Asia-Pacific Conference on Library & Information Education & Practice, 2009

FILTERING SOFTWARE IN JAPANESE PUBLIC LIBRARIES AND THEIR PERFORMANCE

KEITA TSUJI

Graduate School of Library, Information and Media Studies,University of Tsukuba
1-2 Kasuga, Tsukuba-city, Ibaraki-ken 305-8550, Japan
keita@slis.tsukuba.ac.jp

AKIKO YAHAGI

Shibaura Mechatronics Cooperation
2-5-1 Kasama, Sakae-ku, Yokohama-city, Kanagawa-ken 247-8610, Japan
kokia4155@yahoo.co.jp

Introduction: This study clarifies the current status of implementing filtering software in Japanese public libraries and their performance.
Method: 253 municipal libraries and 47 prefectural libraries that provide Internet terminals to their users were chosen as the target, and a mail survey was conducted from November 2007 to January 2008. We received usable responses from a total of 155 libraries. As for the performance of filtering software, we manually checked 4,640 web pages by using PCs on which i-FILTER and InterSafe were installed. We determined whether they blocked content that "should be blocked" and, conversely, whether they excessively blocked content that should not be blocked.
Result: We found that 85.3% of prefectural and 76.9% of municipal libraries installed filtering software. The most chosen software applications were i-FILTER, InterSafe, and CYBERsitter. Among the complaints from users, some mentioned that they could not access the pages of online bookstores, which are considered useful to library users. The survey revealed a trade-off relationship between i-FILTER and InterSafe regarding the blocking rate for harmful pages and the incorrect blocking rate for non-harmful pages. Moreover, block leaks are common for relatively new concepts such as 学校裏サイト (student run school site).
Conclusion: Filtering software is now common among public libraries, but vendors' offerings perform differently. Public libraries should select and set up their software in correspondence with their own information supply policies.

1. Introduction

The Internet has become flooded with information which is thought to be harmful to minors. However, enforcing legal restrictions on the people uploading this kind of information is problematic because doing so infringes on the freedom of expression. The use of filtering software to impose automatic restrictions on Internet access to those who are underage is growing in popularity. In 2006, the Japanese government launched the Filtering Popularization & Education Action Plan with the goal of preventing underage access to harmful information, and it is currently promoting the usage of filtering software along with mobile phone companies, filtering software companies, and computer companies. However, implementing filtering software in public libraries is problematic. Because any one Internet terminal in a typical public library can be used by both minors and adults, if filtering software is introduced across the board, adults will also experience a certain amount of limitations to information access. Though there would be no problems if filtering software could completely reflect the intentions of the library staff, in reality, filtering software is developed by companies outside the library's influence. This means that the intentions of the library staff are not directly reflected in the software design. If companies that develop filtering software are biased in their views towards information access restriction, then a form of censorship will wind up being planted in the heart of our libraries.

In light of this situation, the present study (1) clarifies the status of implementing filtering software in Japanese public libraries in terms of numerical quantity and (2) clarifies the functionality of the software investigated, i.e., is it thoroughly blocking "content that should be blocked" and, conversely, is it not excessively blocking "content that should not be blocked"? By investigating these two aspects, we wish to clarify the current state of the Internet information supply in public libraries and contemplate the appropriateness of filtering software implementation.

This study is beneficial to three parties: libraries, library users, and filtering companies. Firstly, it clarifies how filtering software enforces information access restrictions in public libraries across Japan.

This can provide information for creating future guidelines at libraries considering introducing software and libraries already implementing the software. Secondly, library users can be made aware of what kind of information can and cannot be accessed at the libraries they use. Thirdly, filtering software companies can benefit by having their products and the products of other companies objectively compared by a third party. This can in turn be used as reference information in future software development activities.

2. Preceding Studies

First, we will mention the previous research on public libraries and filtering software and then previous research on software performance. Regarding the status of filtering software implementation in public libraries in the U.S., the U.S. National Commission on Libraries and Information Science (NCLIS) has been carrying out a nationwide investigation since 1998 (Bertot & McClure (1999)). Also, Kawasaki et al. (2001) conducted an investigation together with the Intellectual Freedom Committee of the Nebraska Library Association. In contrast, Japan does claim a great deal of discussion by researchers on the topic (Kawasaki & Takakuwa (2000); Nakamura (2002a); Nakamura (2002b); Moriwaki (2003); Nawa (2004); Ide (2006)), and it does not have a significant body of investigation into actual conditions in terms of numerical quantity. We only have a scant smattering of investigations into Internet terminals and the implementation of filtering software from an architectural viewpoint (Son et al. (2003)).

Also, most of the investigations into filtering software performance have been limited to checking on leaks in blocking website pages which should be blocked (Consumers Union of the United States, Inc. (2001); Matsumoto (2005)), and there are very few examples of research that cover excessive blocking of website pages that should be accessible. The investigation of Endo (1999) is an example of research which covers both areas; however, the research was done in 1999, before the popularization of filtering software and the situation was most likely different from that of today.

3. Survey on Public Libraries

The number of municipal libraries providing Internet terminals to library users was 1,275 as of 2007. We randomly selected 253 libraries from among these and combined them with 47 prefectural libraries for a total of 300 target libraries. We then conducted a mail survey of these target libraries from November 2007 to January 2008. We received usable responses from a total of 155 libraries. Out of these, 34 were prefectural libraries and 121 were municipal libraries.

The number of libraries implementing filtering software was 29 for prefectural libraries (85.3%) and 93 (76.9%) for municipal libraries, totaling 122 libraries overall. Among these, the number of libraries with multiple Internet terminals and no filtering software in all of the terminals was 5 for prefectural libraries and 2 for municipal libraries.

3.1 Year filtering software was first installed

The year software was first installed is shown on Table 1. The entries "all terminals" and "some terminals" differentiate between libraries that installed filtering software in every terminal and libraries that did not install filtering software in every terminal. From Table 1, we can see that among the municipal libraries, at least one has implemented software since 1997, and most libraries began implementing software in 2001 and 2004. Also, we can see that prefectural libraries have installed filtering software relatively recently compared to municipal libraries (in the results of 2006 onward).

3.2 Filtering software being used

Tables 2, 3, 4, and 5 show what kind of filtering software and functions public libraries use. Firstly, there were two types of filtering software in use: (a) software installed in a server that goes on to execute filtering on multiple computers and (b) software installed in each computer. Table 2 reveals the most used type of software: server types are prevalent at both prefectural libraries and municipal libraries, amounting to nearly half of all libraries overall. We expected prefectural libraries to use server type software more than municipal libraries. However, the result was just the opposite, with municipal

libraries scoring a higher ratio for the server type of software.

The "other" category consists of libraries which answered that they use other sources such as (i) the filtering functionality of anti-virus software, (ii) gateway type UTM (Unified Threat Management), and (iii) Internet Association Japan LB (Label Bureau).

As for the specific filtering software being used in the libraries, Table 3 lists the results for the server type of software and Table 4 lists results for the individual computer type. First, we will discuss the server type. The software most used by municipal libraries is i-FILTER (Digital Arts, Inc.), scoring almost 40% of the total. For prefectural libraries, InterSafe (ALPS System Integration Co., Ltd.) is the most used, again scoring almost 40% of the total.

As for the individual computer type of software, i-filter (Digital Arts, Inc.)[1] was the most common for both prefectural and municipal libraries, scoring up to 70% overall (see in Table 4). Second place went to CYBERsitter (IQS Co., Ltd.) in municipal libraries and InterSafe Personal (ALPS System Integration Co., Ltd.) in prefectural libraries.

The server type consisted of a relatively broad spectrum of filtering software. In contrast, the individual computer type did not show much variety, and seemed to be generally dominated by i-filter.

Filtering software has a variety of functions, and the user can select from among them. We asked the functions used, and allowed multiple answers. The results are shown in Table 5. From the table, we can see that the function that blocks specifically designated URLs and the function that blocks content based on input keywords are used the most.

3.3 Reasons for installing filtering software

We asked questions about the reasons for installing software, and allowed multiple answers. The results are shown in Table 6. The most common reason was to restrict access of minors to harmful information, with municipal libraries scoring up to 79.1% in this category. The second most common reason was to make it impossible to conduct illegal actions from library terminals. As mentioned in Section 1, filtering software is expected to perform functions beyond only restricting access to information. Namely, it is also expected to restrict the actions of individuals uploading material.

3.4 Reasons for not installing filtering software

Out of the libraries which responded to the survey, those which do not use filtering software consisted of 5 prefectural libraries (14.7%) and 28 municipal libraries (23.1%). We asked the reason for not installing software, and allowed multiple answers. The results are shown in Table 7. Three out of five prefectural libraries answered that filtering software is problematic in regards to the issue of safeguarding the right to know (60%).

Regarding the 11 libraries that claimed to be restricting usage to Internet information through other means, we followed up with additional survey questions asking what the methods of restriction were. We received responses from 7 of those 11 libraries. These "other means" were mainly (1) disabling access to specific URLs by directly inputting the target URLs into the application software InforBarrier (one library), (2) restricting the use of free e-mail, etc. by restricting usage of things such as Java script and cookies (one library), and (3) appealing to the morals of users by positioning the Internet terminal layout and posting instructions about prohibited actions (two libraries).

3.5 Reception of complaints due to installing filtering software

We asked libraries if they received complaints from users due to their implementing filtering software. As a result, 54 libraries of all 122 libraries (44.3%) that use filtering software answered that they have received complaints. We followed up with additional survey questions to those libraries asking what kind of complaints were received, and received responses from 33 libraries. The majority of complaints were not from minors, but from adults instead. The most common complaint was that pages the adults wanted to see (clearly not pages containing harmful content) were blocked (31 libraries – 93.9%). Only

[1] In the present paper, "i-FILTER" and "i-filter" designate the server type and individual PC type of software.

3 libraries received complaints stating that the library was taking away the right to know of the user and only 1 library received complaints stating that if a public institution such as a library blocks information, this will lead to censorship. Thus, complaints regarding user rights were relatively low.

The blocked Web pages in the complaints included pages of NHK (a public TV station) and Nara Prefecture websites. Also, dissatisfaction with the blocking of online shopping was relatively strong. We asked what kind of response was given to such complaints, and allowed multiple answers. As a result, 24 out of 33 libraries answered that they explained the current situation and reasons for using software and strove to gain understanding from users. Nine libraries chose the answer of taking the URL for which a complaint was received off the block list.

4. Survey of Filtering Software

We conducted a performance survey targeting i-FILTER and InterSafe. There are two reasons why we chose to go with i-FILTER and InterSafe: (1) these two software packages are the most used of the server type and (2) because they are developed by the same companies that develop i-filter and InterSafe Personal, the software packages which scored as the most used in the category of the individual computer type, they should exhibit a similar filtering functionality.

4.1 Survey method

The survey was conducted by using Internet terminals in public libraries. We investigated i-FILTER at the Nakano Ward Central Library (Tokyo) and InterSafe at the Tokyo Metropolitan Library in January 2008. Below is a summary of the survey process.

(1) We selected search words which are thought to lead to Web pages with harmful content for minors under 18 when they are used in a search engine. The 29 words listed in Table 8 were selected. As for the pages harmful to minors under 18, we referred to the information on adult Websites in the rating system called SafetyOnline3, which was created by the Internet Association Japan.

(2) We entered the search words selected in step (1) into the Google Japan search engine (http://www.google.co.jp) on Internet terminals at the Nakano Ward Central Library and Tokyo Metropolitan Library and collected the URLs of the top 40 pages for each 29 words (i.e., 1,160 URLs). We used these 1,160 pages mainly to investigate the amount of blockage of harmful pages.

(3) We combined the 29 words from step (1) with the terms, "current situation," "rights and wrongs of," and "what is" and conducted a simultaneous "AND search" with two search words on Google Japan. We then collected the top 40 URLs and blocked URLs. Namely, for i-FILTER and InterSafe respectively, we gathered $29 \times 40 \times 3 = 3{,}480$ URLs. Though these 3,480 pages did include harmful material, academic reports containing the phrases "current situation," "rights and wrongs of," and, "what is" were also included. Thus, the possibility was high for these pages to not be harmful overall. In contrast to step (2), we used these 3,480 pages to mainly investigate the amount of over blocking of non-harmful pages.

(4) Lastly, we tried accessing the URLs on computers without filtering software installed. We checked for the presence or absence of harmful content for minors under 18 with our own eyes. Determination of the presence of harmful content was based on (i) SafetyOnline3 and Internet Hotline guidelines, (ii) URL domains, and (iii) file formats. However, because it was difficult to classify all of the pages into the two categories of "contains harmful content" and "does not contain harmful content," we set a grey area in between the two categories. Namely, we set the three categories of (a) pages which anyone would easily judge as being harmful to minors, (b) pages which are as easily distinguishable as being non harmful to minors, and (c) pages which contain content that could arguably fall into either of the other two categories. We will only consider categories (a) and (b) herein.

4.2 Results

The extent of pages blocked by the filtering software is shown in Table 8. To see a specific example, let's look at the case of the search word 援交 (juvenile prostitution). Among the 40 pages hit by this search word, the authors judged 37 pages as harmful. The table shows that i-FILTER blocked 70.3% of the 37 pages (i.e., 26 pages) and InterSafe blocked 73.0% (27 pages). Another example is the search word 裸 (nude). Among the 40 pages hit by this search word, the authors judged 22 pages as non-harmful. The table shows that i-FILTER incorrectly blocked 13.6% of these 22 pages and InterSafe incorrectly blocked 9.1%.

Overall, i-FILTER had a higher rate of blocking harmful pages compared with InterSafe. For example, out of the total of 199 harmful pages hit, InterSafe only blocked 80.4%, but i-FILTER blocked 87.4%. The flip side of that coin, however, was that InterSafe wound up scoring lower than i-FILTER overall for the incorrect block rate of non-harmful pages. For example, out of the 504 non-harmful pages hit, i-FILTER incorrectly blocked 16.3%, but InterSafe only incorrectly blocked 5.8%. Although we have not listed the detailed results concerning the previously mentioned AND search result, it also showed similar tendencies.

Below, we will first discuss some of the finer points of the tendencies in blocking harmful pages and then talk about incorrect blocking of non-harmful pages. First, search terms which exhibited low blocking rates for harmful pages were 闇サイト (websites that offer criminal services) and 学校裏サイト (student-run online bulletin boards for certain schools that often become arenas for bullying). 闇サイト scored as the term with the lowest block rate for both i-FILTER and InterSafe, and 学校裏サイト scored as the term with the 3rd lowest block rate in both software packages. Compared to terms pertaining to sex and crime, these terms are relatively new, which is most likely why they are not completely covered by filtering software companies. In contrast, though 援交(juvenile prostitution) is a term pertaining to sex that has existed for quite a while, it exhibited a low block rate. Among the pages hit with the search term 援交, many pages related to mobile phone dating systems. The mobile phone is relatively a new device for Internet access, which might have led to the low blocking rate by the filtering software.

Next, we will discuss some of the finer points of the tendencies in incorrect blocking of non-harmful pages. Table 9 shows representative examples of pages which were blocked even though they were non-harmful and the search terms used to hit those pages (the search terms are in parentheses). Among incorrectly blocked non-harmful pages, the following types of pages were frequently observed: (a) pages that use the non-harmful meaning of search words with multiple meanings (for example, "AV" with the meaning of "audio video" instead of the pornographic "adult video" or "吉原" as the name of a person instead of the brothel area), (b) pages of shopping websites where the product name includes the search term (Amazon.co.jp, etc.), and (c) news pages related to the search term. Particularly on the topic of shopping websites, there were no pages thought to be harmful among the shopping pages blocked by i-FILTER. Out of 96 pages judged as non-harmful, 63 were pages introducing book information for online bookstores such as Amazon.co.jp. Online bookstore pages are very useful for gaining information about books. It is better to set the blocking function not to automatically block pages in the shopping category when using i-FILTER.

Though not shown in Table 8, the correlation between "the non-harmful page count" and "the incorrect blocking rate for non-harmful pages" in the AND search was negative. The correlation coefficient was 0.6 or more. In other words, we found that when the amount of non-harmful pages rises, the incorrect block rate decreases. Thus, if search terms which usually hit non-harmful pages are used (in place of the search terms we actually used for this survey), there is a possibility that the incorrect blocking rate will decrease.

5. Conclusion

The present study surveyed the state of usage of filtering software in Japanese public libraries and the performance of the software used. Web found that approximately 80% of public libraries are currently using filtering software, and that among libraries that use filtering software, i-FILTER, i-filter, and InterSafe are commonly used software packages and specific URL blocking and keyword blocking functions are commonly used functions. Also, approximately half of the libraries using filtering

software have received some kind of complaint, with complaints about blockage of shopping sites being the most common. Online bookstores were included among the blocked shopping sites. These online bookstores are very useful to library users. Regarding this point, libraries should put more effort into understanding the functions of their filtering software and fine tune the restrictions and permissions of blocks to align with the needs of library users. Also, filtering software companies should create more detailed settings so that shopping site restrictions do not block entire websites, but instead only block the shopping forms ("checkout" pages).

The performance survey for filtering software clarified that block leaks are common for relatively new concepts such as 学校裏サイト and there is a trade-off relationship between i-FILTER and InterSafe regarding the block rate for harmful pages and incorrect block rate for non-harmful pages. Public libraries should select and set up software in a way that corresponds to their own information supply policies.

We will carry out an investigation with a larger range of search terms. We also want to conduct surveys on the state of usage of filtering software at public libraries in countries other than Japan.

References

Bertot, J. C., & McClure, C. R. (1999). The 1998 National Survey of U.S. Public Library Outlet Internet Connectivity, Final Report. Washington, DC. The American Library Association, The National Commission on Libraries and Information Science. http://www.ii.fsu.edu/~cmcclure/Finalsurvey_rev4-99.pdf.

Consumers Union of the United States, Inc. (2001). Filtering Software: Test Digital Chaperones for Kids. Consumer Reports. p. 20-23.

Endo, M. (1999). Filtering and how to judge the harmfulness of Internet information (Graduation thesis, Faculty of Letters, Keio University). 60p. (text in Japanese).

Ide, A. (2006). Filtering problem as social system. The Journal of Information Science and Technology Association, 56(10), p. 446-451. (text in Japanese).

Kawasaki, Y. et al. (2001). Internet and intellectual freedom: survey of all public libraries in Nebraska State (Nov. 2000). Kyoto University Library and Information Science Study Group. 88p. (text in Japanese).

Kawasaki, Y., & Takakuwa, Y. (2000). Libraries, Internet and intellectual freedom: philosophy and practice in U.S. public libraries. Kyoto University Library and Information Science Study Group. 207p. (text in Japanese).

Matsumoto, H. (2005). Blocking harmful website by filtering software. Nikkei Personal Computing, 492, p.30-33. (text in Japanese).

Ministry of Internal Affairs and Communications. (2007). Result of second survey of telecommunication service monitors 2006. http://www.soumu.go.jp/s-news/2007/pdf/070522_6_bt.pdf. (text in Japanese).

Moriwaki, A. (2003). Requiring libraries to install filtering and restriction of expression in Internet in future: from the example of CDA, COPA, CIPA. Osaka Law Review, 53(3/4), p.1015-1041. (text in Japanese).

Nakamura, Y. (2002a). Restriction from the First Amendment to the U.S. Constitution to the use of filtering software in public libraries and school libraries: What content can librarians eliminate?. Journal of Japan Society of Library and Information Science, 47(4), p.163-167. (text in Japanese).

Nakamura, Y. (2002b). Light and shadow of Internet terminal in libraries. http://www.avcc.or.jp/d-sigoto/d-01/05-8/d-01-05-8.html. (text in Japanese).

Nawa, K. (2004). Filtering in libraries. Information & Documentation, 47(9), p. 636-638. (text in Japanese).

Son, J. et al. (2003). Current status of Internet service in public libraries: from architecture perspective. Library, Information and Media Studies. 1(2), p. 1-14. (text in Japanese).

	All terminals		Some terminals		Total
	Prefectural	Municipal	Prefectural	Municipal	
before 1997	0	2	0	0	2
1998	0	2	0	0	2
1999	1	5	0	0	6
2000	1	2	0	0	3
2001	2	17	0	1	20
2002	3	13	2	0	18
2003	1	8	0	0	9
2004	3	17	0	1	21
2005	3	14	1	0	18
After 2006	9	11	2	0	22
I do not know	1	0	0	0	1
N	24	91	5	2	122

Table 1: When the filtering software was first installed

	All terminals		Some terminals		Total
	Prefectural	Municipal	Prefectural	Municipal	
Server software that controls multiple PC terminals	11	44	2	1	58
Software for individual PCs	10	31	3	1	45
Filtering service by ISP	1	6	0	0	7
Other than the above	2	4	0	0	6
I do not know	0	6	0	0	6
Total	24	91	5	2	122

Table 2: Types of filtering software being used

	All terminals		Some terminals		Total
	Prefectural	Municipal	Prefectural	Municipal	
i-FILTER	3	16	0	0	19
InterSafe	4	8	0	1	13
InterScan WebManager	1	8	0	0	9
SmartFilter	1	1	0	0	2
AD-Guard	0	1	0	0	1
Other than the above	1	3	2	0	6
I do not know	0	5	0	0	5
Secret	1	2	0	0	3
Total	11	44	2	1	58

Table 3: Server-type filtering software being used

	All terminals		Some terminals		Total
	Prefectural	Municipal	Prefectural	Municipal	
i-filter	7	22	3	1	33
CYBERsitter	0	6	0	0	6
InterSafe Personal	2	1	0	0	3
Other than the above	1	2	0	0	3
Total	10	31	3	1	45

Table 4: Individual-type filtering software being used

	All terminals		Some terminals		Total
	Prefectural	Municipal	Prefectural	Municipal	
Block specific URL	8	37	2	0	47
Block specific keyword	7	33	1	0	41
Allow specific URL	7	19	1	0	27
Block writing and commenting on BBSs	5	22	0	0	27
Add a category to block	4	20	1	1	24
Make access log secretly, without showing user that filtering software is in operation	1	12	0	0	13
Block specific file extension	2	8	0	0	10
Other than the above	0	3	1	0	4
We do not use special functions	4	5	2	1	12
I do not know	0	8	0	0	8
Number of respondent libraries	24	91	5	2	122

Table 5: Functions being used

	Prefectural		Municipal	
To protect minors from harmful information	1	(62.5%)	7	(79.1%)
To prevent illegal actions by library terminals	1	(62.5%)	4	(53.8%)
To prevent uploading and downloading files	1	(4.2%)	1	(18.7%)
To block viruses	3	(12.5%)	1	(16.5%)
To prevent trouble involving shopping and auctions	2	(8.3%)	1	(11.0%)
Local authorities required the installation	0	(0.0%)	4	(4.4%)
Users required the installation	0	(0.0%)	1	(1.1%)
Other than the above	6	(25.0%)	7	(7.7%)
No answer	5	(20.8%)	5	(5.5%)
Number of respondent libraries	2	(100.0%)	9	(100.0%)

Table 6: Reasons for installing filtering software

	Prefectural		Municipal	
We trust users' own moral sensibilities	2	(40%)	11	(39.3%)
We restrict Internet access by another method	1	(20%)	10	(35.7%)
Filtering software interferes with people's right to know	3	(60%)	5	(17.9%)
Filtering software is too expensive	1	(20%)	5	(17.9%)
We do not trust filtering software	1	(20%)	1	(3.6%)
We did not know of such software	0	(0%)	1	(3.6%)
Users or community raised objections to the installation	0	(0%)	0	(0.0%)
We do not have a definite reason	0	(0%)	4	(14.3%)
Other than the above	2	(40%)	5	(17.9%)
I do not know	0	(0%)	5	(17.9%)
Number of respondent libraries	5	(100%)	28	(100.0%)

Table 7: Reasons for not installing filtering software

Search term	Harmful pages	Non-harmful pages	i-FILTER		InterSafe	
			Harmful page block rate	Non-harmful page incorrect block rate	Harmful page block rate	Non-harmful page incorrect block rate
風俗(sex business)	38	1	100.0	0.0	100.0	0.0
援交 (juvenile prostitution)	37	1	70.3	0.0	73.0	0.0
即ハメ (instant sex)	36	0	86.1	-	75.0	-
おっぱい (tits)	33	1	97.0	100.0	100.0	0.0
セックス (sex)	25	4	92.0	0.0	64.0	0.0
監禁 (bondage)	7	16	100.0	18.8	100.0	0.0
裸 (nude)	5	22	80.0	13.6	80.0	9.1
学校裏サイト (student run school sites)	4	27	75.0	3.7	25.0	3.7
闇サイト (criminal service sites)	4	23	0.0	4.3	0.0	0.0
WAREZ (illegally shared software)	3	21	100.0	9.5	100.0	14.3
自殺サイト (suicide sites)	2	21	100.0	4.8	50.0	9.5
死体 (corpse)	2	20	100.0	5.0	50.0	10.0
彼女募集 (girlfriend search)	1	20	100.0	20.0	0.0	15.0
AV (pornographic video for adults)	1	30	100.0	33.3	100.0	0.0
セクロス (s-x)	1	5	100.0	60.0	100.0	20.0
ウリ (prostitution)	0	21	-	23.8	-	4.8
2ちゃんねる (2-channel)	0	21	-	38.1	-	19.0
氏ね (die)	0	2	-	0.0	-	0.0
ブートレッグ (bootleg)	0	16	-	87.5	-	0.0
リッピング (ripping)	0	27	-	11.1	-	3.7
成りすま (impersonation)	0	17	-	17.6	-	11.8
クラッキング (cracking)	0	25	-	12.0	-	4.0
殺害 (murder)	0	25	-	8.0	-	8.0
青酸カリ (potassium cyanide)	0	10	-	20.0	-	10.0
拉致 (kidnapping)	0	31	-	6.5	-	0.0
大麻 (marijuana)	0	24	-	8.3	-	8.3
クスリ (drugs)	0	26	-	15.4	-	0.0
吉原 (brothel area)	0	21	-	4.8	-	0.0
あそこ (cunt)	0	26	-	11.5	-	3.8
Total	199	504	87.4	16.3	80.4	5.8

Table 8: The ratio of blocked pages by single search word

Introduction to religious Taima (hemp rope) of Ise Shrine （大麻） http://www.isejingu.or.jp/taima/taima1.htm
News article about using a compound in marijuana for medical treatment （大麻） http://slashdot.jp/science/article.pl?sid=07/11/24/0017220
Introduction to the audio visual journal "AV REVIEW" （AV: In Japan "AV" sometimes represents "pornographic video for adults"）http://www.fujisan.co.jp/Product/205/
Introduction to iChat AV from Apple （AV） http://www.apple.com/jp/ftp-info/reference/ichatavJ.html
Glossary from Microsoft （クラッキング） http://support.microsoft.com/kb/878995/ja
How to find and act against student-run school sites（学校裏サイト） http://bcnranking.jp/news/0711/071107_8927.html
Introduction to potassium cyanide（青酸カリ） http://www.rarara.co.jp/kawaraban/kako/2001/0219.html
Book review （援交　是非） http://www.junkudo.co.jp/syohyo200407/syohyo4-tyosho.htm
Book review （死体　是非） http://book.asahi.com/trendwatch/TKY200701210154.html
Introduction to journal article about organ transplants （死体　現状） http://www.bitway.ne.jp/ejournal/club/1402101658.html
Report of International Congress on AIDS in Asia and the Pacific Region（吉原　現状） http://api-net.jfap.or.jp/siryou/2001_repo/repo_16.htm
Online shop selling Hokkaido vegetables （吉原　とは） http://www.fujimotoseika.com/jyuku_j/g_j001.html
Interview with Rakuten employee (company) （WAREZ　現状） http://enterprise.watch.impress.co.jp/cda/meister/2007/10/17/11375.html
Introduction to leetspeak（WAREZ　とは） http://www.hmx-12.net/~virgil7/netEchat/leetspeak.htm
News article from CNET Japan （成りすま　現状） http://japan.cnet.com/blog/blindspot/2006/07/05/windowsophcrack_9fd6/
Fantasy - romance novel（成りすま　とは） http://www2.ocn.ne.jp/~horn/carol/ss/koisuru2.html
News about a new car from Honda （あそこ　とは） http://response.jp/issue/2004/1008/article64442_1.html
Introduction to medicine in general （クスリ　とは） http://www.sagamikanpo.co.jp/kusuri.html
Collection of links to news articles about criminal service sites （闇サイト　とは） http://zapanet.info/tundere/popular/%E9%97%87%E3%82%B5%E3%82%A4%E3%83%88.html
Column about external hard disk drives （リッピング　の是非） http://www.netmania.jp/colum/internet/000460.php

Table 9: Example of non-harmful pages that were blocked (with searching words)

About the Authors

Akiko Yahagi earned her BA in LIS from the University of Tsukuba in 2008. She is currently working at Shibaura Mechatronics Cooperation, and her special area of study is patent information.

Keita Tsuji earned his BA, MS, and Ph.D. in Education from the University of Tokyo in 1994, 1997, and 2003. He is currently an associate professor at the University of Tsukuba. His research interests include automatic thesaurus construction and LIS education.