Testimony July 21, 2000

David Biek
Main Library Manager
Tacoma Public Library

Internet Use at the Tacoma Public Library: Our Findings and Experience

Background

Very few data exist that describe the people who use Internet services in public libraries.

Although controversies have reigned over Internet services in public libraries, especially with respect to children and pornography on the Web, the only demographic information about Web users that does exist is has been derived from user surveys and anecdotal accounts. The key question – To what extent do children in public libraries find pornography on the Internet? – cannot be answered from without actual user data.

My aim in testifying before the Commission on Child Online Protection is to share the data that we have collected at the Tacoma Public Library that bears upon this question.

One reason we wrote our own Web browser, which we call Webfoot, was to be able to design exactly the reporting system we wanted. As a consequence, we are able to collect statistics that no commercial browser software can deliver. The Appendix explains in more detail our decision to write our own browser.

This paper offers data collected in the course of day-to-day use of the Internet in all ten branches of the Tacoma Public Library, covering the period from October 1, 1999, to June 30, 2000. The data exist because each Internet user during that period, as a part of the login process, was required to enter his or her library card number as a password.

When a Tacoma Public Library card is used, whether to check out materials or to sign on to the Internet, the automated system logs certain non-identifying demographic information that is a part of each library card record: year of birth, gender, and census tract of residence. (No information is released in violation of the Washington statute protecting library patron confidentiality. The Library validates the patron data once each year to ensure that address or other changes of information are recorded.) Data pertaining to the Internet session itself is also logged, including the terminal number, the branch location, the time of day the

session began, the length of the session, and characteristics of the session itself, including pages loaded, pages failed, and URLs entered.

We can compare data from the United States Census, the Tacoma Public Library circulation system, and the Webfoot Internet browser to analyze how Internet use might match or might differ from other library uses for which a library card is required. Correlating these data sources helps us answer the question, voiced often by public librarians in the Internet era: "We're busier than ever and checking out no books!" We can also compare our data, on a census tract basis, to see how representative of city residents as a whole library users and Internet users might be.

In deciding to offer public Internet access in the Library, the Board of Trustees was aware that pornography and obscenity were issues to be dealt with. The Board decided, in essence, that graphic material of the sort described in the State of Washington "harmful to minors" law was no more suitable on Library computers than they would be on billboards on city streets.

Here's the relevant passage from the Board policy:

> The Library's acquisition of Internet materials to be made available to Library patrons does not include graphic materials depicting full nudity and sexual acts which are portrayed obviously and exclusively for sensational or pornographic purposes.

The Library's full Internet policy may be found at <www.tpl.lib.wa.us/v2/using/net.htm>

The Library's implementation of filtering software is unique. **No website is blocked and all text is delivered.** The emphasis in the Board policy is upon **images** ("graphic materials"). When the CyberPatrol software detects a site on its lists in the "Sexual Activity" or "Full Nudity" categories (the only categories we implement), Webfoot takes over and offers the user a choice to connect to the site with the images inhibited from display. All the text is presented, with the image files are represented by placeholder icons. Webfoot also meets our requirement that user feedback be made as easy, comfortable, and speedy as possible, by popping up an email message box for the user should he or she wish Library staff to review the filtering of the site.

Findings

During the survey period, public Internet access was provided at 184 terminals. A total of 56,743 user sessions were recorded and almost 7,000,000 web pages were loaded. In the tables that follow, "Web Sessions" is a count of unique user sessions. "CyberPatrol Sessions" shows the count of user sessions in which at least one CyberPatrol filter intercept was encountered; there were 3,556 CyberPatrol sessions over this period, 6.2% of the total sessions. The "City Population" is the

1999 estimate and the figure for "Circulation" is the total circulation for all library materials in 1999.

Gender

	City Population	Circulation	Web Sessions	CyberPatrol Sessions
Male	48%	41%	63%	75%
Female	52%	59%	37%	25%

Most observers of Internet phenomena would say that men and boys far outnumber women and girls at the terminals. Our data support this. At the same time, females are a slight majority of the population of the city of Tacoma and a significant majority of those who check out library materials.

On the Web, the ratio of male to female Web users is reversed, and then some. Even more extreme is the preponderance of males when sessions involving CyberPatrol intercepts are involved. The Internet has, plainly, brought a new male audience in to the library.

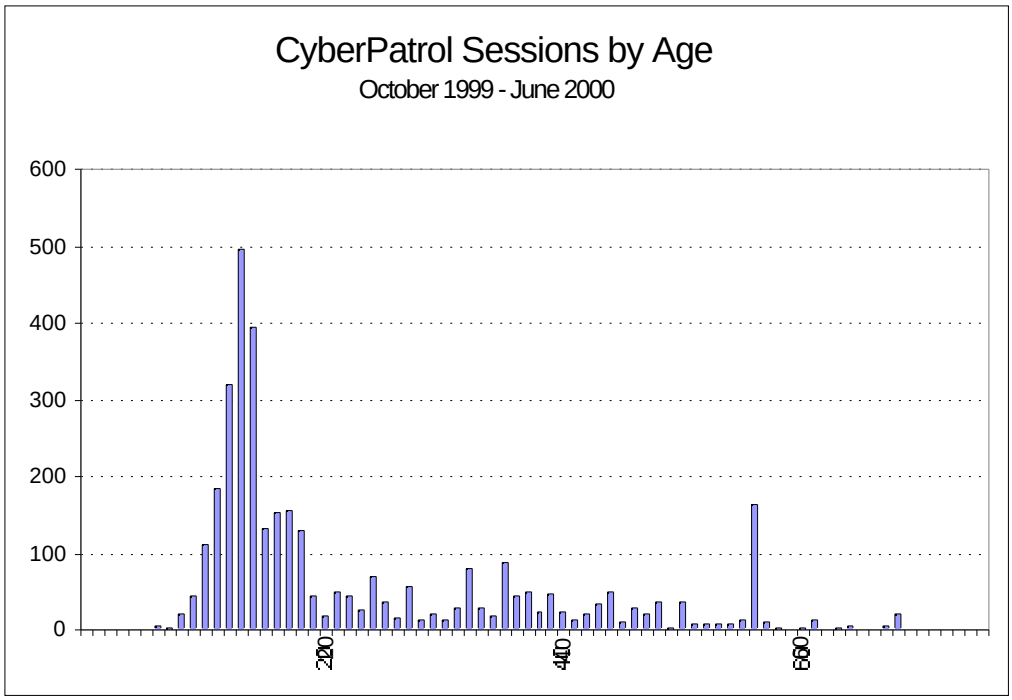
Age

	City Population	Circulation	Web Sessions	CyberPatrol Sessions
0-04	8%	0%	0%	0%
05-14	14%	17%	27%	44%
15-24	15%	16%	19%	23%
25-34	19%	17%	15%	9%
35-44	15%	21%	19%	11%
45-54	9%	15%	12%	5%
55-64	7%	7%	5%	6%
65-74	7%	4%	2%	2%
75+	7%	3%	1%	0%

Median	32	34	27	16

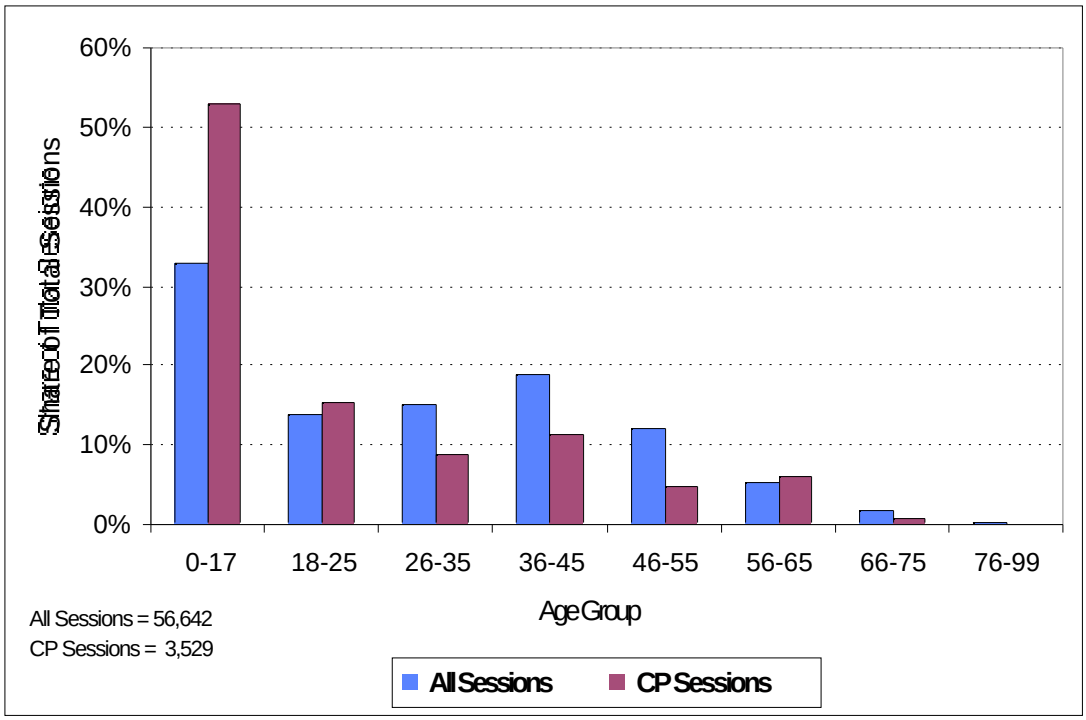
There is a reasonably good match between library card use and the population of the age group. The median ages of these two groups are also close to the same, 32 and 34 years, respectively. Web users are markedly younger, with a median age of 27. Close to half of all Web users are between 5 and 24 years of age.

The most startling finding is that users who register at least one CyberPatrol intercept during a session are so much younger than Web users in general. The median age where CyberPatrol intercepts are involved is just 16 and the single age most frequently seen is 13.



A person 17 years of age is considered to be a minor in the State of Washington, and so the age data may be aggregated to show:

	City Population	Circulation	Web Sessions	CyberPatrol Sessions
0-17	26%	23%	34%	57%
18 +	74%	77%	66%	43%



The next chart presents that data in more detail:

In recent months, reports have included cross tabulations for gender and age. The same picture is seen for boys and girls: the 10 to 17 age group for each sex accounts for the majority of filter activity.

	Male		Female	
	Sessions	CP Sessions	Sessions	CP Sessions
0-17	36%	63%	36%	70%
18-25	10%	8%	13%	7%
26-35	13%	3%	15%	6%
36-45	23%	12%	15%	8%
46-55	10%	6%	16%	7%
56-65	5%	7%	3%	2%
66-75	1%	0%	2%	0%
76-99	0%	0%	0%	0%

A major theme of the Internet policy of the Tacoma Public Library is stated terms of the protection of minors. The data show that the concern of the Board of Trustees in setting this policy was well founded.

Census Tract of Residence

The City of Tacoma comprises 39 census tracts. Our patron registration process automatically assigns a census tract to each patron record based upon the residence address. Analysis of the web activity data is suggestive but highly tentative, since we suspect that the year 2000 Census of the city will show significant demographic changes from the 1990 census.

We do think that the year 2000 information will show that neither total Web sessions nor CyberPatrol sessions are strongly correlated with census tract population numbers. Rather, we believe we will find stronger correlations between: 1) web usage and median age and educational levels, and 2) CyberPatrol intercepts and median age and income levels.

Effectiveness of the Library's Internet Implementation

One reason the Tacoma Public Library chose CyberPatrol was that the software provides an easy mechanism for system administrators to override the website addresses set by the producer. What we expected, and what we have found, is that modifications have indeed been necessary.

We implement two of the CyberPatrol categories: "Sexual Activity" and "Full Nudity." Because of the highly circumscribed requirements of our Internet policy, some sites that are correctly listed according to the producer's criteria should not be flagged on our Library terminals.

The major reason for this is that these sites do not include illustrations of a nature prohibited by Library policy. These may be sites featuring text-only erotic or pornographic stories, adult personals ads, collections of dirty jokes, sites with music lyrics that merit "parental advisory" warnings, sites featuring photographs of women in bikinis or skimpy wrestling apparel not involving nudity, and "warez" sites which include cracked triple-x passwords. These sites are added to the override "allow" list. We learn about sites such as these through user feedback and my own review of the system logs.

Recall that when a site flagged by CyberPatrol is encountered, the user is presented with an email form that can be used to request that the Library review the site. This is analogous to the process we have always offered library patrons for the reconsideration of books in the collection. The form may be submitted anonymously or the user may choose to include contact information for a reply. I review each of these requests and aim for a 24-hour turnaround in my decisions.

Over the period of time described in this study, our users have submitted 1,153 of these requests. In a typical month, 85% of them are from sites, which do indeed include image content that violates Library policy. Almost all of the rest are correctly flagged according to CyberPatrol's category definitions but fail to meet

the more strict requirements of Library policy. These are added to the override list. A few each month do not appear to me to meet the software producer's own definitions; these are also added to the override list and we take the extra step of notifying the company of the problem.

The second source for modifications is my own review of logs of filter intercepts. I check any URLs that I do not recognize as correctly flagged and take appropriate actions. Typically, 95% of all the URLs listed in these logs are properly treated by our system.

The result is that the software list of flagged sites is constantly refined and improved.

Comments on the Library's policy and procedures from members of our community have, with a bare handful of exceptions, been positive and supportive. Some of these comments have been made in connection with well-publicized incidents involving neighboring library systems and Internet pornography.

Conclusions

Month after month, we find that about 6% of all sessions at public access Internet stations at the Tacoma Public Library involve websites flagged by the filtering software. That sounds like a small number, but when the number of sessions is large (now about 7,000 per month and growing), then the number of incidents in an unfiltered environment would be intolerable.

But it could be worse. That 6% is in an environment where a filtering procedure is in place. If nothing were in place, then we would expect the number of incidents to be much higher.

We know that the majority of filter incidents involve minors, with a significant share involving 13 and 14 year olds. Our experience says that the stereotyped "dirty old man" is not looking for Internet pornography at our library.

At the same time, the unique implementation at the Tacoma Public Library ensures that users are able to access any URL on the web and that all text at a flagged website can be delivered.

User satisfaction and community acceptance are high and an efficient mechanism for user feedback is in place.

The Internet has brought a new audience to the Tacoma Public Library. With solid numbers, we can show our governing Board of Trustees how Internet services, in the context of public library service, are used. The information also helps us to craft policies and procedures to ensure that these new services are provided in an effective and responsive manner.

Appendix

Why the Tacoma Public Library Wrote Its Own Web Browser

We began investigating browser software in 1996, at a time when Lynx was the only browser available at the Main Library for the public. Lynx is a text-only tool, difficult to use and to teach.

Mosaic and Netscape were the major commercial browsers available, although Internet Explorer was by then becoming a serious contender. We found that none of these commercial products could deliver even a fraction of the features we wanted.

It's important to emphasize that Internet implementation at the Tacoma Public Library began with many discussions among staff, administration, and the Board of Trustees about the way that Web-based information fit in with the Library's overall service plan and sense of mission. Building upon that, we attempted to build the best possible hardware and software suite to implement these local decisions. The situation, of course, is fluid, and our system has evolved to meet user needs and new technical possibilities.

We had five goals that could not be achieved with a commercial browser:

- We needed a way to control printing costs.
- We wanted to gather reliable statistical information.
- We knew that a number of issues related to browser use by neophytes would need to be addressed.
- Based upon our experience in providing a computer lab in a public library setting, we knew that security and user privacy had to be ensured.
- And, we knew we had to fix obvious deficiencies with available filtering software in order to implement the Library's policy of inhibiting certain graphics from display.

In the years since 1997, when Webfoot made its debut at the Main Library, the commercial browser market has shaken out, leaving Internet Explorer dominant, with Netscape also a major player. Neither of these products, however, is any closer to offering the feature set we desired.

Printing Control

Users have two options for printing from the Web. (They can also save files to floppy discs and can send data to any valid email address.) A text-only print can be made at the ink jet printer alongside each terminal. These prints bear a statement telling the user the amount due for the print. Payment is on the honor system.

Text with images or any degree of formatting, however, is better done on a laser printer. If a user chooses the laser print, the local printer produces a small ticket that shows the price of the print (at ten cents per page) and the job number in the laser printer's queue. We print the required material only when the user has decided to pay for and take the material. We save money on printing since nothing is printed (except the ticket) without being paid for to recover costs. Users are served well, too, since they incur no costs until they are sure they do want the printed material.

Statistics

Web surfing only seems free of cost. In fact, every library that provides Internet access incurs major costs. Rarely are new, untapped sources of funds available to provide the service. Detailed usage reports are the only reliable and objective way to ensure that the Library's investment represents value to the community.

By writing our own browser, we were able to design exactly the reporting system we wanted and therefore we are able to collect statistics that no commercial browser software can deliver. For each public terminal, for each library branch, and for the system as a whole, we have counts of:

- hours of use
- user sessions
- pages loaded
- page loads failed (and the reasons for the failures)
- the occurrences of CyberPatrol flags
- various user commands issued
- downloads and printing activity

- uses of the Library's own website

Because users enter their library number of begin a session, we can collect aggregated demographic data, including age, gender, and census tract of residence.

Ease of Use

Webfoot is not an elaborate piece of software. Consistent with the Library's purpose in making the Web available, the feature set provided is the minimum necessary for efficient web-browsing. We also wanted a browser that would be easy to learn to accommodate the technophobic element of our public. For example, Webfoot has de-emphasized the need to use a mouse by providing keyboard shortcuts for many frequently performed operations.

Security and Privacy

At the time Webfoot was written, Web users in a shared environment such as a public library faced risks to their privacy. Identifying information related to a user's web session was recorded in a number of places, including the cache, history files, settings files, and cookies. With little difficulty, the following user could have access to this information. Webfoot automatically clears or deletes these files at the end of each session.

System security in a shared environment was also a concern. By 1997, the Library had ten years' experience operating a Computer Lab - eight personal computers in a network for word processing and similar uses. We knew from this experience that we could expect vandalism and mischief from hackers at the Internet stations. Webfoot includes many design features to protect the Library's networks. The "clearing" done at the end of each session also reduces security risks.

Webfoot and CyberPatrol

When we began to design Webfoot, web-filtering software had advanced beyond crude word blocking but the state of the art still left much to be desired. We were not willing to accept the only outcome that filtering software then available could deliver: that is, a site that was on the software's list of URLs was absolutely unreachable. Furthermore, mechanisms for users to request reconsideration of blocked sites were not integral parts of the filtering software systems. We also wanted to be able to override the software as needed by adding to or deleting from the list of sites without imposing on staff too heavy an administrative burden.

From our experience in a text-only Lynx environment we knew that our users were not be reluctant to seek out pornographic materials on the Web and we expected that such interest would be much greater when the graphical browser was introduced. In the new graphical environment we wanted only to inhibit the display of those particular images which were not allowed under our Board policy for Internet use.

Following close scrutiny of the software filtering products then on the market, we selected CyberPatrol. In this program, the lists of URLs are categorized so that we were able to approach our policy goals by selecting just two categories for implementation; so-called word blocking is not used. Overriding the software is easy for our system administrator to do. Finally, we saw that we would be able to write the Webfoot software so that the literal and exact purpose of our policy could be achieved.

When a user requests a site that appears on one of the two CyberPatrol category lists, Webfoot intervenes to explain that the requested site has been flagged by the filtering software and asks the user whether he wants to see a text-only version of the site. If so, then the requested site is displayed, with small placeholder icons replacing the. No text at the site is blocked.

Webfoot also meets our requirement that user feedback be made as easy, comfortable, and speedy as possible, by popping up an email message box for the user. Staff review is expedited since the email message contains the URL of the site; clicking on it in an unfiltered browser and go right to the site for the reconsideration.