. X-Stop, *X-Stop XLM for Microsoft NT Proxy Manual*, http://www.xstop.com/docs/Manual_xlm_msnt30b.pdf
7. Cyber Patrol, *Category Definitions - 1/20/99*, http://www.cyberpatrol.com/cybernot/criteria.htm
8. SurfWatch, *How we filter*, http://www1.surfwatch.com/about/filter.html
9. Chris Christiansen, "Worldwide Market for Corporate Internet Access Control", *International Data Corporation*, July 1999.
10. Quality Education Data, *Internet Usage in Public Schools, 4th edition*, 1999.
11. Library Journal, *Library of the Year*, June 15, 2000.
12. WebSense, *WebSense 4: Database Categorization Criteria*, http://www.websense.com/products/categories/version4.cfm
13. URLabs, *URLabs content category definitions*, http://www.symantec.com/urlabs/public/support/faq/categories.html
14. David Lake, "The Web: Growing by 2 Million Pages a Day", *The Industry Standard*, February 28, 2000, http://www.thestandard.com/research/metrics/display/0,2799,12329,00.html
15. Lake.
16. Jonathan Lambeth, "Burger King gives away porn addresses", *UK Telegraph*, June 26, 2000.
17. SurfWatch "Test-a-Site", http://www1.surfwatch.com/testasite/. Cyber Patrol "CyberNot Search Engine", http://www.cyberpatrol.com/cybernot/. WebSense, "Site Look Up", http://database.netpart.com/site_lookup.html.
18. Christopher Hunter, *Cyberporn, Filters, and Public Policy: A Content Analysis Research Proposal study proposal*, 2000.
19. EPIC, *Faulty Filters: How Content Filters Block Access to Kid-Friendly Information on the Internet*, December 1997. http://www.epic.org/Reports/filter_report.html.
20. EPIC, *Filters and Freedom*, http://www.epic.org/bookstore/filters&/.
21. Karen Schneider, *The Internet Filtering Assessment Project*, 1997, http://www.bluehighways.com/tifap/learn.html.
22. Schneider.
23. Schneider.
24. Schneider.
25. Censorware Project, *Welcome to Censorware.org*, http://www.censorware.org/intro/
26. Censorware Project, *Censorware Project Special Reports*, http://www.censorware.org/reports/

27. Censorware Project, *Blacklisted by Cyber Patrol: From Ada to Yoyo*, December, 1997, http://www.censorware.org/reports/cyberpatrol/ada-yoyo.html.
28. Censorware Project, *Protecting Judges Against Liza Minnelli: The WebSENSE Censorware at Work*, June, 1998, http://www.censorware.org/reports/liza.html
29. Censorware Project, *The X-Stop Files: Deja Voodoo*, http://www.censorware.org/reports/xstop/. The Censorware Project may have found other misblocks as well. Just before the X-Stop report was released the ACLU issued a "statement of undisputed facts" that stated there were "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". See http://www.aclu.org/court/loudoun_facts.html.
30. Censorware Project, *Passing Porn, Banning the Bible: N2H2's Bess in public schools*, http://www.censorware.org/reports/bess/
31. Christopher Hunter, *Filtering the Future? Software, Filters, Porn, PICS and the Internet Content Conundrum*, 1999.
32. Hunter.
33. David Burt, *ALA touts filter study whose own author calls flawed*, 2-18-2000, http://www.filteringfacts.org/hunter.htm.
34. David Burt, "A Guide to Filtering Software", *Dr. Laura Perspective*, July, 1999, p. 12. And David Burt, "An Update on Filtering Software", *Dr. Laura Perspective*, October, 1999, p. 14.
35. Michael Sims, "Censored Access in Utah Public Schools", *Censorware.org*, March 1999, http://www.censorware.org/reports/utah/main.shtml.
36. Jamie McCarthy, "Lies, Damn Lies, and Statistics: A followup to our March 1999 Utah SmartFilter report", June 28, 1999, http://www.censorware.org/reports/utah/followup/.
37. David Burt, *Dangerous Access 2000 Edition*, March 1999.
38. Sims. http://www.censorware.org/reports/utah/appendix.shtml.
39. David Burt, *Testimony before the National Commission on Library and Information Science*, November 10, 1998, http://www.filteringfacts.org/nclis.htm.