Chapter 4

Internet Filtering

Sarah Houghton-Jan

Abstract

Laws and library practices regarding content filtering have a significant impact on library customers’ constitutionally protected access to information, their privacy, and their right to free speech. Libraries have a responsibility to be informed about relevant laws, technologies, and best practices in order to protect the intellectual freedom rights of our customers in an increasingly digital information landscape. This chapter of Privacy and Freedom of Information in 21st-Century Libraries aims to help librarians begin to fulfill that responsibility.

Internet filtering, also referred to as content filtering or censorware, is one of the most pervasive and recurring intellectual freedom challenges for libraries worldwide. For some libraries, and indeed some entire countries, there is no question of whether or not to filter—filters in libraries are simply mandated by the government. In the United States, however, there is still at least a cursory nod to intellectual freedom and privacy from our governing agencies. But filtering in the United States is tied to optional, and tempting, library funding.

A divisive issue in libraries, the question of content filtering is central to our communities’ future of information access. Content filtering involves values of a fundamental nature, so any solution results in parties feeling that they are giving up something of profound importance. There is no compromise, no middle ground. The American Library Association (ALA) has stated unequivocally that any type of restriction on a person’s, including a child’s, access to any type of content is unacceptable. The ALA Library Bill of Rights says very clearly: “A person’s right to use a library should not be denied or abridged because of origin, age, background, or views.”1 To the ALA and many librarians, content filters of any kind are antithetical to the mission of the library to provide free and open access to all information.

Simply put, filters block legitimate content with no threat to children while simultaneously allowing access to graphic sexual content that does pose a threat. Internet filters effect content-based restrictions on free speech. To most librarians, especially to those who champion intellectual freedom, Internet filters are constitutionally unacceptable for use in libraries, or anywhere else for that matter.

Laws and Court Cases Related to Internet Filters

There are several laws and court cases that affect library use of Internet filters. Details about the court cases are explicated in the online sources listed in the gray box.

- First Amendment of the United States Constitution
- Child Online Protection Act (COPA) [passed as law, but overturned in courts as unconstitutional]
- Children’s Internet Protection Act (CIPA)
- state or local codes
- ACLU v. Gonzalez
- ACLU v. Miller
- ACLU v. Reno
- ALA v. Pataki
- ALA v. U.S. Department of Justice
- Mainstream Loudoun v. Board of Trustees of Loudoun County Library
- Bradburn v. North Central Regional Library District
- United States v. American Library Association
All filters function by filtering content based on some combination of the domain, IP address, keyword, and file type. Because the amount and dispersion of the content on the Internet is growing so quickly, filtering products start with a list of domains (website address) and IP addresses (where those websites are hosted) and add into the equation some element of the content (trigger words, phrases, file types, etc.).

Products that filter based on domains and IP addresses typically use a search engine (Google in almost every documented case) to run canned searches for trigger words or phrases, such as “sex videos.” That results list is then run through an algorithm that creates a blacklist of blocked pages for that topic or subject matter. Other algorithms block entire domains or IP addresses. Some companies have a staff member spot-check the auto-generated list for errors, but many have no human intervention at all. These domain blacklists generally include 250,000–2,000,000 domains or IP addresses, which are then blocked by the filtering software when a user attempts to access them.

Of paramount intellectual freedom concern to libraries is the methodology behind how content is classified in the filtering software. The automated classification processes and the whitelists and blacklists that filtering software companies develop are ferociously protected and never made publicly available to their customers. Filtering software companies do not tell their customers the types of things or what specific sites they block in each category. No examples are given, and no information beyond a one- or two-sentence description is offered for any company’s product. Their methodology is considered a company trade secret and vital to their continued success.

There are numerous workarounds to content filters that experienced users will fast-learn and easily use. Sites like Peacefire.org are dedicated to helping individuals get around filters. Another method of bypassing filters is through proxy servers, such as Psiphon and StupidCensorship. Some filtering sites therefore choose to filter proxy-avoidance sites, URL translators, and other workaround sites. This raises a new and wholly different intellectual freedom concern beyond the protection of children from sexually explicit material. Many political dissidents and others attempting to hide their identities or locations (sometimes not for wholly idealistic reasons) use these tools as a way to mask their information from government agencies and others seeking to do them harm. As a result, by disallowing the use of these sites in our public libraries, libraries have made it impossible for this group of users to gain access to the tools they need, sometimes for life-or-death reasons.

Technology of Filters

Content filters can be extremely powerful. Filters today employ artificial intelligence, image recognition, and complex keyword analysis algorithms to an extremely granular level. Filters still cannot successfully evaluate and determine the actual content, context, and intent of Web content of various media types—text, still images, video, audio, and more. As a result, filter performance is highly dependent on the program’s artificial intelligence content recognition and any possible administrative human intervention, as well as the chosen settings and features.

These laws and cases should be consulted for further information about local statutes related to Internet filtering, the application of CIPA, and the constitutional limitations placed on filtering implementations.

**Court Cases Relevant to Internet Filters in Libraries**

**ACLU v. Gonzalez**
www.paed.uscourts.gov/documents/opinions/07D0346P.pdf

**ACLU v. Miller**
www.aclu.org/technology-and-liberty/federal-district-court -decision-aclu-v-miller

**ACLU v. Reno II**
www.aclu.org/technology-and-liberty/feature-aclu-v-reno-ii

**ALA v. Pataki**
www.aclu.org/technology-and-liberty/feature-ala-v-pataki

**ACLU v. U.S. Department of Justice and Reno v. ACLU**
www.ciec.org

**Mainstream Loudoun v. Board of Trustees of Loudoun County Library**
www.tomwbell.com/NetLaw/Ch04/Loudoun.html

**Preliminary Injunction Against Child Online Protection Act and Judge Lowell Reed’s Decision**

**United States v. American Library Association (CIPA)**
administrator sees is different. What flexibility exists is different. It is of the utmost importance that libraries considering these products review all of the various factors at play in deciding if a product can work for them, and if so, which product will meet their needs the best.

Filter Accuracy

The accuracy of filters is key to the discussion of how Internet filters work in libraries, and everywhere else for that matter. All filters overblock (incorrectly blocking something unobjectionable) and underblock (incorrectly allowing something objectionable). The question is: how much do they do both, and is that failure rate an acceptable cost?

In filter accuracy studies from 2001 to 2008 (none were done in 2009–2010), the average accuracy success rating of all the tests combined is 78.347 percent. This means that on average, 78.347 percent of the time, the filtering software did what it was supposed to do. Bear in mind that these studies measure only text content, with only one exception of a study that examined filtering efficacy on images.2

If you look only at other studies done from 2007 to 2008 to get the best of the most recent software, we see a nominally higher accuracy percentage—83.316 percent—but the number of studies is limited and therefore leaves a larger margin for error. While filters may be getting a little better . . . they’re still wrong at least 17 percent of the time for text content, and wrong 54 percent of the time on image content.3

An interesting study was done on the effectiveness of home computer Internet filters in preventing unwanted exposure of children to harmful material. The researchers found through a longitudinal study that “the use of filtering and blocking software was associated with a modest reduction (40%) in unwanted exposure, suggesting that it may help but is far from foolproof.”4

Again and again, studies show that content is both overblocked and underblocked at consistent and equivalent rates, no matter what filter or what settings. Seventeen percent of the time, content is overblocked (i.e., benign sites are blocked incorrectly). Seventeen percent of the time, content is underblocked (i.e., sites deemed “bad” get through anyway).

The lesson that this teaches our regular Internet users is this: when you come to the library, your Internet use is going to often be blocked, usually incorrectly, and we won’t tell you why. The lesson that this teaches to our hardcore Internet adult site users is this: try, try again. Exchanging those statistics, ask yourself as an information professional if an overall accuracy rating of 83 percent is okay for websites? If an average accuracy rating of 46 percent is okay for images?

Privacy and Filters

Filtering provides several challenges to the library’s key principle of personal privacy and privacy of information needs. Before considering the implementation of filters or reassessing a current implementation, libraries need to consider issues of data collection, library privacy policies, confidentiality of information needs, and alternatives to filters.

The very nature of filtering software means that there are vast libraries of data about users’ Web and other computer use habits. Libraries need to find out from filtering product vendors what information the vendors are collecting about users’ surfing habits, if and how that information is connected to their computer session login or patron record, if the filtering company has access to that information, if the data is retained on the library’s servers, and if so for how long and in what format. As we are mindful about other user data, so should we be with this data. Whatever can be not collected, should be not collected. Whatever data can be anonymized, should be anonymized. Whatever data is left should be protected solemnly, and access to that data should be extremely limited.

A big part of libraries’ campaigns for Internet use safety and privacy has been to put Internet use policies into place. An associated policy has been the library privacy policy, largely created by libraries in response to the PATRIOT Act. Library privacy policies dictate what a library will do or not do with a customer’s data. Most state what information is collected and saved, how or where it is stored, and who can access it under what circumstances (e.g., a subpoena). The utilization of filters should create an added section to library privacy policies. Libraries need to state what information they are collecting about users and who has access to it.

Another impact of filters on privacy in libraries is a user’s need for confidentiality when it comes to his or her information needs. One of the principles on which libraries pride themselves is that anyone can ask a library staff person any question, access any resource, and the library will not freely make that information available to their friends or law enforcement (without a subpoena). Their information needs are secure and private when it comes to the library’s physical collections and library-selected digital collections, and most librarians believe that this principle should extend into the library’s de facto digital collection: the Internet. However, if a library places filters on computers and requires a library staff member to intervene to approve a blocked site, then this confidentiality evaporates. The user has to tell the library staff member what he or she wants to look at. And that information might pass through several library staff members’ hands before it makes it into the whitelist database or even through some of the filtering company’s employees’ hands as they make that change for the library instead. Customers with sensitive needs are very unlikely to be willing to ask a library staff member to
unblock a site for them about, for example, male impotence or divorce. By requiring staff intervention, libraries violate the principle of confidentiality of information needs, a key tenet of library user rights.

Beyond library policies for privacy and Internet use, there some alternatives to Internet filtering in use in the library world that help to protect users’ privacy, and which should be considered in the place of or in addition to filtering software. These alternatives are almost universally used by libraries that choose to not filter their users’ access.

Libraries can teach classes to self-registered attendees in the library, at school visits, during parent nights, and during visits to local Rotary Clubs and similar organizations. Internet safety for children, privacy, data security, and social media and privacy are common topics. Helping to protect users’ information is an important role of the library, and teaching users best practices is the most successful way to encourage data security and privacy.

Privacy screens can help, although in only a limited way, to keep what users are looking at on their screens private. Research at the San José Public Library into various types and brands of screens found that their physical zones of successful blocking behind the monitor were quite small. For all of the screens SJPL tested, one could see what users were viewing for about a 30 degree angle area behind them. Thus, privacy screens should be utilized only with the caveat that the computer screen is still visible to those seeking to view it.

Libraries can also consider the placement of computers. Placing computers in isolated areas will allow users to maintain their privacy. Placing computers in busy walkways with the screens facing out creates a problem, as they are very visible.

Some libraries create profiles that are age-based, allowing users who are under 18, or under 12, to log in only on certain computers. Placing children’s computers in an isolated area can help to protect the data that children are entering on the computer as fewer adults are likely to be wandering that area.

None of these is an ultimate solution for protecting user information. The library should do what it can to coach users to protect their information and privacy, but still rely on individual responsibility for data security and privacy.

**Intellectual Freedom and Filters**

Four of the six statements of principle in the American Library Association’s *Library Bill of Rights* are relevant to Internet filtering (emphasis the author’s):

- Materials should not be excluded because of the origin, background, or views of those contributing to their creation.
- Libraries should challenge censorship in the fulfillment of their responsibility to provide information and enlightenment.
- Libraries should cooperate with all persons and groups concerned with resisting abridgment of free expression and free access to ideas.
- A person’s right to use a library should not be denied or abridged because of origin, age, background, or views.

If a librarian believes in the ALA *Library Bill of Rights* and is willing to stand up for these core professional values, then it would appear that in no case is it acceptable for a library to filter the Internet for any group of people. Internet filters exclude material. By excluding material, they are censoring information and abridging free access to ideas. Libraries should work with other groups concerned with the same issue. And finally, and perhaps most poignant, no one’s library use or access to information should be affected by his or her age. According to the *Library Bill of Rights*, it is not acceptable to lessen a user’s library access because of the individual’s age, and filtering Internet access based on age is definitely lessening access. These are the principles that we librarians in the United States have agreed to champion. If we take this document seriously, then any Internet filtering implementation or legislation should continue to be a key point of contention for all ALA members.

A key issue brought up in the court decisions surrounding CIPA and Internet filtering in libraries is the idea of “selection versus censorship.” Some courts contend that installing filters is equal to library selection of materials, or collection development decisions, and that each individual library has the right to make those selection decisions and they do not violate First Amendment rights as a result.

I’d like to deconstruct that statement. Installing filters on your public access library computers is a bit like outsourcing to a single company all of your collection development—that one entity will be making all collection decisions about what is immediately accessible and what is not. You might be able to tweak it a little bit, but your default collection is already decided by someone else. Even more, because the process of selection for what sites are blocked by the filter is 99 percent automated by a computer script, installing a filter is like entrusting the entirety of your collection development to an automated computer system instead of to human beings.

If you’ll permit me to carry the analogy a little further, like requiring users to use interlibrary loan to get an item that is not in your collection, Internet filters require users to ask a librarian to override the system so that they can access an item that is not in the library’s “acceptable” Internet collection. And like the many users who are stopped by the hassle and delay of ILL, users of filtered
library computers are stopped in their tracks by the hassle
and delay of the filter’s big “no, no, no!” warning when they
try to access something that may very well be a totally legit-
imate site that in fact does not violate the library’s policies.

There is sufficient legal precedent that libraries can-
not legally adopt or enforce private rating schemes, which
is what Internet filtering software uses. The American
Library Association states that “when libraries restrict
access based on content ratings developed and applied
by a filtering vendor, sometimes with no knowledge of
how these ratings are applied or what sites have been
restricted, they are delegating their public responsi-
bility to a private agency.” The legality of this issue
is still being fought out in the courts, most recently in
Washington State, where it was decided that Internet fil-
ters are not censorship because filtering is equivalent to
collection development (see directly above for contradict-
ing argument). There are contrasting court opinions,
and the issues will likely be fought out for a long time, or until
the Congress passes new legislation.

Deciding what is on the software’s core blacklists
and whitelists is up to machines and filtering software
company staff who are untrained on freedom of informa-
tion, constitutional issues, or best practices for informa-
tion objectivity. Library staff have the ability, usually,
to add pages and domains to the whitelist or blacklist.
Subjectivity is a part of human nature, so who on the
library staff gets to decide what is bad and what isn’t?
What is the library’s procedure for adding something
to either list? How do we “select” in the future when the
software rolls out updated lists? How do we know how
the software makers decide what gets blocked and what
doesn’t? How do we ensure that this is an objective pro-
cess, an accurate process? The answer: we cannot.

How do we know that the morals and values of the
company CEO aren’t making their way into the software’s
lists, in violation of the library’s core principles of equal
and open access to all ideas and points of view? For
example, San José Public Library’s research showed that
WebSense filtering categories have subcategories, some of
which are divided into political positions. With one click
you can block only pro-choice or only pro-life websites or
choose to block only occult or nontraditional religious
sites. X-Stop was shown to block sites such as the Quaker
website and the National Journal of Sexual Orientation
Law, while CYBERsitter blocked sites like the National
Organization for Women. Libraries have a duty to ask
ourselves what values are already in the software’s algo-
rithms, and what procedures build the blacklists.

Internet filters do not constitute selection, and pay-
ning for the Internet does not constitute paying for pornog-
raphy, according to past ALA President Mitch Freedman.
Freedman wrote, “Just like buying the dictionary doesn’t
just pay for certain words, paying to provide public
Internet access doesn’t just pay for just the best or worst
of this amazing communications and information tool. We
don’t rip out unsavory interviews in Rolling Stone or edit
photography books—why would we cut swathes through
the Internet? . . . The filters cut with all of the subtlety of
a meat-cleaver.”

Many feel that automated artificial intelligence selec-
tion of what content to block does not constitute selection.
Chris Hansen of the ACLU is often quoted as saying that
mandating filters in libraries is like mandating that some
stranger randomly pull books off shelves and make them
unavailable, all the while not telling librarians or customers
why the books aren’t available or what books were pulled.

Another issue to consider is the Big Brother Factor.
It’s well-demonstrated that users behave differently if they
feel they are being watched. Libraries must consider what
effect filters will have on their users. Users may not try to
access sites they think might be blocked, worrying that
their use is being tracked. Users may not even try to go to
that site about incontinence, or the video showing women
how to conduct a breast self-exam. At libraries that filter,
library customers report often not being willing to ask
for something to be unblocked for them because they are
embarrassed as the library has already deemed what they
want to be unsavory. We must ask ourselves: how many of
our library customers walk away without the information
they need because of the Big Brother Factor?

Finally, the influence of outside lobbying groups on
local Internet filtering policies in libraries should not be
understated. Some groups, such as the Values Advocacy
Council and SafeLibrarians.org, have local affiliate organi-
izations that they expect to get Internet filters into local
school and public libraries. These groups provide local poli-
ticians in their like-minded political party with template
proposals for Internet filtering ballot measures, city council
resolutions, policy changes, and so on. This often provides
the politician, in his or her mind, with a clear winning plat-
form for the next election. These prewritten policy-change
templates require the politician to insert only his or her
city, school, or county name. With such an easily presented
fast lane to election supremacy, libraries and intellectual
freedom advocates must stand vigilant and constantly
remind politicians that their constituents include people
who believe in the right of choice, not only people who
believe in their right to remove everyone else’s choice.

Additional Library Challenges with
Filtering

There are several additional issues and challenges that
libraries face with Internet filtering, including the ques-
tion of libraries acting in loco parentis, the false sense of
security created by filters, the de-emphasis on education
from our government leaders and the courts, the need for Internet use policies, and finally the cost-benefit analysis of complying with CIPA.

Libraries often challenge Internet filtering, as they challenge limiting children’s access to certain books, as a problem of parents wanting the library to act in loco parentis—in the place of the parents. Libraries have traditionally not wished to fill that role, instead tending toward education of library users about the issues so they can decide for themselves. Libraries promote that only free and open Internet access can address both the First Amendment rights of youth and the right of the parent to guide his or her children.

Filters are also believed to create a false sense of security for people using them, or for parents whose children are using the filtered computers. Looking back at our accuracy ratings for Internet filters (83 percent for text, 46 percent for images), one might imagine that the number of sites or images that would be harmful to minors that still make it through the filters would give proponents pause. For some parents and guardians, placing a filter on a computer is like an announcement saying, “Hey, we have free professionally supervised daycare at the library—just plunk the kid down at the computer!” The parent is left with a sense that his or her child will definitely not encounter any unwanted material on the computer, and the parent may not even consider what unobjectionable material that the child needs might get blocked. Thus, installing filters might be associated with even more exposure to harmful material because the parents will give the children free rein and latitude to use the computers—which still have access to this harmful material. This false sense of security is of great concern to librarians and can be combated only with parent and library user education about what the filters can and cannot do.

There has been little emphasis on education of library users about Internet content, safe searching, and filter capabilities. Most libraries that do not filter have some sort of Internet safety classes for customers (usually geared toward parents), as well as literature and webpages devoted to educating customers about the reality of safety online and access to adult and other unwanted materials.

Numerous panels assigned to explore the issue of obscene Web content and Internet filters have returned with recommendations for additional education of citizens as an alternative or addition to the filtering technology itself. The National Telecommunications and Information Administration recommended in its 2003 study of Internet filtering technology that “technology protection measures are most effective when teachers and educational institutions can customize technology and use it in connection with other strategies and tools.” The NTIA also recommended “new legislative language that would clarify CIPA’s existing ‘technology protection measure’ language to ensure that technology protection measures include more than just blocking and filtering technology.”

This education about other strategies and tools and modification to CIPA language have never happened, and we still have only the CIPA language mandating filters as the sole protection measure against accidental child exposure to harmful materials.

Libraries and the American Library Association openly share their Internet use policies, usually on the library’s website. ALA and other organizations have chimed in with recommendations on what makes a good Internet use policy. ALA’s “Libraries and the Internet Toolkit” is an excellent place to start if your library needs to write or update your own Internet use policy. CIPA itself also outlines several required elements for the Acceptable Use Policies (usually called “Internet Access Policies” by libraries) that are mandated in the law.

- The policy must have offered the opportunity for public input.
- The policy must state the use of filters.
- It must offer ways to monitor student use of the Internet.
- It must provide for a way to block or filter visual depictions of material that is obscene.
- It must discuss safety and security principles for minors with electronic communications.
- It must discuss responses to access by minors to inappropriate sites, and it must discuss responses to hacking or other unauthorized workarounds to the software.

**ALA's Libraries and the Internet Toolkit**

www.ala.org/ala/aboutala/offices/oif/ftoolkits/litoolkit/default.cfm

One additional key issue to consider is that the policy is not just about the Internet—it’s about the use of your public computers, and you may also choose to include the use of your public network as well, including all of those laptop users. If so, the policy’s name and context needs to change in accordance with its expanded reach, perhaps to the “Library Computer, Network, and Internet Use Policy.”

Libraries, whether they are currently filtering or not, may also wish to conduct a cost-benefit analysis comparing the cost of Internet filters to the funding they receive in exchange through ERATE, LSTA, or other federal grants. The money received for filtering is fairly straightforward in nature. Figuring out the true costs of filtering requires...
a bit of math ability. We must look at not only the cost of the filtering software itself, but the cost of support and maintenance, any server or network slowdown cost, the staff time it takes to be trained plus any staff time spent unblocking or blocking sites, IT staff time to maintain the system if necessary, the cost in staff time and marketing about the library’s policy, and any other costs you might encounter. Beyond the straight math, we must also consider that filters can cause the library customers to lose access to information, lose time as they work around the filters, and lose confidence in the library’s relevance and ability to meet their needs. Those costs are incalculable.

Most libraries discover that they actually lose a substantial amount of money when they choose to install filters. Government commissions often see ERATE as free money and do not see the hidden implementation costs it takes to comply with ERATE requirements. The San José Public Library had $35,000 to gain in ERATE funding by implementing filters. Estimated start-up costs for the filtering software technology, staff training, hardware, and software totaled $400,000 per year with ongoing annual costs of $275,000–$300,000. Therefore, for our library, filtering for the purposes of ERATE funding was not a net-profitable situation by any means.16

As all librarians know, libraries that are well-funded have always provided better collections to their users. The Internet, however, provides an opportunity to level that playing field. This opportunity is stymied by CIPA because installing filters requires libraries to pay money to, in fact, reduce their collection’s size by limiting access to the Internet’s resources. Poor communities can either turn down ERATE funding and preserve the size of their Internet “collection” through unfettered access, or they can accept funding and spend money on filters to reduce access. Sadly, the poorer the community, the more it has to lose as ERATE discounts increase in proportion to the library’s financial challenges. Additionally, poorer communities fear lawsuits even more and sometimes choose to filter simply in order to avoid the possibility of litigation from religious and other special interest groups. As a result, it is often our most economically disadvantaged communities that find themselves filtering out of a monopoly situation. It is also these communities that would most benefit from unfettered access to the Internet to help level that intellectual and societal playing field.

Conclusion

Intellectual freedom advocates believe that Internet filtering is censorship. Proponents for filters believe that Internet filtering protects our children. I encourage both sides to examine the data on filters’ effectiveness. I believe that such data analysis will change the debate entirely, from a philosophical debate to one of technological capabilities and the costs that are incurred with imperfect technologies. Sadly, the technology has not caught up with our expectations for how it should work. Until it does, the debate about Internet filtering in libraries needs to change from one purely about philosophical principles to one also including the hard data demonstrating these filters’ serious flaws. Providing access to information, a library’s primary goal, cannot be accomplished through draconian governmental regulation over libraries restricting access. Instead, librarians, parents, and thoughtful individuals everywhere in our communities should work together to find ways to educate, prepare, and support our community members as digital citizens.

Notes

6. ALA, Library Bill of Rights.
Resources for Further Information


14. Ibid.


16. Light and SJPLibrary, “Policy Options and Staff Report.”


