

RESEARCH OUTPUTS / RÉSULTATS DE RECHERCHE

What can be regulated on the Internet by control/filtering software

d'Udekem-Gevers, Marie

Published in:

Proceedings of fifth World Conference Human Choice and Computers HCC-5, Geneva, Switzerland 25-28 August 1998

Publication date:

1998

Document Version

Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for pulished version (HARVARD):

d'Udekem-Gevers, M 1998, What can be regulated on the Internet by control/filtering software. in S Munari, G Krarup & L Rasmussen (eds), *Proceedings of fifth World Conference Human Choice and Computers HCC-5, Geneva, Switzerland 25-28 August 1998: Computers and networks in the age of globalization*. University of Lausanne, Lausanne, pp. 315-334.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

What can be regulated on the Internet by control/filtering software ?

Marie d'Udekem-Gevers

mgevers@info.fundp.ac.be

<http://www.info.fundp.ac.be/~mge/>

Cellule Interfacultaire de Technology Assessment (CITA)

Facultés Universitaires Notre-Dame de la Paix, Namur, Belgique

with financial support of the Belgium *Services fédéraux des affaires scientifiques, techniques et culturelles*
Programme "Interuniversity Poles of Attraction - Phase IV - Project 31"

Draft Version of 17 July 1998

© Copyright Marie d'UDEKEM-GEVERS

0. INTRODUCTION	3
1. GENERAL FRAMEWORK	5
1.1. Kinds of control/filtering	5
1.2. Possible locations of control/filtering	5
1.3. Control scopes and corresponding possible technical solutions	6
2. EXAMPLES OF CURRENT FILTERING SOFTWARE PACKAGES AND CORRESPONDING CONTROL SCOPES	6
3. PICS RATING SERVICES	8
3.1. PICS Third-Party Rating Services	8
3.2. PICS Self-Rating Services	12
3.3. Some General Remarks and Conclusions about these Observations	14
3.3.1. General Remarks	15
3.3.2. Conclusions about these Observations	15
3.4. Generalisation : What could be rated with PICS?	16
4. ETHICAL ISSUES AND QUESTIONS	17
4.1. Outside PICS	17
4.2. With PICS	17
4.2.1. 'Set labelling vocabulary and criteria for assigning labels'	17
4.2.2. 'Assign labels'	17
4.2.3. 'Write filtering software'	18
4.2.4. 'Set filtering criteria'	18

0. Introduction

On no account is this paper moralistic or prescriptive. It is only an analytic description of off-the-shelf control/filtering¹ software and a synthesis of the literature on this topic.

The following explanations are needed to understand this complex domain.

PICS (Platform for Internet Content Selection) is a set of technical standards which have been developed since summer 1995 by the MIT's (Massachusetts Institute of technology) World Wide Web Consortium. Currently, PICS becomes more and more important and control software like, for example, Cyber Patrol² does not hesitate to become PICS compliant.

« The first and most important distinction that PICS introduced is a **separation between labelling and filtering**³. A label⁴ describes the content of something. A filter makes the content inaccessible to some audience»⁵ In other words, thanks to PICS « Consumers choose their selection software and their label sources (called rating service) independently. »⁶ Moreover, PICS standards facilitate « *self rating* (enable content providers to voluntarily label the content they create and distribute) *and third party rating* (enable multiple, independent labelling services to associate additional labels with content created and distributed by others. Services may devise their own labelling systems, and the same content may receive different labels from different services. »⁷

« More generally, there are six roles that could be filled by different entities » (see table 0).

¹ 'Control' and 'filtering' are considered here as synonymous.

² According several comparative reviews (see Munro C. 1997, Internet Filtering Utilities, *PC Magazine*, April 8 1997, pp. 235-240.; Parental Control Software at a Glance, October 97 issue of *FamilyPC* <http://www.zdnet.com/familypc/9709/noway/table.html>; Meeks Ch., 8 programs to porn-proof the Net, 4/3/96; updated 5/28/97 <http://www.cnet.com/Contenu/Reviews/Compare/Safesurf>), Cyber Patrol is the best among the tested packages.

³ « PICSRules is a language for expressing filtering rules (profiles) that allow or block access to URLs based on PICS labels that describe those URLs » PICS FAQ p. 10

⁴ « PICS labels can be attached or detached » (and stored on a separate server called a 'label bureau'). (Resnick 1997)

⁵ Resnick 1998 PICS FAQ

⁶ Resnick and Miller 1996 p. 88

⁷ (see <http://www.w3.org/PICS/principles.html>).

Table 0: The 6 roles implied by filtering software (according to Resnick 1998)⁸

1. Set labelling vocabulary and criteria for assigning labels
2. Assign labels
3. Distribute labels
4. Write filtering software
5. Set filtering criteria [for example by parents]
6. Install/run filtering software.

⁸ see <http://www.w3.org/PICS/PICS-FAQ-980126.html>

1. General framework

1.1. Kinds of control/filtering

Table 1.1.6 different kinds of control/filtering defined by Cranor & Resnick (12 March 1998)⁹

1. SUGGEST	Recommend appropriate content for children [ex CyberYes list (of CyberPatrol)]
2. SEARCH	Select a content that is appropriate for children and match a query
3. INFORM	Provide information about the content (ex. Evaluweb which displays a banner on the web pages)
4. MONITOR	Record for later inspection a list of the content accessed or attempted to be accessed by a user; this may be a complete list or include only the content deemed inappropriate for children (ex in Net Nanny)
5. WARN	Provide information about content and recommend against accessing that content before it displayed
6. BLOCK	Prevent children from accessing content [ex. Microsoft Internet Explorer (which filters on the basis of any PICS labels), Cyber patrol, Net Nanny,...]

1.2. Possible locations of control/filtering

Table 1.2 . 6 different locations of control/filtering (adapted from Cranor & Resnick, 12 March 1998)

1. Personal computer	ex. Cyber Patrol, SurfWatch, Net Shepherd, MS Internet Explorer
2. LAN or local proxy server	ex. Cyber Patrol, SurfWatch, SafeSurf Internet Filtering Solution N.B. « harder for individual to tamper with »
3. Internet Service Provider	ex. SurfWatch, SafeSurf Internet Filtering Solution ¹⁰
4. Remote proxy server	ex. Bess
5. Search engine	ex. AltaVista engine at Net Shepherd
6. Web site	ex. EvaluWEB providing both PICS labels and banner graphics

⁹ Adapted from <http://www.research.att.com/~lorrie/pubs/tech4kids/t4k.html>

¹⁰ See <http://www.safesurf.com/ssfaq.htm#basic>

1.3. Control scopes and corresponding possible technical solutions

Table 1.3. Control scopes and corresponding possible technical solutions

Control general scopes	Control specific scopes	Possible current technical solutions
1. Control at the level of the entry point to an address or a file	1.1. <u>anything with an URL</u> ¹¹ i.e. for example: <ul style="list-style-type: none"> • WWW (HTML Protocol) • FTP (File transfer Protocol) • GOPHER (for information research) • Usenet News groups and individual messages (NNTP Protocol) • TELNET (for terminal access) • [IRC Internet Relay Chat] (in the near future) (N.B. e-mail messages do not have an URL) 	PICS labelling and filtering
	1.2. <u>local or on-line applications</u> ex. games, personal financial managers, etc.)	filtering on the basis of lists of <ul style="list-style-type: none"> • URL or • names of newsgroups, chat etc.
2. Control at the level of the content itself	2.1. <u>Incoming information</u> N.B. for example for e-mail (including their attachments)	<ul style="list-style-type: none"> • Key word/string filter (or word-matching) • Intelligence artificial based software
	2.2. <u>Outgoing information</u> (for ex. personal information via IRC, website questionnaire, e-mail, ... or offensive words in search of sexually explicit sites or conversations)	
3. Time control	3.1. Hours/day	
	3.2. Days/week	
	3.3. Total by week	

2. Examples of current filtering software packages and corresponding control scopes

A control software package can include several scopes and several technical solutions.

Table 2. Control scopes and 2 examples of current control software packages

¹¹ Adapted from Resnick, latest updated 98/01/03 Platform for Internet Content Selection (p. 10).

	Cyber Patrol 4.0 ¹²	Net Nanny ¹³ 3.0
1.1 <u>anything with an URL</u> ¹⁴ i.e. for example: <ul style="list-style-type: none"> • WWW (HTML Protocol) • FTP (File transfer Protocol) • GOPHER (for information research) • Usenet News groups and individual messages (NNTP) • TELNET (for terminal access) • [IRC Internet Relay Chat] (in the near future) 	<ul style="list-style-type: none"> • Yes (site level and page level) • Yes • Yes • Yes • Yes • Yes 	<ul style="list-style-type: none"> • Yes • Yes • ? • Yes • ? • Yes
1.2. <u>local or on-line applications</u> (ex. games, personal financial managers, etc.)	<ul style="list-style-type: none"> • Yes 	<ul style="list-style-type: none"> • Yes (non Internet BBS etc.)
2.1. <u>Incoming information</u> N.B. for example for e-mail (including their attachments)	<ul style="list-style-type: none"> • NO (?) 	<ul style="list-style-type: none"> • Yes (e-mail, chat rooms etc.)
2.2. <u>Outgoing information</u> (for ex. <i>personal</i> information via IRC, website questionnaire, e-mail,...or <i>offensive</i> words in search of sexually explicit sites or conversations)	<ul style="list-style-type: none"> • Yes (via IRC or website questionnaire or words in search of sexually explicit sites or conv.) 	<ul style="list-style-type: none"> • Yes (e-mail, chat rooms etc.)
3.1. Hours/day	<ul style="list-style-type: none"> • Yes 	<ul style="list-style-type: none"> • NO
3.2. Days/ week	<ul style="list-style-type: none"> • Yes 	<ul style="list-style-type: none"> • NO
3.3. Total by week	<ul style="list-style-type: none"> • Yes 	<ul style="list-style-type: none"> • NO

¹² See The CyberPatrol Fact sheet at <http://www.cyberpatrol.com/cyber/fact.html> and Cyber Patrol Online Demonstration tour at http://www.cyberpatrol.com/cp_demo/default.html

¹³ See F.A.Q. at http://www.netnanny.com/NNFAQ/nnfaq_bottom.html

¹⁴ Adapted from Resnick, latest updated 98/01/03 Platform for Internet Content Selection (p. 10).

3. PICS Rating Services

The rating services included in tables of § 3 are listed (with their hyperlinks) at <http://www.3org/PICS/raters.html>¹⁵. According to our vocabulary, 'levels' inside a labelling category can be used for setting filtering criteria. But 'sub-categories' or 'examples' allow to explain the whole meaning of a category but cannot be used for setting filtering criteria.

3.1. PICS Third-Party Rating Services

Table 3.1.1. PICS Third-Party Rating Services : Who has set labelling vocabulary and criteria for assigning labels ?

CyberNOTlist	evaluWEB	Net Shepherd's Rating (of websites, FTP, NNTP and IRC protocols)	SurfWatch 's Rating (for kids)
linked to Cyber Patrol ¹⁶ , a Microsystems software	no link with any filtering software	linked to the filtering software 'Net Shepherd' ¹⁷ of NetShepherd Inc.	Linked to 'SurfWatch', the Spyglass' filtering software
<i>Microsystems' CyberNOT Oversight Committee</i> , « a diverse group of advisors from a wide range of civic organizations representing all parts of the social and political spectrum »	free (and experimental) system online developed by Sserv (?) ¹⁸ « presenting itself as 'the web's leading daily starting point for news, web search and what's happens on line »	Net Shepherd, « a Calgary based Internet company » « employing virtual Internet communities for relevant information retrieval, increased subscriber loyalty & enhanced e-commerce performance » ¹⁹	« Developed and regularly reviewed by SurfWatch's community-based Advisory Committee » ²⁰ which « is a group of community members ranging in age from their 20s to their 50s. The group includes teachers, parents, publishing professionals, members of the clergy, corporate human resource managers, drug and alcohol prevention counsellors and child abuse prevention agents. They represent diverse ethnic and religious backgrounds, and have in common an interest and passion for the issues that surround filtering content on the Internet. » ²¹

Table 3.1.2. PICS Third-Party Rating Services : Who is in charge of assigning labels ?

¹⁵ Except for ORC (Objective Ratings Criteria) mentioned in the list of the webpage but not found on the Web.
¹⁶ It must be noticed that Cyber Patrol has its own criteria of ratings but also supports the RSACi's rating standard.
¹⁷ See <http://www.zdnet.com/familypc/9709/noway/table.html>
¹⁸ This information has to be confirmed.
¹⁹ See <http://www.netshepherd.com/Corporate/Presreport.html> (dated June 15 1998).
²⁰ See <http://www.surfwatch.com/filteringcriteria/> (visited on 11 July 1998).
²¹ See http://www1.surfwatch.com/submit/faq_index.html

Cyber Patrol (4.0) (CyberNOTlist)	evaluWEB	Net Shepherd's Rating	SurfWatch (for kids)
<p>« Cyber Patrol's proprietary CyberNOT list is selected and continuously reviewed by a skilled research team that works full-time to identify sites inappropriate for kids. Researchers use specialised sorting software to remove previously viewed sites and scan the remaining sites, displaying 24-character phrases that include suspect words as 'sex' or 'satanic'. This helps eliminate thousands of innocuous sites. The identified sites are reviewed by the research staff and sorted into 12 categories.... » (see next table)</p>	<p>artificial intelligence software (which does <i>not</i> take into account the context) + evaluation by a human reviewer if needed</p>	<p>labels assigned by third party communities composed of end-users (> 1200 volunteers) who vote for the rating of a site (→ <i>World Opinion's Service Data Base</i> with more than 500 000 ratings opinions) « A Rating Community is a virtual, on-line community of people who represent the <i>Internet's general population.</i> »</p>	<p>« SurfWatch employs a team of skilled Internet surfers to locate questionable sites. Each sit identified by our surfers is also evaluated by SurfWatch staff members and may also be reviewed by an Advisory Committee of parents, teachers and community members. This 'eyes on page' content evaluation is supplemented by our pattern blocking technology which detects words in URLs, chat, or newsgroup names that indicate inappropriate content »²² « The SurfWatch Surfer team contains individuals ranging in age from their 20s to their 40s. The surfers live all over the United States, including California, Indiana, Michigan, Illinois, Texas, Utah, Pennsylvania and Washington, and abroad in <i>Japan</i>. They work at all hours of the day and night to catch sites that may only operate during certain times of the day. They are independent contractors whose primary occupations include airline pilot, mystery novelist, teacher, homemaker and accountant. »</p>

²² See <http://www.surfwatch.com/filteringcriteria/index.html> (visited on 30 January 1998).

Table 3.1.3. PICS Third-Party Rating Services : breaking down of the labelling categories

Cyber Patrol (4.0) (CyberNOTlist) 11/5/97	evaluWEB	Net Rating	Shepherd's	SurfWatch (for kids) = core
Violence/ Profanity (Graphics or Text)				Violence (6 sub-cat.)
Partial Nudity				
Full Nudity				
Sexual Acts (G/T)				Sexually Explicit (7 s-c.)
Sex Education (G/T)				
Gross Depictions (G/T)				
Intolerance (G/T)				Hate Speech (4 sub-cat.)
Satanic/Cult (G/T)				
Drugs/Drug Culture (G/T)				see below
Militant/Extremist (G/T)				
Questionable/Illegal & Gambling (G/T)				Gambling (5 sub-cat.)
Alcohol & Tobacco (G/T)				Drugs/Alcohol (7 sub-c.)
[+ Sports & Entertainment (cf. soft. for firms)]				
	Age based: 3 levels: <ul style="list-style-type: none"> • GV = general viewing (suitable for all ages) • PA13 = Parental Advisory (unsuitable for young children) • XC = explicit content (intended only for legal adult) 	Age based: 6 levels: <ul style="list-style-type: none"> • General • Child • Pre-teen • Teen • Adult • Objectionable 		
		Quality : 5 levels: <ul style="list-style-type: none"> • poor • fair • good • very good • excellent 		
				ChatBlock (all)

Table 3.1.3.bis

PICS Third-Party Rating Services : breaking down of the labelling categories

SurfWatch ⁷ profession edition ²³ or Productivity category criteria ²⁴
astrology/mysticism (4 sub-categories)
entertainment (4 sub-categories)
games (4 sub-categories)
general news (4 sub-categories)
glamour or intimate apparel (4 sub-categories)
hobbies (4 sub-categories)
investments (4 sub-categories)
job search (3 sub-categories)
travel (3 sub-categories)
motor vehicles (4 sub-categories)
personals or dating (4 sub-categories)
real estate (3 sub-categories)
shopping (4 sub-categories)
sports (4 sub-categories)
Usenet News (all)
ChatBlock (all)

325

²³ See <http://www.news.com/News/Item/0%2C4%2C21751C00.html?sas.mail>

²⁴ <http://www1.surfwatch.com/filteringcriteria/productivity>

3.2. PICS Self-Rating Services

Table 3.2.1. Current PICS Self Rating Services: Who has set labelling vocabulary and criteria for assigning labels?

Adequate.com ²⁵ (provides a visual cue) (to be installed in Microsoft's Internet Explorer)	IT-RA (Laprima Rating Agency Italiana)	RSACi (i stands for Internet) (integrated in Microsoft's Internet Explorer 3.0 - as the <i>option by default</i> - and in CyberPatrol) (N.B. Nominated for the Carl Bertelsmann Prize 1998)	Safe For Kids ²⁶ (N.B.: no certification mark)	SafeSurf's Internet Rating Standard (integrated in Microsoft's Internet Explorer and Netscape ²⁷)	Vancouver Webpages Rating Service
no link with any filtering software	no link with any filtering software	no link with any filtering software	no link with any filtering software	initially linked to SafeSurf Internet Filtering Solution (no longer available)	no link with any filtering software
free system developed by Adequate.com (entertainment web sites provider firm)	developed by the University of Bologna in collaboration with GENITORI (discussion list) of the 'Citta Invisible'	developed by the Recreational Software Advisory Council (RSAC), « an independent, non-profit organisation based in Washington DC » (which initially performed a computer game rating)	developed by a <i>British physicist</i> (weburbia.com) without commercial goal but « to provide a simpler system of rating pages which did not require me or others to answer complicated ill posed questions and register our answers with SafeSurf or RSACi » ²⁸	voluntary system - « developed with input of thousands parents & netcitizens world-wide » by SafeSurf, « a no-fee parents' organization » « dedicated to making the Internet safe for your children without censorship ²⁹ » (and founded by 2 men out « of concern for children accessing adult material »)	experimental service developed by Vancouver-Webpages.com (a part-time small business) The categories were created by one individual. ³⁰

Table 3.2.2. PICS Self Rating Services : breaking down of the labelling categories

Adequate.com ³¹	IT-RA	RSACi	Safe For Kids	SafeSurf's Internet Rating Standard	Vancouver Webpages Rating Service
-----------------------------------	--------------	--------------	----------------------	--	--

²⁵ <http://www.adequate.com>

²⁶ <http://www.weburbia.com/safe/>

²⁷ <http://www.safesurf.com/press> (June 17 1998).

²⁸ Cf. e-mail dated 13 July 1998 from Phil Gibbs.

²⁹ « Proof that the Internet community will implement solutions on its own can convince them [legislators] of alternative solutions to censorship. »

³⁰ Cf. e-mail dated 17 July 1998 from Andrew Daviel.

³¹ <http://www.adequate.com>

see age	Violenza (5 levels)	Violence (5 levels)		SS--005 Violence (9 levels)	Violence (6 levels)
				SS--001 Profanity (9 levels)	Profanity (5 levels)
		Nudity (5 levels)		SS--004 Nudity (9 levels)	
see age	Sesso (5 levels)	Sex (5 levels)		SS--006 Sex, Violence & Profanity (9 levels)	Sex (8 levels)
				SS--002 heterosexual Themes (9 levels)	
				SS--003 Homosexual Themes (9 levels)	
	Razzismo (5 levels)			SS--007 Intolerance (9 levels)	Tolerance (7 levels)
				SS--008 Glorifying Drug Use (9 levels)	
				SS--00A Gambling (9 levels)	Gambling (4 levels)
Age based: <ul style="list-style-type: none"> • Web Y = all children • Web Y7 = older children (fantasy violence) • Web G = general audiences • Web PG = parental guidance suggested (moderate violence/sexual situations/ infrequent coarse language/ suggestive content) • Web 14 = parental guidance strongly suggested (intense violence/ intense sexual situations/ strong coarse language/ intensely suggestive content) • Web MA = mature audiences only (graphic violence/ explicit sexual situation/ crude 			Age: 0. Safe for kids 1. Parental guidance (9 examples) 2. Adults only	SS--000 Age Range (9 levels)	

indecent language					
				SS--009 Other Adult Themes (9 levels)	
see age	Linguaggio (5 levels)	Language (5 levels) (including profanity and hate speech)			
	Pubblicità (5 levels)				
	Religione (5 levels)				
	Politica (5 levels)				
	Didatticità (5 levels)				Educational Content (5 l.)
					Multiculturalism (6 l.)
					Environmental Awareness (6 levels)
					Safety (5 levels)
					Canadian Content (3 l.)
					Commercial Content (4 l.)

3.3. Some General Remarks and Conclusions about these Observations

3.3.1. General Remarks

Criteria for assigning labels which allow *the more choices and the greater adaptability at the customization level* are with

- the more categories
- the more levels inside each category
- not only age-based categories (so as to let to the person in charge of setting the criteria the possibility to judge by himself/herself what is convenient for a given age) .

Criteria can be *more or less subjective* (and thus possibly inconsistent):

- age-based categories are rather subjective and linked with culture
- quality categories are also subjective
- the more levels inside a category, the more nuances and thus the more risks of subjectivity.

3.3.2. Conclusions about these Observations

Following comments can be made from the tables of § 3 and, of course, concern only the ten analysed ratings. (Unfortunately, available information about evaluWEB are insufficient to be used here except for the breaking down of labelling categories.)

Several conclusions can be drawn first **about who fill the different roles and about the link between rating and the commercial sector** .

- Some ratings (CyberNOTlist, Net Shepherd's Rating, SurfWatch's rating and SafeSurf Internet Rating Standard) have been linked, at least initially, to the development of a given filtering software (respectively by Microsystems[U.S.], the Canadian company NetShepherd, Spyglass [U.S.] and SafeSurf [U.S.]).
- Firms either ask 'representative' committee to set labelling vocabulary and criteria for assigning labels (for example, Microsystems, Spyglass/SurfWatch) or do it themselves (for example, NetSepherd and Adequate.com [U.S. entertainment web sites provider firm]). The case of SafeSurf seems special: it will no longer be offering the SafeSurf Internet Filtering Solution and currently present himself as a no-fee parents' organization which has developed a rating standard.
- The two firms in charge of defining criteria and vocabulary for assigning labels do not realise themselves the labelling: Net Shepherd lets labels assigned by third party communities composed of end users and have built a very famous database of ratings opinion and Adequat.com provides self rating service.
- On the other hand, four self ratings services (the Italian IT-RA, the U.S. RSCA and the U.K. Safe For Kids and Vancouver Webpages' Rating) have not been developed by firms but by non-profit organisations (the first two) or an individual (the last two).

From a **cultural point of view**, let us notice that the great majority of the ratings is:

- *in English* but one is written in Italian

- *with criteria defined in U.S.* except for Net Shepherd's Rating and Vancouver Webpages Rating Service (both made in the Western part of Canada), Safe For Kids (made in U.K.) and, of course, IT-RA (made in Italy). Let us add about third party ratings services that only *U.S.* people intervene to *assign labels* in the case of the CyberNOTlist and evaluWEB but the SurfWatch Surfer team includes also Japanese people. And let us notice that the Rating Communities of Net Shepherd « represent the Internet's general population »!

About the **rating categories**, let us make the following conclusions:

- Only the self rating services provide *levels* inside categories other than age-based categories.
- *Age based (or quality) ratings* are provided either by commercial firms (Adequate.com and Net Shepherd) or by non-profit organization.
- One of the ratings allowing the most choices for the setting (for example, by parents) is provided by a no fee *parents'* organization (SafeSurf).
- The most frequent labelling categories are: sex, violence, and then age.

And let us as make a suggestion: *Why not include in the categories specific labels for child pornography and calls to violence or hate in order to really and efficiently always censor it (already at the level of ISP)?* This suggestion is in accordance with the conclusions of MAPI (see : <http://www.info.fundp.ac.be/~mapi/mapi-eng.html>) but is not made by CPSR (Computer Professional for Social Responsibility) (1997), for example, nor any government so far!

3.4. Generalisation : What could be rated with PICS?

« PICS labels can describe any aspect of a document of a Web site... »: indecency (cf. current labels), privacy, intellectual property, subject categories, reliability of information from a site, software's safety (« to help protect computers from exposure to viruses ») (see Resnick 1997).

4. Ethical Issues and Questions

Ethical issues with filtering software will be discussed from the breakdown of table 0. Let us first remark that 'set labelling vocabulary and criteria for assigning labels' is not value-neutral and that 'assign labels' and 'set filtering criteria' can imply moral judgements.

4.1. Outside PICS

Outside PICS, it happens that several roles (particularly 'set labelling vocabulary', 'assign labels', 'write filtering software' and even 'set filtering criteria') are filled by one sole commercial entity. Ethical issues are obvious with this kind of software: users are linked to the subjective value judgements of this firm!

4.2. With PICS

With PICS, as explained above, the 6 roles can be filled by different entities. This can improve the situation from an ethical point of view but cannot delete any issue.

4.2.1. 'Set labelling vocabulary and criteria for assigning labels'

- First of all it must be noticed that to set labelling vocabulary and criteria for assigning labels is a really important role : as pointed out by CPSR (1997) « in general, the use of a filtering product involves an implicit acceptance of the criteria used to generate the ratings involved. » *Who is in charge of this role?*

4.2.2. 'Assign labels'

- *Who is in charge of the very sensitive role of assigning labels ?*
- *Which of the two approaches (self-rating and third party rating) is the best ?*
- When a third party rating service is involved, the next questions are to be raised :
Who is effectively represented by this third party ? Is this party representative of a values-oriented organisation or of a given population ?
How are the ratings attributed ?
- With self rating, the questions are :
How to oblige or, at least, to incite people to self rate ?
And on the basis of which principle ?
How to solve the problem of mislabelled pages, and particularly of deliberately mislabelled pages ?
[As suggested by Resnick (Technology inventory), « The Internet community will need to co-operate in the creation of either vouching services, which vouch for authors who are honest in their self-labelling, or blacklisting services which keep track of authors whose labels are not reliable. »]

- If people assign labels and if labelling is not compulsory all over the world, it is obvious that many sites will stay unlabelled. The question is then : *What to do with unlabelled sites* ? If the software allow unrated sites, then the global control will not be efficient but if it does not, then innocuous and very interesting sites will be not accessible. (see the discussion of Weinberg on this subject). And thus in this case, « blocking software could end up blocking access to a significant amount of the individual, idiosyncratic speech that makes the Internet a unique medium of mass communication. Filtering software, touted as a speech protective technology, may instead contribute to the flattening of speech on the Internet. » (Weinberg 1997)

4.2.3. 'Write filtering software'

- *Who is in charge to decide:*
 Which kind of control/filtering ? (see the 6 possibilities [suggest, search, inform, monitor, warn, block] in table 1.1)
 Which location ? (see table 1.2)

4.2.4. 'Set filtering criteria' (= customize)

- *Which kind of customization* ? In fact there is a dilemma : the more choices you give to the final users the more difficult it is to set ! A solution in the future could be, as suggested by Resnick (Technology Inventory), the use of PICSRules (a format for writing filtering settings).
- *Who is in charge of this role?* Initially this role was dedicated to parents to control their children. But filtering software are used also by libraries (in U.S.), for instance to control adults (see Weinberg).
- *Is it ethically justified to give such a power of control to this kind of entity ?*

Sources

Parental control SW comparisons

- Munro C. 1997, Internet Filtering Utilities, *PC Magazine*, April 8 1997, pp. 235-240.
- Parental Control Software at a Glance, October 97 issue of *FamilyPC*
<http://www.zdnet.com/familypc/9709/noway/table.html>
- Meeks Ch., 8 programs to porn-proof the Net, 4/3/96; updated 5/28/97
<http://www.cnet.com/Contenu/Reviews/Compare/Safesurf>

Synthesis:

- Cranor L.F. & Resnick P., Technology Inventory, 12 March 1998, <http://www.research.att.com/~lorrie/pubs/tech4kids/t4k.html>
- Resnick P. 1997, Filtering Information on the Internet, *Sci. Amer.* 03 97, <http://www.sciam.com/0397issue/0397resnick.html>
- Resnick P. (last updated) 1998 01 03, Platform for Internet Content Selection (PICS), <http://www.w3.org/PICS/>
- Resnick P. PICS Third-Party Rating Services (last update March 5, 1996), <http://www.w3.org/PICS/raters.htm>
- Resnick P. last revised 26 01 1998, PICS, Censorship & Intellectual Freedom FAQ, <http://www.w3.org/PICS/PICS-FAQ-980126.html>
- Resnick P. & Miller J., 1996, PICS: Internet Access Controls without Censorship, *CACM* 39 (10): 87-93, October 1996.
- Weinberg J. 1997, Rating the Net, <http://www.msen.com/~weinberg/rating.htm>
- CPSR 1997, Filtering FAQ, Version 1.1, 12/25/97, written by Hochheiser H., <http://quark.cprs.org/~harryh/faq.html>
- MAPI (Mouvement Anti-Pédophilie sur Internet) <http://www.info.fundp.ac.be/~mapi/mapi-fr.html>

Web sites with specific technology description

<http://www.cyberpatrol.com>

<http://www.microsys.com/cyber/fact.html>

<http://www.sserv.com/evaluweb/pics.html>

<http://www.sserv.com/evaluweb/ratings.html>

<http://www.surfwatch.com>

<http://www1.surfwatch.com> : Surfwatch PICS Label Bureau

<http://www.news.com/News/Item/0%2C4%2C21751C00.html?sas.mail:Lipton,B.>, Software to filter workers' access, 4 May 1998

<http://www.safesurf.com>

<http://www.rsac.org>

<http://www.netshepherd.com> (November 97)

<http://www.ratings.netshepherd.com>

<http://www.adequate.com>

<http://www.citinv.it/>

<http://www.rsac.org/>

<http://www.weburbia.com/safe/ratings.html>

<http://vancouver-webpages.com/VWP1.0/>

Acknowledgements

The author thanks Jacques BERLEUR for his comments and suggestions.