Reply to Copyright Office DMCA 1201 Censorware Exemption Question

June 29, 2003

Seth Finkelstein (sethf@sethf.com) and James S. Tyre (jstyre@jstyre.com)

David O. Carson
General Counsel
Copyright Office
Library of Congress
101 Independence Ave., S.E.
Washington, D.C. 20559−60000

Re: Docket No. RM 2002−4
Exemption to Prohibition on Circumvention of Copyright Protection Systems for Access Control Technologies

Dear Mr. Carson:

By emailed letter dated June 5, 2003, you asked the following question:

> Please clarify, as specifically as possible, the types of applications you believe should or should not be subject to an exception for the circumvention of access controls on filtering software lists, if such an exception is recommended.

This will be the joint response of Seth Finkelstein and James S. Tyre to that question.

In the DMCA circumvention exemption proceeding of 2000, the Librarian Of Congress granted just two exemptions. The first such exemption was:

"1) Compilations consisting of lists of websites blocked by filtering software applications;",

It was noted in part, in the rulemaking discussion, that "... the first class exempted fits comfortably within the approach to classification discussed above, ...".

The process of defining the "class of works" for these purposes is notoriously complex. So in the 2003 proceedings, Mr. Finkelstein thought it prudent to keep closely to the language which had been accepted so well earlier. After all, if it "fits comfortably", why meddle with success?

In the renewal proposal, Mr. Finkelstein introduced what he considered to be a clarification. He stated the relevant "class of works" as

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

This was in fact specifically intended to address some of the over−uses of the word "filtering", and guard against any spurious objections. The word "filtering" has too many usages, e.g.

1. censorware
2. anti−spam
3. virus−scanning
4. personalization
5. mail killfiles

As testified (April 11, page 53, line 10):

> MR. FINKELSTEIN: Oh, yes. I think the best public relations that the censorware companies ever did was to get the word "filter" attached to their products. When you think of a spam filter, for example, you think of something that you do not want to see.
>
> But, again, as I said earlier, censorware is not like a spam filter. What censorware is, is an authority wants to prevent a subject under their control from viewing material that the authority has forbidden to them. This description is general.

A proposed definition of "censorware" is:

> "Programs *designed* and *optimized* for use by an authority to prevent another person from sending or receiving information."

Some examples of such programs are N2H2's "BESS" or "Sentian", Websense, SmartFilter, SurfControl (successor to CyberPatrol), CyberSitter, and so on.

In the United States, this issue of control by authorities is most popularly framed as parents over children for sexual material. But there is nothing to the definition of censorware which restricts it in that way. Recently, in the United States, much legislation has raised the issue of control of government over citizens with regard to public libraries. Often the exact same products are sold to totalitarian governments for use against their citizens. See, for example:

"Companies Compete to Provide Saudi Internet Veil"
http://www.websense.com/company/news/companynews/01/111901.cfm

"Internet censorship goes to the Wall"
http://www.websense.com/company/news/companynews/02/asia/092702b.cfm

For this reason, the definition offered by Mr. Metalitz in testimony is, as he himself did admit possible, inadequate:

(May 14, page 45, line 13)

> MR. METALITZ: Yes. Sure. We have put something in writing to say we think the filtering software that was covered by the evidence that's been presented here, and it's on page 13 of our joint reply comments. "Filtering software used to prevent access to Internet sites containing material deemed objectionable to children or otherwise inappropriate for some segment of the public or for display in a public setting."
>
> Now, that may not be a very good definition, and I would think that people who have the word "censorware" in their name would have probably a sharper definition of what kinds of material they're talking about. But the burden, of course, is on the proponent throughout this proceeding and this panel can't recommend an exemption unless there's evidence to support it that shows a substantial adverse impact on the availability of something, some copyrighted work or noninfringing purposes. So I would suggest that, you know, we've taken a stab at it and I'm sure Mr. Tyre can do a lot better. But we just think that whatever finding is made here ought to conform to the evidence and not extend much more broadly to get into areas that aren't covered by the evidence.

Since the exemption was granted once, and the class was specifically noted as "fits comfortably", we would thus submit there is no evidence that it is too broad.

We believe the proper way to deal with any legitimate concerns would be in the text of the rulemaking, as opposed to trying to amend the class. Especially at this point late in the proceedings.

Moreover, amending the class brings up the following problem as articulated by James Tyre:

(May 14, page 46, line 15)

> MR. TYRE: Certainly we can provide a more precise definition of censorware. I don't have one in writing in front of me, but that can be done. That's not the problem.
>
> The problem is dealing with the other aspects of what Mr. Metalitz proposes, and that these things other than what would be defined as censorware. And one of the specific reasons why that's a problem, is because there's been so much consolidation in the industry, the relevant industry segment, that it's not a surprise that you have companies such as Symantec which are offering integrated products which consist both of traditional censorware and of firewall protection, antivirus protection things of that nature.
>
> And what I'm asking for, I don't know whether I'll get it, but what I'm asking for is something from Mr. Metalitz that tells us how we deal with something like that, how we deal with an integrated product. And further, how we deal with what I would call a pure censorware company such as N2H2 not suddenly grasping onto this newly limited category and by making a few minor changes into its database, suddenly turning itself into a company that in addition to doing censorware has some minor security functions, some minor virus protection. And all of a sudden because of however this definition may work, finds itself because of imprecise wording or any other reasons no longer subject to an exemption, assuming of course that there's going to be an exemption at all.
>
> So I'm really troubled by how all of this will play out. And that's why, though I may not get my wish, I am wishing that you will put the burden on Mr. Metalitz to give us something far more concrete to consider than what has been given.

Indeed, it seems the "integrated" products are becoming more and more popular. See:

http://www.symantec.com/press/2003/n030602c.html

> CUPERTINO, Calif. − June 2, 2003 − Symantec Corp. (Nasdaq: SYMC), the world leader in Internet security, today announced Symantec Mail Security for Microsoft Exchange, a comprehensive, integrated mail security solution consisting of content filtering, spam prevention, and industry leading antivirus protection.

Now, this integration does not contradict the framework about "filtering" given above. For example, a corporation, as an authority, usually doesn't want to be sent spam or viruses, and at the same time may want to control its employees as to what they are permitted to read. Again, the issue is not whether this control over employees is ideologically proper or not, but rather, the differences in meaning over the different senses of the word "filtering".

In fact, the deep distinctions between the various applications have profound architectural implications. The following testimony is noteworthy for examination:

(May 14, page 25, line 3)

> MR. METALITZ ... We know that filtering software that may fit the description that appears in the exemption that exists now, is one of the key tools in keeping our network safe and secure. And many of those filtering software packages may include lists of websites that either are the sources of viruses or the source of spam, which is of course is a scourge that we're all having to deal with increasingly now.

It is a bit difficult to discern the concern here. It relates back to the point made in testimony that censorware blacklists aren't particularly good as source of sex sites (because they are filled with out−of−date sites, duplication, no quality assessment, and so on). Spam−source websites are as far from secret as can be. They announce their presence with thousands, millions, of notices! These sites are in fact widely publicized, as a means of attempting to shut them down. The Spamhaus Project, http://www.spamhaus.org/ , states that:

> Spamhaus tracks the Internet's worst Spammers, known Spam Gangs and Spam Support Services, and works with ISPs and Law Enforcement Agencies to identify and remove persistent spammers from the Internet.

Spammers can be looked−up there in great detail, particularly
http://www.spamhaus.org/rokso/index.lasso

> ROKSO collates information and evidence on known hard−line spam operations that have been thrown off a minimum of 3 consecutive Internet Service Providers for serious spam offenses.

Perhaps the concern was regarding anti−spam programs.

One of the best anti−spam software packages, SpamAssassin, http://spamassassin.org/ is completely open−source. All websites, data, patterns, and so on, are open for inspection and evaluation. And it's become a better system for it. It's actually possible to see why a false positive occurred. Moreover, because of this "transparency", the it's−not−on−the−list game is not possible. So the creators have an incentive to fix problems rather than possibly deny their existence. Any idea that examining spam lists will lead to increased problem with spam, is refuted by SpamAssassin's success. Examining some spam lists is likely to be far more a scourge to companies selling poor products ("snake oil") than to anti−spam efforts.

By contrast, there is no significant open−source censorware blacklist. The function of banning sites is programmatically trivial. But virtually nobody is using any freely−available blacklist. All efforts in this direction have languished. Why? Because spam is something you do not wish to receive yourself, so open−source people work on making the list accurate, as it personally affects them. Whereas censorware is a prohibition by an authority on someone else, so there's no incentive to make it accurate. The authority using it (on someone else) doesn't have the same deep incentive to make absolutely sure it's accurate.

Turning to viruses, it's important to understand that computer anti−virus programs do not, in fact, contain viruses (they are different from organic anti−virus vaccinations, which are made from dead or weakened organic viruses). For each virus, they contain a few bits of characteristic information which indicates whether the virus is present. It would be akin to raising the spectre of lists of fingerprints of various dangerous criminals. The criminals may be dangerous, but their fingerprints are extremely inert, and of no use whatsoever in planning a crime.

See, for example, the description at:
http://www.trendmicro.com.au/download/pattern_update_help.htm

> Virus Pattern File Updates
>
> The virus pattern file is used together with the scanner to detect whether or not a file is infected with a virus. It does this by comparing the "signatures" of viruses which are stored in the virus pattern file to the code inside the file being scanned. If there is a match, then that file is said to be infected.
>
> Whenever a new virus is discovered, it's "signature" is added to the virus pattern file, that's why newer pattern files are always being released.

Moreover, computer viruses are not an obscure topic. Their extensive dissection in security literature provides ample sources of examples for malicious use. Someone inclined towards hostile code−writing might start with, for example, the July 2003 issue of _Wired_, which has annotated source code and an extensive discussion of the functioning of the "Slammer" worm:
http://www.wired.com/wired/archive/11.07/slammer_pr.html

So in sum, to respond to the following point: (May 14, page 48, line 3)

> MR. CARSON: ... I assume you're not saying that there is a reason why people should be able to have access to lists of what a virus swapping software blocks? Is that true or is that of interest to you?
>
> MR. TYRE: Speaking for myself and for the Censorware Project, that is not of interest to us. Whether it would be of interest to other security researchers, I have no knowledge or comment.

While we may have no interest in investigating such topics, their existence should not be used as a back−door to create exceptions which swallow the exemption. Given the failure of Mr. Metalitz to even articulate a problem in a way that can be clearly understood, much less provide any evidence of a difficulty, there seems no reason to alter such a delicate construction as the class definition.

One additional aspect bears mention concerning the censorware blacklists. In the testimony, there was evidently some confusion between having an ordinary evaluation copy of censorware, with encrypted blacklists, and being permitted to examine the unencrypted blacklists.

(May 14, page 41, line 23)

> MS. PETERS: I asked about private agreements, and you just basically cited and said that Mr. Finkelstein basically had the list but no longer did. Is that a comment on what agreements might be reached that maybe you can get an agreement to get it once, but having continuous access is a problem?
>
> MR. TYRE: The practices vary somewhat from company−to−company. But the normal practice is that you fill out a form, you give them your information. Anytime I've ever done this, I've used truthful information, no fictitious identity. And I believe that the same is true for Seth and other people I know who have done this. ...

And later (May 14, page 67, line 5) [emphasis added]

> MR. METALITZ: ... And the testimony I heard, and I don't know that this is correct, that basically it's very easy for someone to get at least one free bite at this database *without going*

> *through decryption*. It seems to relevant to me and it indicates that perhaps means other than an exemption would help to cure whatever adverse impact you find in this area.

Which was related to: (May 14, page 40, line 21)

> MR. TYRE: ... You know that during the first hearing Seth Finkelstein did have on one or two occasions access to the N2H2 blacklist. But then N2H2 stopped letting him have it, not surprisingly, but they stopped.

To clarify what Mr. Tyre meant, N2H2 allowed Mr. Finkelstein to have an ordinary, encrypted−blacklist, 30−day−limited, evaluation version of their software on a few occasions, typically when they did not check his background. One time their automated registration process approved him, but N2H2 later checked his background and threatened to revoke his evaluation credentials ("Letting you evaluate the product would be the same as working with the opposition. I have yet to read an article that you wrote that had anything good to say about filtering."). Currently, they will not let him have even their standard evaluation software (and are extremely nasty about their refusals too!). When Mr. Finkelstein was allowed access to the evaluation versions, he had to − and did − figure out how to circumvent the encryption of the blacklist so that he could study it.

Note, since the blacklist changes, it is important to periodically re−investigate it, especially to debunk spurious claims of improvement. Remember, claims of improvement of the blacklist are not made under oath, and there is no penalty for puffery. As Mr. Tyre noted, the X−Stop censorware blacklist was decrypted and studied several times during the course of the *Mainstream Loudoun* case, as it was important to see how the list changed over time.

(May 14, page 11, line 7)

> MR. TYRE: ... David Burt makes a technically correct statement but very misleading statement in his joint reply to the effect of there's nothing in the court record to indicate that the Censorware Project in general or Seth in particular had anything to do with developing the evidence in the case. That statement is 100 percent correct and 100 percent misleading. Because what happened was Seth decrypted the list not just once, but on many, many, many different occasions because you want to see what happens as they find out about new bad blocks, whether they unblocked them, what new they've added to the blacklist, things like that. Through the Censorware Project we were analyzing the lists, we were going through the lists. We were feeding the list bad blocks to the appropriate people involved in the case.

No censorware company has *ever* allowed a critic access to the unencrypted blacklist. As Mr. Tyre stated, for a critic to request such access to the unencrypted blacklist, would be, (page 43, line 15) "the ultimate idle act".

One final note: As this reply was being written, the Supreme Court issued opinions upholding the CIPA library censorware law. That outcome will undoubtedly be portrayed by censorware companies as an endorsement of their products. But it can also be argued to make structural, architectural, investigation of censorware, even more important. In particular, the following comment of Justice Kennedy has profound implications:

> " ... if it is shown that an adult user's election to view constitutionally protected Internet material is burdened in some other substantial way, that would be the subject for an as−applied challenge, not the facial challenge made in this case."

So this may not be the final word for the Constitutional challenges. The Rulemaking documentation speaks of a burden of showing "likely" adverse effect. Given the splintered nature of the CIPA decision, where no opinion had a majority of the court, further litigation seems very likely indeed. And a future "as−applied" challenge will likely benefit from as much knowledge as possible regarding the actual properties of censorware blacklists.

Sincerely,

Seth Finkelstein and James S. Tyre