

Testimony Prepared for the COPA (*Child Online Protection Act*) Commission

Gio Wiederhold

Professor

Computer Science Dept. and Dept. of Medicine, Stanford University
www-db.stanford.edu/people/gio.html

Technology demo by

James Wang (PhD summer 2000)

As of fall 2000 at Pennsylvania State University
http://jxw.stanford.edu/cgi-bin/zwang/wipe2_show.cgi

August 2000

Thank you for the privilege of presenting some statements concerning the technology relevant to the protection of minors before this commission. I will not attempt to present the full range of technological choices, threats, and candidate solutions. Instead I will focus on two points, one being specific technology in image recognition (WIPE™) that we have developed at Stanford, and, second, some systematic setting of the issues in a producer-transmitter-consumer framework, which will illustrate requirements and barriers to technological solutions. I will actually start with the second aspect, but close with some suggestions for managing the system issues.

Dealing with the system chain

To deal effectively with any problem we have to consider complete systems and the feedback loops that enable their functioning. Applying corrections or constraints at isolated points will just cause systems that are in some sense stable, to adapt. We have to consider here the entire chain of flow from Producers via Transmitters to Consumers. In each of these categories exist a variety of groups, with different means and objectives. We often hear complaints that some technology will not solve the whole problem, but such an observation, even though nearly always true, does not mean that it cannot be applied to affect the system in some focused and desirable way. However, its effect on the overall system must always be assessed.

The producers of objectionable material can be, and are likely to be anywhere; not only where US law applies. However, physical predators are likely to be within reach of the laws being considered here. While many producers of pornography are financially motivated, and will respond to certain measures, there is also a significant remainder for whom it is a hobby or a disease. Their costs in setting up a pornographic website are small. The computer technology is cheap, and poverty and exhibitionism provides low-cost content material. Substantial cost is incurred in marketing and collecting income, but for the hobby sites marketing is not a concern. If commercial explicit websites are closed, the hobby sites will gain popularity as long as transmission and consumers remain available.

Transmission capabilities via the Internet of attractive images and videos of all kinds are improving. Most commercial producers and consumer sites in industry and higher education have broadband access (ISDN, DSL, cable, T1, ...). Today, halfway into the year 2000, it is estimated that already 1.4M homes have broadband access as well. By 2004 more than 20M homes will have such capabilities [Yankee group per Internet World, August 1, 2000]. The spread of broadband transmission capabilities to the home is, among other, motivated by the entertainment industry and is unlikely to be stopped.

Distance has become a minor factor in transmission costs, so that material can move from any where to anywhere at nearly equal cost, and intermediate nodes can easily be on countries where U.S. law does not apply, even when producers and consumers are relatively local. Costs incurred for transmission are further being reduced by caching voluminous material at the periphery of the Internet. Such technology has been initiated by Akamai and is used by major distributors of video and images such as CNN. There is no reason why objectionable material cannot be made available similarly.

The density and complexity of the Internet backbone connections, as well as the many ISP shortcuts is such that constraints in transmission paths that are seen as onerous can be bypassed.

Consumers the systems of our concern are the children, and the parents and the libraries who are paying for Internet access and their equipment. As Commissioner Telage has observed, these communities are not willing today to pay for and install restriction tools, as Netnanny and Cyberpatrol. Libraries are reasonably concerned about not limiting freedom of speech and its access. Parents are *not* motivated to impose restrictions on their children, especially when they are teenagers, because of concerns for trust and awkwardness in dealing with the issues. There may also be an unwillingness by some adults to deny access to explicit material to themselves. Furthermore, as also pointed out during these hearings, many of the tools available today are naïve and inadequate.

We have found indeed that the businesses that are marketing restriction tools do not have the income to substantially improve their products. Lack of quality feeds the cycle that inhibits their sales in the first place. It is unlikely that any of them can deal in isolation adequately with the problem. The growth of the Internet, the agility of the producers, and the complexity of dealing with all the paths and the diversity of consumer equipment makes their task harder than they are willing to admit. My final suggestions will address this issue.

The consumers are a critical link. If they do not cooperate, the chain will not be broken, as seen in

- *Illegal copying of copyrighted music*
- *Non-payment of sales taxes*
- *Drug use condoned by well-off urban and suburban households*
- *Watching Pornography - at one point ~50% of Internet use*

Tools for consumers must be made desirable, simple, reliable, and adaptable. And they must be free or very cheap, since there is the expectation that everything on the Internet is free -- once you bought your computer. This model that continues the tradition of free radio and TV reception in the United States, all supported by advertising.

Our technology

As a byproduct of combining research in privacy protection and image recognition, supported by NSF HPCC and Digital Library initiatives, James Wang and I have developed and demonstrated a fast wavelet-based image recognition technology that can identify objectionable images with a higher reliability than other, less mathematical methods. This technology, Wavelet Image Pornography Elimination or WIPE™ is available for licensing through Stanford. The distinction made by WIPE are learned automatically, by feeding into WIPE the contents of known pornographic sites versus other image libraries. WIPE can now identify objectionable pictures with 97% correct hits, while misclassifying 7% of non-objectionable images. The actual values, of course, depend on ratios of images presented, we must note that the web overall contains many more desirable images than objectionable ones. We will display a few examples of failures, since they provide a better insight into the WIPE technology than showing a thousand successes. The system is available for testing on-line, at the URL given in the heading, where you can submit your own pictures for evaluation [Wang, James Ze, Jia Li, Gio Wiederhold, and Oscar Firschein: "System for Screening Objectionable Images"; Computer Communications Journal, Vol.21 no.15, pages 1355-1360, Elsevier Science, 1998].

Since pornographic web sites contain many objectionable images WIPE, can be used to classify sites with very high confidence based on image contents alone. A practical system would combine WIPE technology with text-based analysis and good-site lists, so that, say, a museum would not be classified as an objectionable site [Wang, James Ze, Jia Li, Gio Wiederhold, Oscar Firschein: "System for Classifying Objectionable Websites"; Proceedings of the 5th International Workshop on Interactive Distributed Multimedia Systems and Telecommunication Services (IDMS'98), Thomas Plagemann and Vera Goebel (Eds.), Oslo, Norway, 113-124, Springer-Verlag LNCS 1483, September 1998].

We have used our technology also to extract words hidden in images and graphics. That approach has been applied to removing patient-identifiers from X-ray films to be released to research. We are interested in applying it also to extract terms hidden in images and banners, so that further filtering can take place.

We did learn from trying to transfer our technology into commercial tools that the current vendors of filtering software did not have sufficient income to consider augmenting their products.

A suggestion

It appears clear that the web is too dynamic and large a place for simple solutions to the problems, that COPA is addressing to be easily implemented, marketed and maintained. No single approach is or can be perfect. Various presenters have advocated a classification of the providers. that is certainly one desirable step. Two choices were given as alternatives:

1. Greenfields, namely web sites certified to be suitable for kids, perhaps tagged with a domain suffix as *green.kid*. Everything not green could be forbidden, or closely monitored by parents.
2. Redlightareas, namely web sites known to contain prurient material, perhaps tagged with a domain suffix as *red.xxx*. Everything not red would be freely accessible to kids, unless additional constraints were imposed by parents.

Unfortunately, this simple division is naïve. The majority of material on the web, namely business, entertainment, and scientific sites, would not want to consider themselves red, green, nor explicitly non-red or non-green. The volume, furthermore, of all web material is so much that no greenfield could guarantee that it covered all green material, with the converse holding for the redlightareas. They have no interest, nor the expertise to classify themselves in that manner. This means that greenfields and redlightareas can both exist, with the labeling managed by selected authorities, but that a very large gray area will remain. There are already organizations maintaining greenfields. They should gain more recognition and acceptance, and a *green.kid* label may well help. They should make clear, however to the parents, that they can only include a small portion of the web, and that parent-guided forays out of the greenfield will be important for the kids education.

Dealing with redlight areas is a level more complex. We can assume that most commercial purveyors of objectionable material would join a *red.xxx* domain, unless it seriously inhibited their business. Employers and many public places may choose to do so. I have no data about what percentage of their business derives from customers at such sites. Transmitters and ISPs may not want to restrict transmission of such material, since it is likely to represent a substantial customer base.

But we also have to deal with redlight sites that are not motivated to adopt the *.xxx* domain name, as hobby-based sites and the like. Foreign sites may also not comply.

Redlight sites outside of the *xxx* domain could be recognized largely automatically, by combining existing text based technology and novel techniques such as WIPE. But keeping a list of objectionable non-*xxx* domain sites up-to-date will exceed the capabilities of the vendors of filtering products. Here a consortium, set up to publish such a list, seems to be a solution. It would run the spiders, crawlers, and analyzers, provide some human monitoring prior to moving sites into the list. It would never need to crawl the *red.xxx* nor the *green.kid* domain sites. As a consortium it should be exempt from anti-trust concerns, and also from any particular ideology. Provision for appealing mistakes will need to exist, even integrated technology supervised by well-meaning individuals will not be perfect.

Companies providing filtering software could now adopt or select from that published list. They would not compete on content, but on ease of use, installation, parental adaptability, and the like.

There might be legal means to encourage redlight sites to join the *xxx* domain classification, but again that is an issue that I cannot address. .

In closing

No single solution will be adequate, but a combination of technologies inserted at a public place can do much to provide the information needed to protect children from undesirable influences. However the technologies have to be applied within a well-understood producer-transmitter-consumer model to be effective and economic for all participants, and avoid easy workarounds. There is a cost to breaking the chain through which undesirable influences are exerted, and there must be willingness by parents and other participants to bear that cost, both financially and socially.