Name of the commenter: Seth Finkelstein

Proposed class or classes of copyrighted work(s) to be exempted

Compilations consisting of lists of websites blocked by censorware ("filtering software") applications.

Brief summary of the argument(s) in support of the exemption proposed above

Discovering what is truly banned by censorware has been a matter of public debate. Such evidence has played an important role in litigation such as the _Mainstream Loudoun v. Loudoun County Library_ library censorware case, or the Children's Internet Protection Act (CIPA) case. Studies of censorware blacklists are vastly hindered by not being able to access those blacklists. In particular, studying structural, architectural issues, such as "loophole" sites, requires access to the decrypted blacklist.

Full version:

To the Librarian of Congress:

Pursuant to the rulemaking proceeding mandated by the Digital Millennium Copyright Act, which provides that the Librarian of Congress may exempt certain classes of works from the prohibition against circumvention of technological measures that control access to copyrighted works, the following is my comment arguing for a DMCA exemption regarding circumvention accessing censorware blacklists.

In order to lay out the argument as required, I will quote the relevant passages from the notice of inquiry, and then supply the appropriate material. The argument proceeds first in outline and then to detail.

Outline

First, a proponent must identify the technological measure that is the ultimate source of the alleged problem, and the technological measure must effectively control access to a copyrighted work.

The "technological measure that is the ultimate source of the alleged problem" here, is the encryption or scrambling which serves to keep secret the censorware blacklists. The encryption or scrambling effectively control access to the blacklists of websites blocked by censorware ("filtering software") applications, by requiring in the ordinary course of the censorware program's operation, a decryption process in order to gain access to the censorware blacklist.

The basic operation of censorware was described by the district court in the CIPA decision.

> [Censorware] programs function in a fairly simple way. When an Internet user requests access to a certain Web site or page, either by entering a domain name or IP address into a Web browser, or by clicking on a link, the [censorware] checks that domain name or IP address against a previously compiled "control list" that may contain up to hundreds of thousands of URLs. The three companies deposed in this case have control lists containing between

> 200,000 and 600,000 URLs. These lists determine which URLs will be blocked.

These extensive blacklists are transmitted from the censorware company in an encrypted or scrambled format and stored on the receiving machines. The censorware program itself loads the blacklist from an encrypted or scrambled file, and then uses the blacklist internally for the checking process described above. In virtually all commercial censorware programs, the blacklist is considered proprietary information and is never viewable by the user.

Second, a proponent must specifically explain what noninfringing activity the prohibition on circumvention is preventing.

The prohibition on circumvention prevents a wide range of noninfringing fair−use activities in criticism, comment, news reporting, teaching, about the contents of the censorware blacklists.

Note this is not an assertion that *every possible* criticism, comment, news reporting, teaching, about the contents of the censorware blacklists is prevented. There certainly are some such activities which are not affected by the prohibition. However, as will be detailed extensively below, many investigations are inhibited, and at least one investigator (me!) has been chilled at times.

As I explain in more detail later, while it's possible to test some sites without examining the decrypted blacklist, investigating other aspects of the blacklist, such as finding secret categories, require the blacklist plaintext.

In order to illustrate specific noninfringing, fair−use activity, let me recount some examples of criticism, comment, news reporting, teaching, about the contents of the censorware blacklists which were particularly related to decryption of these blacklists.

News reporting : In 1996, I was the then−anonymous source behind the very first expose of what censorware in fact truly banned, the CyberWire Dispatch "Keys to the Kingdom" .

Teaching: In 1997, I coFounded Censorware Project, and was the source for the decryption−based reports on various products (I'm no longer a member of Censorware Project, ironically stemming at heart from problems of legal risk mostly due to the DMCA).

Criticism: To quote attorney James Tyre's account, discussing the library censorware case of Mainstream Loudoun v. Loudoun County Library :

> I have mentioned the lawsuit against the Loudoun County Public Library, and the filing of that lawsuit itself, let alone the favorable result is, in many ways, perhaps the most tangible evidence of Seth's good works. In September 1997, in direct response to a plea for help from a member of Mainstream Loudoun, the group which would become the Plaintiff in the lawsuit, Seth decrypted X−Stop, the censorware which the Loudoun County Library was about to commence using. He and I analyzed the results, found a plethora of "bad" blocks, and Jonathan Wallace of The Ethical Spectacle (who also became a founder of CWP) wrote a devastating article, "The X−Stop Files", about the results. The article is on the Net at
>
> http://www.spectacle.org/cs/xstop.html
>
> Once the lawsuit was filed, we (by then, CWP had been formed) continued to feed new evidence of bad blocks to the attorneys handling the case, all of that evidence coming from

> Seth's repeated decrypts of updates of the X−Stop blacklist. I cannot know for certain if the lawsuit ever would have been filed in the absence of Seth's decryption of X−Stop, or if it would have been won in the absence of the continuing work, but I do know how appreciative those on the inside were of these efforts.

Note also the EFF Statement −− Dec. 22, 2000 on mandatory library censorware :

> Seth Finkelstein, the programmer principally responsible for the investigation of X−Stop filtering software and its flaws, vital to the landmark Mainstream Loudoun victory, ...

Comment: In 2001, my extensive decryption work was finally recognized and honored by my winning an EFF Pioneer Award :

> Seth Finkelstein − Anti−censorship activist and programmer Seth Finkelstein spent hundreds of unpaid and uncredited hours over several years to decrypt and expose to public scrutiny the secret contents of the most popular censorware blacklists. Seth has been active in raising the level of public awareness about the dangers that Internet content blocking software and rating/labeling schemes pose to freedom of communication. His work has armed many with information of great assistance in the fight against government mandated use of these systems.

All of the above noninfringing activity would have been prevented by a prohibition on circumvention to access censorware blacklists. The passage of the DMCA in 1998 in fact put a halt to the work described above. Granted, although section 1201(a)(1) did not go into effect immediately, the complexity of the then−new law, and apparent skyrocketing legal risk, made it untenable for me to continue such investigations. It was only after the Library of Congress explicitly granted an exemption that it seemed even possible to continue such decryption−based work.

In the statutory areas discussion below (see (iii) the impact ...), I detail more recent work which would be killed if the censorware exemption is not renewed, as part of the negative impact which would result.

Third, a proponent must establish that the prevented activity is, in fact, a noninfringing use under current law.

The investigations I have done, go to the heart of fair−use criticism, comment, news reporting, teaching, etc. Let's note the four−factor test for fair−use, from section 107 of title 17, copyrights:

> 1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;

All my work against censorware has been for a nonprofit educational purpose.

> (2) the nature of the copyrighted work;

Censorware blacklists are only subject to "thin" copyright, in terms of selection and categorization. They are factual compilations, not works of literature or performances.

> (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and

Quoting a few examples is a very small amount in relation to the whole list.

> (4) the effect of the use upon the potential market for or value of the copyrighted work.

As noted below, "biting criticism" is not "a harm cognizable under the Copyright Act."

Statutory areas and further details

It's stated that the nature of the Librarian's inquiry is further delineated by the statutory areas to be examined:

(i) the availability for use of copyrighted works;

Virtually all censorware blacklists are only available for use in encrypted or scrambled form. This is not a situation where the work is available in a variety of formats, such as DVD vs. VHS, or eBook vs. paper. There is one format, an encrypted or scrambled file.

However, to pre−emptively address any argument by a censorware company that the DMCA is necessary for their purposes, it should be noted that many censorware products were being marketed years before the DMCA was passed. For example, CyberPatrol was marketed in 1995. The DMCA circumvention prohibitions were clearly not a prerequisite for censorware.

Moreover, the existence of the DMCA circumvention exemption for censorware blacklists has been in effect for years, and the censorware companies do not seem to have suffered from it. I would ask that any claims to the contrary by censorware companies, of negative effects, be held to the same standards here. Namely, that here, the censorware companies "must show that such problems are or are likely to become of such significance that they would constitute a substantial adverse effect", and "conjecture alone would be insufficient to support a finding of "likely" adverse effect", and so on.

(ii) the availability for use of works for nonprofit archival, preservation, and educational purposes;

A secret blacklist file can of course be archived in encrypted form. But after a while, it is very difficult to use such an archived file even in the corresponding censorware program. Some programs won't even load encrypted blacklists which are too old. Others try to immediately download the latest version. Program versions change. Encryption algorithms change. It quickly becomes impractical to use an old encrypted blacklist. It's a strong example of the "dead media" problem (where formats can no longer be read), as a "dead algorithm" problem.

Note that decrypting a censorware blacklist is sometimes not a matter of having the encryption algorithm explicitly. There are some indirect methods which rely on having a functioning program. Such methods will not work when the program itself can no longer be run in practice.

As censorware issues become more debated and figure in more legal actions (e.g. government mandates, such as the CIPA law), there might be a longitudinal, time−based, study on how a blacklist changes overall (perhaps addressing such issues as what percentage of the list is no−longer−existing URLs, how the overall size of the list changes, what is the rate of additions and deletions, and so on). Without an exemption, building an archive of blacklists for such "archival, preservation, and educational purposes" would be admitting to a count (that's "count" in the sense of indictments) of circumvention for each blacklist file.

It's noteworthy here to observe that the one OTHER exemption granted by the Librarian was for "Literary works, including computer programs and databases, protected by access control mechanisms that fail to permit access because of malfunction, damage or obsoleteness.". Note especially "databases" and "obsoleteness". The censorware blacklist is a database, and the aging of the relevant software into old versions, no longer supported by the maker, can raise issues of obsoleteness.

The Librarian recognized this all as a "genuine problem that the market has not adequately addressed" (and this existence of a "genuine problem" would remain true even if for statutory reasons, the Librarian found that future exemptions could not be justified due to the definition issues in "class of works"). Thus, because of that problem, it would be useful to have the censorware blacklist available in decrypted form for "nonprofit archival, preservation, and educational purposes", BEFORE the relevant censorware blacklist became practically unreadable in encrypted form due to obsoleteness.

Though I refer to the "obsoleteness" etc. exemption, this justification is independent. It's in fact a kind of pre−obsoleteness requirement. The Librarian has cautioned that the "obsoleteness" etc. exemption may not necessary be renewed in triennial rulemaking. And again, the archiving of censorware blacklists in a human−readable form often needs to be done before the associated censorware program becomes obsolete.

(iii) the impact that the prohibition on the circumvention of technological measures applied to copyrighted works has on criticism, comment, news reporting, teaching, scholarship, or research;

Let me begin by noting that there is now certainly one line of censorware investigation which does not rely on decryption. It was first extensively mined by Bennett Haselton, who discovered that with the rise in virtual−hosting (many websites sharing one Internet address), many censorware programs became so inaccurate that one could find wrongly banned sites by in essence throwing darts at the blacklist. This was later refined and put to great good use by Ben Edelman in expert−witness testimony for the CIPA trial, with the realization that even better results for legal−evidence purposes could be obtained by using not random, but only high−value sites, as the darts.

But the fact that *some* types of blacklist investigation can still be done without decryption, is no basis to justify the prohibition of deeper, more in−depth, blacklist investigations. I'm more interested in focusing on what I think of as intrinsic, structural, *architectural* issues, where it's not easy to make the excuse of fixed−in−next−release.

Perhaps the best example of this type, is from an investigation I performed of the censorware program BESS (made by N2H2), This resulted in a report called
BESS's Secret LOOPHOLE (censorware vs. privacy & anonymity)
My report exposed what was, at the time, a then−secret (or more politely, "undocumented") category in this censorware. This was a category called "LOOPHOLE", which contained sites which provide services of anonymity, privacy, language translation, humorous text transformations, even web page feature testing, and more. These sites were not pornography or any type of prohibited content. All these sites were banned, at all times, because they functioned as a means of escape from the control of censorware (even the Google cache was prohibited!). Again, this "LOOPHOLE" category *cannot* be deactivated or disabled, and indeed, from the perspective of the needs of censorware to control the reader, such sites constitute threats to the necessary blinder−box. This report of mine was cited in expert−witness testimony for the CIPA trial, and my overall work in this area seems to have been a factor in the District Court CIPA decision striking down the law as applied to libraries, e.g.:

> As noted above, filtering companies often block loophole sites, such as caches, anonymizers, and translation sites. The practice of blocking loophole sites necessarily results in a significant amount of overblocking, because the vast majority of the pages that are cached, for example, do not contain content that would match a filtering company's category definitions. Filters that do not block these loophole sites, however, may enable users to access any URL on the Web via the loophole site, thus resulting in substantial underblocking.

But I could not have exposed this secret category, and seen that it could not ever be deactivated, as a structural, architectural, "feature" of the program, without examining the raw blacklist itself. There would be no other way of obtaining that hidden information.

To stress this, repeating myself, remember, the Librarian asks for

> *concrete examples or cases of specific instances in which the prohibition on circumvention of technological measures controlling access has had or is likely to have an adverse effect on noninfringing uses.*

Here is a very concrete example − if I could not have circumvented the technological measure (encryption) controlling access to the N2H2/BESS blacklist, I would not have discovered the secret LOOPHOLE category. This would likely have had an adverse effect on the noninfringing use of my critical examination of the structure of censorware, and have been one less factor for the District Court to consider in its CIPA decision.

With regard to the request

> *It would also be useful for the commenter to quantify the adverse effects in order to explain the scope of the problem, e.g., evidence of widespread or substantial impact through data or supplementary material.*

It's hard to quantify what−might−have−been, what are the effects if a report isn't done ("For want of nail, the shoe was lost; for want of shoe, the horse was lost; for want of horse, the rider was lost; and for want of rider, the war was lost."). I can say that the censorware companies have literally hundreds of thousands of dollars to promote their wares, see e.g. Austin Business Journal, September 25, 2001

> The Austin office of Ignition Strategic Communications will serve as the agency for record for Scotts Valley, Calif.−based SurfControl Inc.
>
> The annual retainer for the Ignition account is $350,000.

When censorware companies have such large sums to promote their products (including to libraries and schools), but critics have to worry about lawsuits which could *cost* them large sums to defend doing research for deep criticism of censorware, this would seem to have a substantial impact on the public debate.

(iv) the effect of circumvention of technological measures on the market for or value of copyrighted works;

Perhaps the strongest evidence for the record comes from a censorware company's own words to its shareholders. It was clearly stated, in an N2H2 annual report, regarding this DMCA exemption, that "N2H2 does not believe that the final rule will affect the value of its lists of blocked Web sites": (emphasis added)

> "The U.S. Copyright Office issued a final rule interpreting the provisions of the Digital

> Millennium Copyright Act, or DMCA, that prohibits the circumvention of technological copyright protection mechanisms. The final rule took effect on October 28, 2000, and created an exemption to the DMCA anti−circumvention provisions for compilations consisting of lists of Web sites blocked by filtering software applications. The consequence of the final rule is that lists of Web sites blocked by filtering software do not receive extra protection under the DMCA, and technological measures used to prevent access to such lists may be circumvented without violating the new "anti−circumvention" provisions Copyright Act.
>
> N2H2 considers its lists of blocked Web sites to be proprietary, valuable information. However, ***N2H2 does not believe that the final rule will affect the value of its lists of blocked Web sites***. N2H2 regards its lists as trade secrets and protects their confidentiality primarily through physical security controls and contractual non−disclosure provisions. The final rule simply exempts lists of blocked Web sites from the new copyright law protections available under the DMCA, so that the level of copyright law protection available for such lists is the same as it was before the DMCA was enacted in 1998."

From another perspective, it's important to keep in mind that negative and unfavorable results of research and commentary are not necessarily the kind of effects from which a work should be shielded. For example, in a recent Supreme Court case involving parody as fair use (Campbell v. Acuff−Rose), when discussing the fair use factor for "the effect of the use upon the potential market for or value of the copyrighted work." the Supreme Court particularly distinguished between "biting criticism that merely suppresses demand and copyright infringement, which usurps it.". This supports a finding that the criticism a censorware company may endure for being exposed as having ridiculous bans, ludicrous overbroadness, many obsolete items, and similar, does not constitute "a harm cognizable under the Copyright Act."

and (v) such other factors as the Librarian considers appropriate.

While the prior discussion has been more distant, let me take this area to try to put a human face on this endeavor.

There is a famous quote about lawsuits:

> "I must say that, as a litigant, I should dread a lawsuit beyond almost anything short of sickness and death."
>
> −− Judge Learned Hand, from "The Deficiencies of Trials to Reach the Heart of the Matter", in 3 "Lectures On Legal Topics" 89, 105 (1926), quoted in Fred R. Shapiro, "The Oxford Dictionary Of American Legal Quotations" 304 (1993).

The personal legal risks for this censorware research work are *huge*. *Enormous*. *Staggering*.

I don't get paid for this. It isn't my job. And if it were my job, (*pace* Dmitry Sklyarov), I'd be squarely within the DMCA section 1204(a) provision for criminal liability!

That section states (my emphasis)

> Any person who violates section 1201 or 1202 willfully and ***for purposes of commercial advantage or private financial gain*** − (1) shall be fined not more than $500,000 or imprisoned for not more than ***5 years***, or both, ***for the first offense***; ...

Note all violations of section 1201 are included, which presumably included 1201(a)(1)(A) circumvention violations.

With regard to the element of *"purposes of commercial advantage or private financial gain"*, I've made no secret myself that, while I don't do this for the (so far, nonexistent) money, I'd certainly welcome being paid. I noted with admiration the sentence in Ben Edelman's expert witness report, where he stated "I received $−−− per hour for my work." (the exact amount was redacted − I noted with even more admiration that it was three figures which were redacted) Yet this simple human wish, to be paid for what is at heart, hard, demanding, highly skilled professional work, in the absence of an exemption, would seem to expose me to, again, CRIMINAL liability.

The government's sample language for jury instructions, regarding this element, are that:

> "... the Government need not prove that the defendant actually received a profit from the infringement. The Government need only show that the defendant acted with the hope or expectation of some commercial advantage or financial gain."

After all, the Department Of Justice's own Criminal Resource Manual quite clearly states that:

> Emphasis should be placed on the word "purpose," because it is not necessary to prove that any profit was realized.

It stresses with emphasis later that the key word in the requirement is "*purpose*" (emphasis in original):

> Evidence of discrete monetary transactions (i.e., the selling of infringing goods for a particular price) provides the clearest evidence of financial gain, but such direct evidence should not be a prerequisite to prosecution. Such a stringent requirement would ignore the plain wording of the statute, which requires only the showing of commercial or financial *purpose*.

The overall requirement seems to be construed broadly (bartering, being an employee, etc.), encompassing "expressed or implied intent of the parties".

Note this aspect of being an employee applies very directly to the Sklyarov case. While that wasn't a 1201(a)(1)(A) case, it's still relevant in that it shows the reach of DMCA criminal provisions (which remember, *do* apply to mere circumvention in 1201(a)(1)(A)), and that the DMCA criminal provisions are by no means reserved, either in theory or in practice, for big−time infringement businesses.

This isn't a situation where the worst that can happen is that a plaintiff company tries to get an injunction (which boils down to a court saying "Stop doing that"), and the defendant can bask in the glory of being a cause−celebre, while civil−liberties lawyers and industry lawyers fight it out. Rather, it just doesn't take much to be faced with, as the worst than can happen, going to prison. Dmitry Sklyarov was jailed even before trial. While it can be argued that his situation was unusual in terms of being a foreign national, it still makes very clear that immediate substantial imprisonment is a very real possibility under DMCA provisions, and again, it just doesn't take much to be risking jail.

As I write this, Jon Johansen, one of the programmers who wrote DeCSS (but not, contrary to myth, responsible for the decryption algorithm), is facing *criminal* charges for his work. He's accused of

> "breaking a protective device or in a similar way unlawfully obtaining access to data or software"

While, since he's Norwegian, this is not the DMCA, the charges certainly are arguably analogous to a DMCA section 1201(a)(1)(A) access circumvention violation (as opposed to DMCA section 1201(a)(2) or 1201(b) tool−making or trafficking violations).

Moreover, censorware companies can be extremely legally aggressive in suing programmers. The most famous lawsuit, Microsystems. v. Scandinavia Online (the CyberPatrol case) involved extensive charges (though not based on the DMCA) against the programmers. Note the lawsuit started off asking $75,000 in damages. Matthew Skala, one of the programmers who was sued in that case, has written one of my favorite passages concerning legal defense: (emphasis added)

> "Of course I was disappointed by this state of affairs. When we published the essay I didn't expect a lawsuit, but I had also thought, "Well, if there is a lawsuit it won't be a problem, because there are organizations that take care of things like that." I fondly imagined that in case of legal silliness, someone would just step in and say "We'll take it from here." What I found out was that those organizations, through no fault of their own, were able to give me a lot of sympathy and not enough of anything else, *particularly money*, to bring *my personal risk of tragic consequences* down to an acceptable level, despite, incredibly, the fact that what I had done was legal. Ultimately, I couldn't rely on anybody to deal with my problems but myself.
>
> Some people learn that lesson a bit less impressively than I had to."

The censorware company N2H2 is another example of extraordinary legal aggressiveness. During the CIPA trial, it sought to prevent even expert−witnesses testimony from being public on "trade secret" grounds − "They say that certain things we talk about them having blocked will show the nature of their software, ..."

If N2H2 is willing to take such legal action involving court−approved experts serving as witnesses in a Federal trial, the risk for a mere programmer (even a civil−liberties award−winning programmer) is terrifying.

There's current pre−emptive litigation (Edelman vs. N2H2) against N2H2, seeking declaratory judgment concerning various censorware decryption legal issues, including but not limited to, DMCA violations. N2H2 has stated in an form 10−Q

> We intend to defend the validity of our license agreement and to enforce the provisions of this agreement to protect our proprietary rights. We also intend to assert all of our legal rights against Mr. Edelman if he engages in future activity that violates the agreement or our proprietary rights.

There's a great quote by Ben Edelman about it all: *"I don't want to go to jail. I want to go to law school."*

I don't want to go to jail either.

And I *have* circumvented the N2H2/BESS encryption.

In 1998, rising legal risks drove me to quit anti−censorware activism. It was only in the context of certain expected legal backing, and later the Library of Congress granting a DMCA exemption, that I felt it was possible for me to return to doing anti−censorware work. The renewed legal aggressiveness of censorware companies, such as N2H2's actions during the CIPA trial noted above, have almost driven me to quit again. I don't want to be playing Russian Roulette with lawsuits, where every extensive report involving decryption is taking a chance at being sued. The prospect of enduring ruinous litigation, even conceivably criminal charges, for *volunteer* work, is simply not worthwhile. ***IT'S TOO MUCH LEGAL RISK!***

Without an exemption, I will be definitely be forced to quit these investigations again.

Sincerely,

Seth Finkelstein