A Study of PICS/RDF-Based Internet Content Rating System
: Issues Related to Freedom of Expression

PICS/RDF 기반 인터넷 내용 등급 시스템 연구: 표현의 자유를 중심으로

김유승(Youseung Kim)*

ABSTRACT

Since the use of the Internet has proliferated, the availability of illegal and harmful content has been a great concern to both governments and Internet users. Among various issues related to such content, Internet content filtering technologies have been developed for enabling users to deal with harmful content. In recent years, commercial filtering has become massively popular. Many parents, teachers and even governments have chosen commercial filtering software as a feasible technical solution for protecting minors from harmful information on the Internet. The Internet content filtering software market has grown significantly. However, Internet content filtering software has led to intense debate among civil liberties groups. They deem this to be censorship and argue that Internet filtering technologies are simply unworkable because they have inherent weaknesses. They are critical of the fact that most filtering has violated free speech rights and will eventually wipe out minor and controversial, yet innocent incidences of free speech on the Internet. In this article Internet content filtering, in particular PICS/RDF-based label filtering, so-called Internet content rating system, will be explored and its advantages and drawbacks relating to end-users' autonomy and freedom of expression will be discussed.

초 록

인터넷의 대중화와 함께, 인터넷의 불법유해정보의 존재는 정부와 인터넷 사용자들에게 큰 관심거리가 되고 있다. 불법유해정보 문제에 대한 다양한 방법들 중에서, 인터넷 콘텐츠 필터링 기술은 사용자들이 스스로 유해정보 문제에 대처할 수 있도록 개발되어 왔다. 지난 몇 년 사이, 상업 필터링 제품에 대한 관심이 높아지고 있다. 부모, 교사, 심지어는 정부 당국도 청소년을 인터넷 유해정보로부터 보호하는 기술적 대안으로써 상업 필터링 제품을 선택하고 있고, 그 시장도 빠르게 성장하고 있다. 하지만 시민단체들은 중심으로 인터넷 콘텐츠 필터링에 대한 비판의 목소리가 높다. 필터링은 기술적 측면에서 태생적인 약점이 가지고 있을 뿐 아니라, 표현의 자유를 위축시키는 결과를 초래할 것이라는 비판이다. 이 논문은 인터넷 콘텐츠 필터링, 특히 일체화 필터링과 구분되어 내용평가시스템으로 불리는 PICS/RDF 기반의 라벨 필터링의 기술적 측면을 분석하고 표현의 자유, 사용자 자율성과 관련된 문제들을 살펴봄으로써, 불법유해정보에 대한 기술적 해법의 타당성에 대하여 논하고자 한다.

Keywords : internet, filtering, PICS, RDF, rating, freedom of expression

* 시간당사, 중앙대학교 문헌정보학과 (yskim@icogito.net)
■ 논문접수일자 : 2007년 8월 21일
■ 계재확정일자 : 2007년 9월 10일
1. Introduction

As the Internet has allowed information to be circulated with the speed of light and regardless of frontiers, contents which are deemed to be illegal and harmful have also been disseminated. Since the use of the Internet has proliferated, the availability of such content has been a great concern to both governments and Internet users. Illegal content and harmful content are significantly different issues. Each issue requires a separate regulatory solution. While in principle the control of illegal Internet content is an issue of enforcement, the management and control of harmful content is an issue of user and consumer choice (Pierlot 2000). In this context, Internet content filtering technologies have been developed for enabling users to deal with harmful content.

In recent years, commercial filtering has become massively popular. Many proponents, including parents, teachers and governments, have chosen commercial filtering software as a feasible technical solution for protecting minors from harmful information on the Internet, such as child pornography and obscene material. They argue that filtering software has enhanced affectivity and reliability for years, as proven by their popularity in the marketplace. Indeed, during the last few years, with the exponential rise in popularity of the Internet, the Internet content filtering software market has grown significantly. According to a research firm, IDC, in 2005 the revenues of Internet content filtering reached 521.4 million USD and this is expected to have grown to more than 929 million USD by 2009 (Burk & Ryan 2005).

However, Internet content filtering software has led to intense debate among civil liberties groups. They deem this to be censorship and argue that Internet filtering technologies are simply unworkable because they have inherent weaknesses. They are critical of the fact that most filtering software has violated free speech rights and will eventually wipe out minor and controversial, yet innocent incidences of free speech on the Internet.

In this article Internet content filtering, in particular PICS/RDF-based rating systems, will be explored and its advantages and drawbacks relating to end-users' autonomy and freedom of expression will be discussed.
2. Definition of Filtering Software

On the Internet filtering is a technical mechanism for sorting content into categories for the purpose of decreasing accessibility of certain type of content. In general, filtering software is designed in order to prevent Internet access by monitoring user requests and by interceding between user and connection to the Internet. Jonathan D Wallace¹ defines filtering software as follows:

Software products published by commercial software publishers which do any of the following: block access to Internet sites listed in an internal database of the product; block access to Internet sites listed in a database maintained external to the product itself; block access to Internet sites which carry certain ratings assigned to those sites by a third party, or which are unrated under such a system; scan the contents of Internet sites which a user seeks to view and block access based on the occurrence of certain words or phrases on those sites (Wallace 1997).

Currently, there are several different filtering technologies on the Internet: keyword screening, blacklist filtering, whitelist filtering, packet filtering, image analysis filtering, label filtering and so on. Amongst them filtering based on database and keyword or phrase are usually referred to as first generation filtering. These first generation filtering systems have been the dominant filtering methods of commercial Internet content filtering products, while the European Commission has preferred label filtering systems that are based on the Platform for Internet Content Selection (PICS) as opposed to first generation filtering software as a technical solution for preventing illegal and harmful content on the Internet. The first generation filtering technologies are still used alone or in combination by most commercial filtering software in the current marketplace. Although first generation filtering technologies have a number of technical limitations, they are apparently considered to be a feasible tool for addressing issues of inappropriate content on the Internet. Before exploring PICS/RDF-based rating systems in depth, various drawbacks of first generation filtering will be briefly discussed as a preliminary study.

¹ Jonathan D. Wallace publishes an online magazine, Ethical Spectacle (http://www.spectacle.org/) and is co-author of Sex, Laws and Cyberspace (Wallace & Mangan, 1997). He was a co-plaintiff in ACLU vs. Reno, which challenged the CDA.
3. A Critique of First Generation Filtering

In recent years, first generation filtering software has gained in popularity as a feasible solution for addressing issues related to problematic Internet contents. However, it has been criticised for its inherent technical weaknesses and even for being a censorship tool. In this section, a critique of first generation filtering will be analysed.

The first criticism that can be levelled against first generation filtering is that it restricts user autonomy. Most types of filtering software allow a user-defined control. Users can choose certain categories of filter lists. Furthermore, users are allowed to modify the filter lists. Indeed, most filtering software provide various user-customised options.

However, their user-customised options are burdensome and limited. First of all, most of their filter lists are not transparent. Although users can manually add or delete a site or word from filter lists one by one, the remaining thousands of blocked Websites and keywords in the filter lists are still unknown. Thus, users cannot know what their filtering software is blocking in practice. In the case of server-side filtering which is usually employed by ISPs, user autonomy may be far more restricted. It would mean that users' rights to choose certain Internet information are virtually in commercial companies' hands. In my view, the responsibility for deciding what is harmful and what is not should rest with individuals, not with commercial companies. It is said that filtering software is a classic case of a privatised censorship scheme. Indeed, most filtering software companies hold their databases of blocked sites as proprietary information, because the greatest commercial value of filtering software consists in blocking databases. Even if all filtering software made their filter lists viewable, this problem might still not be solved. Because a filter list contains hundreds of thousands of Web pages and is constantly updated, only few parents may be capable of reviewing the entire list, and then customising it. Moreover, most filtering software fails to give users any explanation as to why they block a site.

The second criticism to be levelled at first generation filtering is that it continuously and inevitably raises issues concerning under-blocking and over-blocking. The quantity of information on the Internet is simply too vast. More than seven million new Web pages
are added each day (Cyveillance 2000). Even the world-class search engines reflect only a fraction of content available on the Internet — Deep Web which search engines is hardly able to have access is about 550 times larger than surface Web (Lyman and Hal 2003). It is impossible for filtering software to evaluate all content available on the Internet. Thus, under-blocking is an unavoidable weakness of filtering software. Alongside under-blocking issues, over-blocking raises controversy regarding freedom of expression on the Internet. Filtering software blocks not only harmful Internet sites, but also many controversial and even non-controversial sites. For instance, Internet sites concerning AIDS information and education for safe sex, which might be accessed by a wide range of people including teenagers, are blocked by many commercial filtering products (Kaiser Family Foundation 2002). In addition, some filtering software companies are using their products to suppress criticism of them (Commission on Child Online Protection 2000, 19).

The third point of critique of first generation filtering is that the filtering software cannot understand the various contexts in which information appears. Some experts point out that the problem of Internet content filtering stems from its nature, which largely relies on mindless mechanical blocking through identification of key words and phrases (Heins, Cho, and Feldman 2006). Many sites are blocked by keyword filtering which relies on researching only isolated indecent words. For instance, if a filtering software product is set to block sites which contain the word, 'breast', sites which contain a recipe for chicken breast or medical information concerning breast cancer will be blocked by the filtering product. Although this may be an extreme example, in my view, other similar situations may easily occur with any other word or phrase, such as drug and sex. In this sense, keyword filtering cannot relate the words to their broader context, because context is simply too complex for mechanical evaluation (Balkin, Noveck and Roosevelt 2000).

4. Internet Content Rating System

As discussed above, first generation filtering software poses a number of serious problems none of which are likely to be solved in the foreseeable future. In a sense, the Internet content rating system has been developed as an alternative. Furthermore, it has been
endorsed as a technical solution for preventing children from accessing harmful Internet content by a number of Internet self-regulatory bodies and governments. EU 'Action Plan on Promoting Safer Use of the Internet'\(^2\) has supported the development of an International Internet content rating system taking into account Europe's cultural and linguistic diversity (European Commission 1999, 3). In this section the technical aspects of the Internet content rating system will be explored and two leading rating systems, RSACi and ICRA, will be examined. The advantages and disadvantages of the Internet content rating system will be discussed. The issues of two rating systems concerning free speech rights will also be explored.

4.1 Technical Specifications

4.1.1 PICS

In discussing any issue relating to the Internet content rating system, it is necessary to begin by mentioning the Platform for Internet Content Selection (PICS), since it is the dominant standard for label filtering. PICS was developed as a set of software specifications for label formats and distribution methods by W3C with the participation of many companies, organisations and institutions. W3C (1997a) defines that the PICS specification enables labels (metadata) to be associated with Internet content. It was originally designed to help parents and teachers control what children access on the Internet, but it also facilitates other uses for labels, including code signing and privacy.

In August 1995, the development of technical specifications was launched. In early 1996, the final technical

\(^2\) On 27th November 1997 the European Commission (1997) launched a proposal for an 'Action Plan on Promoting Safe Use of the Internet.' Through the Action Plan, which would cover a three year period from 1998 to 2001, the Commission envisaged four main lines of action: creating a safe environment, including the creation of an European network of hotlines to report illegal content by the public and the development of the Internet industry's self-regulatory schemes for combating illegal content; developing international filtering and rating systems to prevent users from potentially harmful content; encouraging awareness campaigns among the public, in particular parents, teachers and children; and monitoring and support for legal developments in the sector (Akdeniz 2001). The Action Plan completed its first phase during 1999 to 2001. Its extended second phase was also completed between 2002 and 2004. As of 2007, the Action Plan is replaced by the Safer Internet plus Programme which takes place between 2005 and 2008.
specifications were completed (W3C, 1998). Since then, PICS has swiftly caught on with the Internet industry. Several PICS-based rating services have been developed, including RSACi and SafeSurf. Moreover, a number of stand-alone filtering software packages have become PICS-compliant. Microsoft Internet Explorer (IE), which currently dominates the Web browser market all over the world, is compatible with PICS. Consequently, most Internet content rating services today follow the PICS specifications.

PICS was designed to provide a technical standard for creating, distributing and using metadata. Associated with a certain URL, PICS equips various people and organisations to create labels which can provide any kind of descriptive information about Internet content, including the rating information. For instance, if a Web page contains an article which is appropriate only for adults, a label might include the statement that there is a certain type of adult information on the page?labelling can be done either by first-party or by third-party. In other words, PICS is a technical standard for dealing with labels of Web documents at certain URLs.

Balkin, Noveck, and Roosevelt (2000 220), members of the Information Society Project at Yale Law School, said, strictly speaking, PICS itself is not a rating system. Indeed, PICS does not rate anything nor provide a specific rating criterion. It merely gives an outline of the basic format for labelling. Thus, for the implementation of the PICS specifications, a certain rating and labelling service and PICS compatible filtering software are essential. Resnick (1999) lists six major tasks of this implementation that can be operated by various parties as follows: 1) Set labelling vocabulary and criteria for assigning labels, 2) Assign labels, 3) Distribute labels, 4) Write filtering software, 5) Set filtering criteria, 6) Install/run filtering software.

Firstly, to establish an Internet content rating system, the development of a standard vocabulary and categories for labels are required. Here, the term vocabulary means any description of Internet content. For instance, RSACi rates Web content in four categories: violence, nudity, sex, and language. Each category's vocabulary elements, so-called descriptors, are assigned scalar values from zero to four. The sex category includes the following descriptors: Level 1—Passionate kissing, Level 2—Clothed sexual touching, Level 3—Non-explicit sexual acts and Level 4—explicit sexual
acts or sex crimes. If Microsoft IE adjusts the rating level of the sex category to Level 2, it will block Web pages which are rated as level 3 or 4 of this category. The filter setting of Microsoft IE does not seem to be easy to access for some parents who are not computer-literate, since it is hidden several layers down in the main menu rather than appearing in the top menu. Fortunately, the RSACi system has been adopted by Microsoft IE as a default feature. In other rating systems users have to manually install a RAT file\(^3\) on their Web browsers. This matter is, in my view, directly related to the current poor popularity of the Internet content rating system.

However, this kind of standardisation and categorisation can be problematic, since contextual factors of vocabulary elements are easily excluded in those processes. Vocabulary elements may reflect a certain community's moral and cultural values, although information available on the Internet holds very diverse viewpoints. For these reasons, W3C has encouraged the development of a wide range of rating systems in order to maximise user choice. Since PICS allows Internet users to have easy access to the widest possible range of content selection products, and a diversity of voluntary rating systems (W3C 1998), any PICS-compliant software can process any PICS-compliant labels which are provided by various entities. In principle, users can choose their rating services and software, according to their different cultural, political, and religious viewpoints.

Secondly, in order to rate a Website, certain labels should be assigned to the site. Rating can be done not only by the site creator, referred to as the first-party, but also by a third-party (W3C 2000b). Unlike first-party labelling, third-party labelling runs through a server, the so-called label bureau, which is separate from a Web document. This server is an HTTP server that understands a particular query syntax and can provide labels for documents that reside on other servers (Miller 1996). Third-party labelling can be done without any acknowledgment of site creators or information providers.

Thirdly, labels should be transmitted to Internet users who request them for filtering. For transmitting labels, first-party and third-party labelling use different methods. In first-party labelling, according to a user’s request,

---

\(^3\) RAT file is a text file with a filename suffix of .rat which contains a description of a rating system.
one or more labels which are embedded in the header of a Web document are fetched by a browser or stand-alone filtering software. A PICS label places at the head of a HTML document, the \langle head \rangle section which usually carries metadata of the document. The following is an example of embedding a PICS label in an HTML document.

According to W3C, in third-party labelling, labels are transmitted through label bureaus. When an end-user asks to see a particular URL, a software filter makes an inquiry to the label bureau to ask for labels that describe that URL. Depending on what the labels say, the filter may block access to that URL (W3C 2000b).

Fourthly, the development of PICS-compliant filtering software is needed. As discussed above, since PICS is a technical standard, it cannot operate without filtering software which deals with PICS labels. Currently the most common type of label filtering software is a Web browser such as Microsoft Internet Explorer.

Fifthly, users who want to use PICS-based rating systems should choose a rating service which can associate with filtering software. Finally, installing and running PICS-compliant filtering software are the last steps. The software can take place not only at end-user level, but also upstream such as a proxy server, a search engine and an Internet service provider. Since PICS allows the possibility of upstream filtering many Internet libertarian organisations, it has been criticised for threatening end-users’ autonomy and rights to freedom of expression.

4.1.2 RDF

The Resource Description Framework, developed by W3C, is another foundation for supporting metadata. It provides common structures that can be used for the Extensible Markup Language (XML) data exchange (W3C 2001). It is applicable in a variety of areas ranging from cataloging to intelligent software agents and digital signatures. However, for the purpose of this study the discussion will be limited to issues relating to content rating.

The basic RDF model consists of three object types: resources, properties and statements. Firstly, resources means all things described by RDF expressions, such as an entire Website, a part of a Web page, and even an object that is not directly accessible via the Web. Secondly, a property is a specific aspect, characteristic, attribute, or relation used to describe a resource. Thirdly, a statement
is a specific resource together with a named property plus the value of that property for that resource. A statement contains these three elements which are called *subject, predicate*, and *object* respectively (W3C 1999). Just as a PICS label contains several different ratings, a single RDF statement is able to assign a number of properties (Balkin, Noveck, and Roosevelt 2000, 228).

Indeed, RDF is quite similar to PICS. It can express anything that PICS can. Furthermore, it provides a model for representing metadata that is even more general than PICS with more express power (W3C 2000c). RDF has a class system. A collection of classes is called a schema. A PICS rating service description is analogous to an RDF schema. In this sense, RDF is referred to as a successor to PICS. Phil Archer (2004), chief technology officer of ICRA, stated that as an XML-based technology, RDF can be deployed just as easily in mobile communications infrastructure as on the fixed Internet, as well as any other medium that has occasional or permanent network access such as games consoles and digital TV. The potential is significant.

ICRA launched a project which is named Customisation and Personalisation thorough RDF in February 2003. Finally, in January 2006 it published a new ICRA labelling system specification which employs RDF (ICRA 2006a).

### 4.2 Internet Content Rating System: Technical Analysis

Now, I will explore two Internet content rating systems: the RSACi system and the ICRA system, mainly from technical aspects. The RSACi system are one of the earliest practical PICS-based rating systems which was developed in early 1996. They were almost simultaneously developed with PICS. Before the advent of the ICRA system, the RSACi system was the most widespread system worldwide. While the RSACi systems are based in the US, the ICRA system is developed under the European Commission’s ‘Action Plan for Promoting Safer Use of the Internet.’

#### 4.2.1 The RSACi System

The RSACi system was developed by the Recreational Software Advisory Council (RSAC) which is an independent, non-profit organisation based in the US. RSAC was established in 1994 to rate video games for violent content, bad language, sex and nudity. The
original RSAC rating system was developed in September 1994 in direct response to the threat of congressional legislation that sought to control levels of violence in the computer game market (W3C 1996). Since then, RSAC has extended its original rating system to the Internet largely in response to the attempts of the US government to regulate indecent information on the Internet. RSACi is an acronym for the Recreational Software Advisory Council on the Internet. In this sense, it can be said that the RSACi system is an offshoot of the former RSAC system. In November 1995 the RSACi Working Group had its first meeting with representatives from Microsoft, ATT, Bell Atlantic, Time Warner and others. In February 1996 RSAC announced the launch of RSACi and since April 1996 the RSACi rating system has been available to the public.

The RSACi system rates Web content in four categories: violence, nudity, sex, and language on a scale of 0 to 4, from None through progressively stronger examples.

RSACi is currently governed by the Internet Content Rating Association (ICRA), since RSAC transferred its assets, including the RSACi system, to ICRA in April 1999. Thus, RSAC no longer exists. However, it does not necessarily mean that the RSACi system is not working any more. The latest version of Microsoft Internet Explorer (version 6.0) still has the RSACi system as its default rating system. It is the only rating system which Microsoft IE has adopted as a default option. For this reason, it is still one of the most widespread Internet content rating systems worldwide. The RSACi label is now provided by ICRA alongside the ICRA label. For instance, CNet.com has both labels as follows:

\[
\text{<meta http-equiv="picc-label" content="\{picc-1,1 http://www.icra.org/ratingsw02.html\} gen true for http://www.cnet.com" \r (cz 1 lz 1 nz 1 oz 1 vz 1) http://www.rsac.org/ratingsw01.html\} gen true for http://www.cnet.com" \r (n 0 s 0 v 0 1))/>\]

Although it has had this advantage, its popularity is rather disappointing. According to a report, by October 2000 about 150,000 Websites have rated themselves with the RSACi system (Keller & Verhulst, 2000). ICRA states that those rated Websites includes a great proportion of the top 100 sites which account for 80 percent of the whole traffic on the World Wide Web (ICRA 1999). However, this number is very small compared to the number of total Web pages on the Internet. During the last years there is no significant change in the number of rated Websites. This kind of poor popularity does not
affect only the RSACi system. All the PICS/RDF–based Internet content rating systems which are currently available have suffered from the same problem. Since search engines and news sites which contain a vast amount of varied information, it may not be easy to rate themselves by a single category. However, it does not mean that there is no need to rate these sites. The number of rated Websites is vital, because the success of the rating system largely depends on it. The rating system needs to reach a critical mass for achieving its practical force. For this reason, ratings of these heavily trafficked, popular and influential sites are important.

Furthermore, it also have drawn some complaints of American cultural bias (Keller and Verhulst 2000). Therefore, many institutions in Europe, such as INCORE, INHOPE and the European Commission, have made efforts to establish Internet content rating systems for the European and International markets. As a result, in December 2000 ICRA introduced the new ICRA labelling system.

4.2.2 The ICRA System

In March 1999 ICRA incorporated as a non-profit organisation in London and was officially launched two month later. It has offices in the UK in Brighton and in the US in Washington, D.C. ICRA has received project funding from the European Commission under the 'Action Plan for Promoting Safer Use of the Internet' and is supported by many non-profit organisations and Internet companies, such as Yahoo!, AOL Europe and Microsoft. Unlike other rating systems, the ICRA system is a multi-party rating system (MPRS) that is theoretically based on the layer cake model which was proposed by the Information Society Project at Yale Law School. The layer cake model can be illustrated as follows (Figure 1).

The plate is the software specification which includes PICS, PICSRules and RDF. The first layer of the cake is a basic vocabulary that is used by first–parties in rating their sites. In the ICRA system first–parties do not rate their Web content sites in certain categories on scalar numbers of levels. Instead, they list all vocabulary elements which are applicable to their Web content.
In other words, the ICRA system separates the vocabulary elements from the construction of rating templates. In my view, this feature makes the system relatively objective and value-neutral as compared to other rating systems, since the construction of rating templates inevitably involves some degree of value-judgment.

The ICRA system has 45 descriptors. Up to 40 descriptors can be selected together. Its labelling questionnaires and filtering interfaces are available in several languages, including English, German, French, Spanish and Chinese (Hong Kong). The ICRA descriptors and codes are shown above (Table 1).

However, since ICRA has introduced RDF-based labelling specification in 2006, the ICRA label can be encoded using any RDF scheme. If the RDF instance is called label.rdf and is located in the root of the website, the Link tag can be shown as follows:

```html
<link rel="meta" href="/labels.rdf" type="application/rdf+xml" title="ICRA labels" />
```

In this case, RDF content label can be encoded as shown in (Table 2).

The second layer consists of rating templates which are created by third-parties. Third-parties take certain vocabulary elements and arrange them into categories and scalar orders. Thus, third-parties do not have to rate enormous numbers of Websites in order to create templates. The cost of creating templates can therefore be significantly reduced. In this sense, it is expected that each third-party may provide different templates that reflect diversity of information on the Internet. An expert report from the Information Society Project states that by combining a basic vocabulary at level one with flexibility at level two users can achieve much greater diversity and provide more end-user choice than in a unitary system (Balkin, Noveck, and Roosevelt 2000, 247).

The third layer is a set of third-party ratings of individual sites. For instance, any URL-based filtering systems which are compatible with PICS can be placed at the third layer. However, it is feared that there is a possibility that this kind of additional third-party rating can have a negative effect on the rest of the layer structures. As discussed above, the so-called first generation filtering technologies pose serious technical shortcomings and have been criticised for violating end-users' autonomy.

In sum, the ICRA system is a more flexible and relatively objective solution compared to other rating systems.
### Table 1: The ICRA descriptors and associated codes (ICRA 2006b)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Chat</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chat</td>
<td>cb 1</td>
<td>ca 1</td>
<td></td>
</tr>
<tr>
<td>Moderated chat suitable for children and teens</td>
<td>ca 1</td>
<td>cb 1</td>
<td></td>
</tr>
<tr>
<td>None of the above</td>
<td>cz 1 / cz 0</td>
<td>cz 1 / cz 0</td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explicit sexual language</td>
<td>sd 1</td>
<td>ia 1</td>
<td>i4</td>
</tr>
<tr>
<td>Crude words or profanity</td>
<td>ls 1 &amp; lb 1</td>
<td>lb 1</td>
<td>12</td>
</tr>
<tr>
<td>Mild expletives</td>
<td>lc 1</td>
<td>lc 1</td>
<td>11</td>
</tr>
<tr>
<td>None of the above</td>
<td>Iz 1 / Iz 0</td>
<td>Iz 1 / Iz 0</td>
<td>10</td>
</tr>
<tr>
<td>Nudity &amp; Sexual Material</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erections or female genitals in detail</td>
<td>se 1</td>
<td>na 1</td>
<td>n4</td>
</tr>
<tr>
<td>Male genitals</td>
<td>nc 1</td>
<td>nb 1</td>
<td>n3</td>
</tr>
<tr>
<td>Female genitals</td>
<td>nc 1</td>
<td>nc 1</td>
<td>n3</td>
</tr>
<tr>
<td>Female breasts</td>
<td>na 1</td>
<td>nd 1</td>
<td>n2</td>
</tr>
<tr>
<td>Bare buttocks</td>
<td>nb 1</td>
<td>ne 1</td>
<td>n2</td>
</tr>
<tr>
<td>Explicit sexual acts</td>
<td>se 1</td>
<td>nf 1</td>
<td>s4</td>
</tr>
<tr>
<td>Obscured or implied sexual acts</td>
<td>sb 1</td>
<td>ng 1</td>
<td>s3</td>
</tr>
<tr>
<td>Visible sexual touching</td>
<td>sc 1</td>
<td>nh 1</td>
<td>s3</td>
</tr>
<tr>
<td>Passionate kissing</td>
<td>sa 1</td>
<td>ni 1</td>
<td>s1</td>
</tr>
<tr>
<td>None of the above</td>
<td>nz1&amp;sz1/ nz0&amp;sz0</td>
<td>nz 1 / nz 0</td>
<td>s0</td>
</tr>
<tr>
<td>Context – Artistic</td>
<td>xa</td>
<td>nr 1</td>
<td>–</td>
</tr>
<tr>
<td>Context – Educational</td>
<td>xb</td>
<td>ns 1</td>
<td>–</td>
</tr>
<tr>
<td>Context – Medical</td>
<td>xc</td>
<td>nt 1</td>
<td>–</td>
</tr>
<tr>
<td>Other Topics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Promotion of tobacco use</td>
<td>oa 1</td>
<td>oa 1</td>
<td></td>
</tr>
<tr>
<td>Promotion of alcohol use</td>
<td>ob 1</td>
<td>ob 1</td>
<td></td>
</tr>
<tr>
<td>Promotion of drug use</td>
<td>oc 1</td>
<td>oe 1</td>
<td></td>
</tr>
<tr>
<td>Gambling</td>
<td>oe 1</td>
<td>od 1</td>
<td></td>
</tr>
<tr>
<td>Promotion of weapon use</td>
<td>od 1</td>
<td>oe 1</td>
<td></td>
</tr>
<tr>
<td>Promotion of harm against people</td>
<td>og</td>
<td>of 1</td>
<td></td>
</tr>
<tr>
<td>Material that might be perceived as setting a bad example for young children</td>
<td>oh</td>
<td>og 1</td>
<td></td>
</tr>
<tr>
<td>Material that might disturb young children</td>
<td>of</td>
<td>oh 1</td>
<td></td>
</tr>
<tr>
<td>None of the above</td>
<td>oz 1 / oz 0</td>
<td>oz 1 / oz 0</td>
<td></td>
</tr>
<tr>
<td>Violence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sexual violence / rape</td>
<td>va 1</td>
<td>va 1</td>
<td>v4</td>
</tr>
<tr>
<td>Blood and gore, human beings</td>
<td>ve 1</td>
<td>vb 1</td>
<td>v4</td>
</tr>
<tr>
<td>Blood and gore, animals</td>
<td>vf 1</td>
<td>ve 1</td>
<td>v4</td>
</tr>
<tr>
<td>Blood and gore, fantasy characters (including animation)</td>
<td>vg 1</td>
<td>vd 1</td>
<td>v4</td>
</tr>
<tr>
<td>Killing of human beings</td>
<td>vh 1</td>
<td>ve 1</td>
<td>v3</td>
</tr>
<tr>
<td>Killing of animals</td>
<td>vi 1</td>
<td>vf 1</td>
<td>v3</td>
</tr>
<tr>
<td>Killing of fantasy characters (including animation)</td>
<td>vj 1</td>
<td>vg 1</td>
<td>v3</td>
</tr>
<tr>
<td>Deliberate injury to human beings</td>
<td>vb 1</td>
<td>vh 1</td>
<td>v1</td>
</tr>
<tr>
<td>Deliberate injury to animals</td>
<td>vc 1</td>
<td>vi 1</td>
<td>v1</td>
</tr>
<tr>
<td>Deliberate injury to fantasy characters (including animations)</td>
<td>vd 1</td>
<td>vj 1</td>
<td>v1</td>
</tr>
<tr>
<td>Deliberate damage to objects</td>
<td>Not defined</td>
<td>vk 1</td>
<td>v1</td>
</tr>
<tr>
<td>None of the above</td>
<td>vz 1 / vz 0</td>
<td>vz 1 / vz 0</td>
<td>v0</td>
</tr>
<tr>
<td>Context – Artistic</td>
<td>xa 1</td>
<td>vr 1</td>
<td>–</td>
</tr>
<tr>
<td>Context – Educational</td>
<td>xb 1</td>
<td>va 1</td>
<td>–</td>
</tr>
<tr>
<td>Context – Medical</td>
<td>xc 1</td>
<td>vt 1</td>
<td>–</td>
</tr>
<tr>
<td>Context – Sports</td>
<td>xd 1</td>
<td>vu 1</td>
<td>–</td>
</tr>
</tbody>
</table>
### Table 2: An Example RDF Instance containing ICRA labels (ICRA 2006a)

The host(s) for which the label(s) are applicable. Subdomains are in scope.

```xml
<?xml version="1.0"?>
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:xsd="http://www.w3.org/2000/01/XMLSchema#
xmlns:label="http://www.w3.org/2004/02/22-rdf-syntax-ns#
xmlns:dc="http://purl.org/dc/elements/1.1"
xmlns:icra="http://www.icra.org/rdfs/vocabulary03#"

<rdf:Description rdf:about=""
	<dc:creator rdf:resource="http://www.icra.org"/>
	<label:authorityFor rdf:resource="http://www.icra.org/rdfs/vocabulary03#"
	<label:authorityFor/>

</rdf:Description>

An additional string that must match the resource's URI for any labels in the RDF instance to be applicable.

```xml
<label:Ruleset>
	<label:hasHostRestrictions>
		<label:Hosts>
			<label:hostRestriction>example.org</label:hostRestriction>
			<label:hostRestriction>example.com</label:hostRestriction>
		</label:Hosts>
	</label:hasHostRestrictions>
</label:Ruleset>

The default label.

```xml
<label:Ruleset>
	<label:hasHostRestrictions>
		<label:Hosts>
			<label:hostRestriction>example.org</label:hostRestriction>
			<label:hostRestriction>example.com</label:hostRestriction>
		</label:Hosts>
	</label:hasHostRestrictions>
</label:Ruleset>

An ordered sequence of rules that should be matched against a resource's URI. If a rule is satisfied it must provide a label that overrides any default.

```xml
<label:ContentLabel rdf:ID="label_1"
	<rdfs:comment>Label for all/most of website</rdfs:comment>
	<rdfs:label>No nudity, no sexual content, no violence, no potentially offensive language, no potentially harmful activities, no user-generated content</rdfs:label>
	<icra:na1/>
</label:ContentLabel>

A description of the RDF instance itself that identifies where additional information about the label can be found, including how its veracity can be assessed.

```xml
<label:ContentLabel rdf:ID="label_3"
	<rdfs:comment>Label for photography section</rdfs:comment>
	<rdfs:label>Exposed breasts, bare buttocks, no sexual content, no violence, no potentially offensive language, no potentially harmful activities, no user-generated content</rdfs:label>
	<label:hasModifier>crest:ha</label:hasModifier>
</label:ContentLabel>

<label:ContentLabel rdf:ID="label_4"
	<rdfs:comment>Label for guestbook and message board</rdfs:comment>
	<rdfs:label>No nudity, no sexual content, no violence, no potentially offensive language, no potentially harmful activities, user-generated content (moderated)</rdfs:label>
</label:ContentLabel>
</rdf:RDF>
```
Furthermore, it provides globally translatable descriptors. However, this does not mean that the ICRA system is a perfect solution. Another concern is that its stand-alone filtering software increasingly emphasises the first generation filtering methods rather than its original labelling system, thus it resembles other commercial filtering software. In the next section the advantages and disadvantages of Internet content rating will be critically discussed.

4.3 Advantages

The rating systems which are based on the PICS specifications are more sophisticated compared to first generation filtering software. While first generation filtering software manually rates individual Web pages only as adult or child safe, or block or no-block, the PICS/RDF-based rating software rates Web pages along multiple dimensions such as violence, nudity, sex, and language. They also allow users to control any number of values for any given dimensions. For instance, a parent can block only sites rated over 3 for violence and 8 for sex. This means that parents are able to create their own filtering rules for their children. It means that the PICS/RDF-based rating software can be customised by any end-user. This flexibility is a very important feature as far as end-users' autonomy is concerned, since not everyone necessarily wants to block the same Web pages. In my view, the PICS/RDF-based rating system is a significant advance in Internet filtering software. The theory of PICS empowers Internet users to control their own access to Internet content, and would reduce the risk of government censorship. Despite these advanced features, Internet content rating systems have been criticised for many reasons, from technical issues to issues of free speech.

4.4 Technical Disadvantages

As regards technical issues, the first point of critique is that there is an easy loophole to circumvent the rating system. While Microsoft IE version 4 or above has supported various PICS/RDF-based rating systems through its Content Advisor, other popular browsers, Opera and Mozilla Firefox, do not support any rating system. Using these Web browsers, children can easily evade the system.

The second point of critique is that the Internet content rating system's filtering coverage is very narrow. The system is currently working only on the
World Wide Web, while first generation filtering software is generally able to filter most types of Internet communications. According to statistics from the Internet Watch Foundation\(^4\) covering the last six years, Usenet is a significantly problematic part of the entire Internet. It cannot be rated by the PICS system. E-mail, chat room, FTP and newsgroup are also beyond the PICS/RDF-based rating’s targets.

The third point of critique of the Internet content rating system is about whether or not the system can be enforced. The success of the Internet content rating system, including the ICRA system, largely depends on the number of rated Websites. In order to achieve a viable rating system, it should reach a critical mass. However, currently the number of rated Websites constitutes too small a proportion of the total number of Websites, though ICRA has made great efforts to promote its rating system with the European Commission’s support. Keller and Verhulst (2000) explain that since the ICRA model relies largely on uncompensated effort by both first-party content providers and third-party list makers, it is important to find means to both encourage participation as easy as possible.

However, so far, no Internet content rating system seems to find that means. Poor participation results in poor enforcement of the rating system, and then this poor enforcement reproduces people’s poor involvement on an enlarged scale. In my view, the best solution to gain the public’s popularity is providing a rating system which is easily accessible and user-friendly. In this context the rating system would be loaded into Microsoft IE as a default top menu. Developing stand-alone software which is similar to many other commercial filtering products cannot be the right answer. At least the ICRA system would have been a default rating system of Microsoft IE. Nevertheless, it is not expected that the ICRA system becomes its default rating system, since the European Union has prevented the US-based company from being an official partner of its ‘Action Plan for Promoting Safer Use of the Internet’ (Archer 2004).

\(^4\) The Internet Watch Foundation (IWF) was launched to address the problem of illegal and harmful content on the UK Internet, with particular reference to child pornography in September 1996. It was established to implement an industry proposal, ‘R3 Safety–Net.’ ‘R3’ referred to the triple approach of the proposal: rating, reporting, and responsibility. The proposal endorsed the establishment of a hotline reporting system and a PICS based rating system. It also emphasised the industry’s responsible services and self-regulatory efforts.
4.5 The Right to Free Speech and the Internet Content Rating System

Apart from the technical defects of the Internet content rating system, many libertarians and civil organisations have argued that the PICS/RDF-based rating system may violate freedom of expression on the Internet. They have argued that third party rating systems do not guarantee transparency and accountability, and therefore may raise private censorship issues. There will be no space for free speech arguments and dissent because the ratings will be done by private bodies and the governments will not be involved "directly." When censorship is implemented by government threat in the background, but run by private parties, legal action is nearly impossible, accountability difficult, and the system is not open or democratic.

The ACLU strongly objects to Internet content rating for reasons detailed in its report, *Fahrenheit 451* 2:5) *Is cyberspace burning?* The ACLU insists that Internet content rating may cause controversial speech to be censored (ACLU 1997).

In fact, the PICS/RDF-based rating systems have faced serious difficulties in dealing with contextual value, in just the same way as the first generation filtering software. The RSACi system excluded contextual factors from vocabulary elements. It could not distinguish between artistic nudity and obscene nudity (Balkin, Noveck, and Roosevelt 2000, 251–254). As previously mentioned, there are so many different standards relating to various aspects of life worldwide on the Internet that it is impossible to apply one subjective standard to the entire Internet community. Therefore, subjective rating categories are highly controversial. Objectivity is needed to retain reliability regarding rating systems.

In this sense, the ICRA system, which is referred to as the multi-party rating system, made an effort to provide potentially objective rating terminologies. It separates the vocabulary elements from the construction of rating templates which inevitably involves some degree of value judgment. Indeed, it has made a significant advance with regard to many aspects of the Internet content rating system.6) However, it does not

---

5) "Fahrenheit 451" is a title of Ray Bradbury's novel which was initially published in 1953. The novel depicts the futuristic world where freedom of thought and speech are gone. Fahrenheit 451 is the temperature at which books burn.
yet provide a perfect solution. The ICRA system's descriptors do not provide absolute objectivity. In my view, its descriptors, such as "passionate kissing" and "material that might disturb young children" are rather subjective. Thus the ICRA system cannot be free from criticism against subjective value-judgment, just like other rating systems.

Furthermore, from the beginning the Internet content rating system has been criticised for empowering governments to control the access of their adult citizens. There is a possibility that PICSRules can be used for the purposes of enabling the development of country profiles to facilitate a global or universal rating system desired by governments, because it can block access to content on entire domains, via the specification of full or partial domain names and/or IP addresses, regardless of the username, port number, or particular file path that is specified in the URL (GILC 1999).

There is always a potential for people to cheat in their self-rating. For instance, someone, who runs a commercial Website for adults realises that many people will not get to his or her site if it is either rated as sexually explicit or not rated at all. He or she may rate the Website OK for minors. In addition, mis-rating can happen unintentionally because many Web pages contain much more complex information than the given rating categories can cope with. In this sense, the PICS/RDF-based rating systems which largely rely on the concept of self-rating may break down in the absence of a penalty system for mis-rating. The rating system may have the potential to lead to heavy-handed government censorship.

5. A Mandatory Internet Content Rating System in Korea

In 2000 the Korean government proposed the Act on Promotion of Information and Communication Network Utilisation and Information Protection, etc., the so-called Tongsin Jilseo Hwakripbeop

---

6) According to the Final Report for the DVB Regulatory Group by Keller and Verbuijl (2000), a multi-party rating system has a number of strengths as follows:

[It] makes possible comparatively thorough coverage of the net, [It] enables individual parental control of content filtering [and] allows flexible adaptation across diverse cultural groups. [It also] draws on existing, globally applicable technological standards [and] operates with no direct cost to parents or content providers. [Furthermore, it] has the backing of major industry participants.
[Communication Order Establishment Law]. The proposed Act presented a mandatory Internet content rating system. According to the proposed Act, a person who provides harmful information to minors on the Internet should self-rate their information and display the rating.

For several reasons, opponents, such as the Jinbo Network Centre, perceived the proposed Act's Internet content rating system as a governmental censorship system (Hong 2001). Firstly, it proposed that it should be managed by a government institution, although the Internet content rating systems in most other countries have been conducted by non-governmental organisations. Secondly, under the proposed Act, rating Internet content was not a recommended option, but a legal requirement. In a sense this Article might sound reasonable in terms of child protection on the Internet. However, the article was clearly contradictory, because the notion of "harmful information" on the Internet is too vague. In the COPA (CDA II) case the US Court held that applying a concept of harmful-to-youth to the Internet could result in blocking a substantial amount of Internet content which is lawful to adults. The Korean Constitutional Court also states that, as regards content regulation, the concept of regulatory object should not be ambiguous, abstract, or comprehensive. Otherwise, it may be resulted in the regulation of communication that should not be regulated and may lead to the violation of the rule against excessive restriction (14–1 KCC 616, 99Hun–Ma480, June 27, 2002). In fact, courts often find it very difficult to judge this kind of issue. The Happy Sara case (94Do2413), the Lie to Me case (98Do679) and the Kim In-kyu case (2001Go–Hap54) are prime examples.

Nevertheless, under the proposed Act all judgments concerning harmful information on the Internet were entirely dependent on the decisions taken by the Information Communication Ethics Committee (ICEC). According to Article 29 of the proposed Act, ICEC was able to require Internet service providers to stop providing services to people who do not rate their Internet content. Worse still, according to Article 33, anyone who thinks Internet content rating is inappropriate could require ICEC to re-examine the rating. Then ICEC could require the content provider to submit information relating to its rating, and order a revision of the rating. Ironically, since "anyone can require ICEC," it was possible that ICEC could require itself to re-examine any Internet content's
rating. Consequently, ICEC was virtually able to control the rating of any Internet information which falls within the South Korean government’s jurisdiction. Furthermore, the proposed Act imposed a heavy penalty for mis-rating or non-rating. According to Article 77, a person who mis-rates on purpose could be sentenced to up to three years’ imprisonment or punished with a fine of 30 million KRW.

Another problem was found in Article 34 which required that all public institutions, including schools and libraries, must install Internet rating software on their terminals. Installing Internet content filtering software in public institutions has proved to be controversial. Article 34 targets all the public Internet access points which youths may use, including the Internet café. In my view, the issues concerning whether Internet content rating software should be installed at a school should be decided by the school, not by the government. Installing Internet content rating software in libraries also raises issues of serious concern. Libraries are institutions not only for children and teenagers but also for adults so installing any content filtering software in libraries may violate adult users’ freedom of expression.

Fierce criticisms from many civil rights organisations and Internet users were made regarding the proposed Act. Finally, the section relating to the Internet content rating system was removed from the final Act which was passed by the National Assembly in December 2000. This case has been the prime example that shows possibilities which any developing nation worldwide may employ the Internet content rating system as a censorship tool rather than an optimal technical solution for self-regulation on the Internet.

6. Conclusion

In summation, no filtering technology is entirely accurate and reliable. They frequently fail in their mission which is to restrict children’s access to harmful information on the Internet effectively. Nevertheless, advocates including many parents and organisations seem to think that these filtering technologies and products are better than nothing. As mentioned above, the filtering software products are widely used in homes, schools, and even libraries and they are gaining in popularity. One of the CIPA’s authors, Ernest Istook, argues that blocking some legitimate information
is a price worth paying to protect minors from unwanted information on the Internet. He said, Filters will never be perfect, but that is no excuse not to try to protect our children (Das & Pike, 2001).

However, in my view, these ideas give rise to serious problems. First of all, the serious shortcomings of filtering are not temporary, but inherent. Why should free speech rights be restricted because of the imperfection of filtering technologies? Freedom of expression is a thing of great value which must not be compromised by efforts to achieve a safe Internet. Moreover, there is a risk that parents will put excessive confidence in commercial filtering software, since most commercial filtering product companies are unlikely to inform end-users that their products have inherent technical limitations, whereas they are quick to advertise how brilliant their products are. In other words, the use of filtering products may give parents and teachers a false sense of security. In this context, the Internet Watch Foundation states that parents and teachers should be aware of filtering software's technical weaknesses and limitations:(IWF 2003)

The most important thing to remember when it comes to considering which tools to use is that no single filtering product can be guaranteed to totally protect your child from accessing inappropriate material. ... Like a seat belt in a car, a filter can help protect you but it cannot guarantee you will not have a crash!

I have no objection to parents deciding to use filtering products at home for their own children, as long as they are aware of its limitations. However, installing mandatory filtering software at public Internet access points, such as public libraries and Internet café, is a different case, as it may breach people's rights to access certain information. The ALA Intellectual Freedom Committee (2000) states that filtering products have created a dissonance with the basic mission of libraries. It claims, libraries are responsible for serving a broad and diverse community with different preferences and views, Blocking Internet sites is antithetical to library missions because it requires the library to limit information access.

In this context, the PICS/RDF-based rating system may be an impressive
second-best solution. However, it is not a satisfactory solution for issues relating not only to freedom and regulation but also to child protection. On the contrary, its advantages are almost negated by its disadvantages. As discussed, its practical effectiveness is still very doubtful, while it involves controversial issues relating to freedom of expression. Indeed, it is questionable whether it is an appropriate regulatory method for dealing with harmful Internet content—illegal Internet content, such as child pornography, is beyond the scope of the rating system, because this kind of illegal content is forbidden for any conceivable audience and should be regulated by the enforcement of laws (Cyber-Rights & Cyber-Liberties (UK) 1997).

As a result, I identify that the Internet content rating system 1) does not have broad public support; 2) may not be enforceable; 3) is not easy to understand, because of complex technical issues; 4) may bring unintended consequences, such as a chilling effect on freedom of expression; 5) is a response to a short-term public concern; 6) may create a false sense of security for concerned citizens, because of its technical defects; 7) may unconditionally prohibit harmful content that is freely available to adults in other media; 8) may not fulfill its public accountability, because of its industry-based nature. The dissemination of harmful content on the Internet is a serious social concern that needs to be addressed. However, in my view, the filtering and rating systems do not seem to be appropriate solutions. As PCMLP (2004, 70) points out, Internet content filtering remains an area where self-regulation has raised far more concerns than solutions. If so, what can be the alternatives? Akdeniz (2004, 120) also claims, there should be more emphasis on promoting the Internet as a positive and beneficial medium and there is urgent need for awareness of Internet usage. As discussed previously, the role of Internet users in controlling harmful Internet content is crucial. Ultimately, parents and teachers have the prime responsibility for the protection of children from accessing potentially harmful content on the Internet.
References


Keller, Daphne & Verhulst, Stefaan. 2000. Parental control in a converged communications environment self-regulation, technical devices and meta-


