Legal Implications of Internet Filtering

LLM Dissertation

Peter A. Craddock

Supervisor: K. Madders

LLM specialism: Regulation & Technology
# Table of Contents

**Introduction** 1

I. **Historical, technical and factual analysis of Internet filtering** 2
   1. The Yahoo! case, an international debut for filters 2
   2. Internet filtering typology 3
      2.1. Filtering methods 3
      2.2. Level of deployment 4
      2.3. Resources targeted 5
      2.4. Interaction with Internet traffic 6
   3. Internet filtering worldwide 6
      3.1. Filters imposed by the State 6
      3.2. Voluntary or State-encouraged ISP-level filters 11
      3.3. Company, library, school and home filtering 13

II. **Legal analysis of filters** 15
   1. Freedom of expression 15
      1.1. Applicability of the right to freedom of expression 16
      1.2. ‘Prescribed by law’ 18
         (i) **Law** 18
         (ii) **Accessibility** 20
         (iii) **Forescability** 21
      1.3. ‘Necessary in a democratic society’ 21
         (i) ‘Pressing social need’ 22
         (ii) ‘Relevant and sufficient’ 23
         (iii) ‘Proportionate to the legitimate aim pursued’ 24
            a) ‘Appropriate and necessary’ 24
            b) ‘Least onerous’ 26
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.</td>
<td>Liability</td>
<td>27</td>
</tr>
<tr>
<td>2.1.</td>
<td>Exclusion of liability</td>
<td>27</td>
</tr>
<tr>
<td>2.2.</td>
<td>Prohibition of any general obligation to monitor</td>
<td>29</td>
</tr>
<tr>
<td>2.3.</td>
<td>Liability for removal of legitimate content</td>
<td>30</td>
</tr>
<tr>
<td>3.</td>
<td>Privacy and data processing</td>
<td>31</td>
</tr>
<tr>
<td>4.</td>
<td>Legitimacy in governance: transparency</td>
<td>31</td>
</tr>
<tr>
<td>III.</td>
<td>Filters mindful of the law</td>
<td>34</td>
</tr>
<tr>
<td>1.</td>
<td>Freedom of expression</td>
<td>34</td>
</tr>
<tr>
<td>1.1.</td>
<td>‘Prescribed by law’</td>
<td>34</td>
</tr>
<tr>
<td>1.2.</td>
<td>‘Necessary in a democratic society’</td>
<td>34</td>
</tr>
<tr>
<td>2.</td>
<td>Liability</td>
<td>35</td>
</tr>
<tr>
<td>3.</td>
<td>Privacy and data processing</td>
<td>35</td>
</tr>
<tr>
<td>4.</td>
<td>Legitimacy and concluding remarks</td>
<td>36</td>
</tr>
<tr>
<td>Epilogue</td>
<td></td>
<td>36</td>
</tr>
<tr>
<td>Bibliography</td>
<td></td>
<td>37</td>
</tr>
<tr>
<td>1.</td>
<td>Articles &amp; studies</td>
<td>37</td>
</tr>
<tr>
<td>2.</td>
<td>Cases</td>
<td>43</td>
</tr>
<tr>
<td>2.1.</td>
<td>Supra-/International courts</td>
<td>43</td>
</tr>
<tr>
<td>2.2.</td>
<td>National courts</td>
<td>44</td>
</tr>
<tr>
<td>3.</td>
<td>Legislation, treaties &amp; declarations</td>
<td>44</td>
</tr>
<tr>
<td>3.1.</td>
<td>Council of Europe</td>
<td>44</td>
</tr>
<tr>
<td>3.2.</td>
<td>European Union</td>
<td>45</td>
</tr>
<tr>
<td>3.3.</td>
<td>National legislators</td>
<td>45</td>
</tr>
</tbody>
</table>
Introduction

Alexander turns on his computer, smiling as he hears the familiar chime. Colours fill the screen, a feast for his eyes, and he clicks on an icon, his gateway to the Internet.

As he submits two words to a search engine and chooses the first result, as if advised by an old friend, Alexander is unaware of the underlying processes.

His computer converses with a network provider and asks whether Alexander may access the website. The network provider turns to a domain name server, to find out on which server the website is located, before finally connecting to the hosting provider to obtain transmission of the website data.

Meanwhile, Alexander blinks. The Internet must be unhappy with him: he is denied access to the website. Alexander sighs, and goes back to the search results. He does not pause to consider whether access was blocked rightfully or whether this limits his freedom; he does not even contemplate complaining to anyone. After all, it’s the Internet, and he doesn’t understand it. How could he, a normal web user?

In this simplified tale of daily Internet use¹, Alexander is confronted with access denial to a website that appeared in search results. As he shares the general population’s lack of understanding of the technology underlying the Internet and the World Wide Web², he does not know why the information embodied in the website is not being transmitted to him. He is unable to assess whether the problem lies with the website owner or with any of the intermediaries between him and the website.

One possible explanation may, however, spring to the mind of an observer with some degree of technical knowledge: this access denial may come from a filter.

In this paper, we shall analyse the legal implications of the use of filters for blocking access to Internet content.

Prior to the legal analysis of any technological apparatus, however, one must examine it from a technical standpoint. In a first part, after a brief history of filters, we shall therefore lay out a typology of Internet filters and examine their use throughout the globe.

In a second part, we shall assess the legality of filtering, in both general and specific situations, notably through the lens of fundamental rights and freedoms, of data protection and of liability.

Finally, drawing lessons from our legal analysis, we shall conclude by proposing a new framework for filters, one that is mindful of the legal environment.


² Chris Reed states that ‘after all, the Internet is one of the most “technical” phenomena around’. Ibid., 3.
I. Historical, technical and factual analysis of Internet filtering

1. The Yahoo! case, an international debut for filters

In April 2000, Mark Knobel, acting on behalf of the International League against Racism and Anti-Semitism (Ligue Internationale Contre le Racisme et l’Antisémitisme), filed a complaint against Yahoo!, then known as the ‘Lord of the Portals’.

Yahoo! offered a number of services on its yahoo.com and yahoo.fr websites, notably auction services, and Mark Knobel had been shocked to find Nazi memorabilia on the yahoo.com auction listings, which were accessible to French citizens. As French criminal law prohibits the public display of Nazi memorabilia, the Parisian Tribunal de Grande Instance ordered in May 2000 that Yahoo! ‘restrict access to listings involving Nazi memorabilia to any and all French citizens’.

Yahoo!, however, did not comply with the order, putting forward the following reasons:

‘there are no technical means capable of satisfying the terms of the order […];

on the assumption that such means existed, their implementation would entail unduly high costs for the company […] and would to a degree compromise the existence of the Internet, being a space of liberty and scarcely receptive to attempts to control and restrict access’.

To these arguments, the plaintiffs responded by discussing geo-location technology, which they claimed Yahoo! already implemented, as ‘Yahoo auctions in France […] were not in fact coming from servers in the United States [but] from Swedish servers’.

The Court assembled a panel of experts, which included Vinton Cerf, often called the ‘father of the Internet’. These experts reported that geo-location services allowed accurate targeting of 70% of Internet users in France, and that geo-location combined with declarations of nationality would allow for sufficiently effective means of targeting French citizens.

---


7 Goldsmith & Wu (n 3), 7.
On 20 November 2000, presiding judge Gomez took the panel’s report into account and stated that ‘it is possible to determine the physical location of a surfer from the IP address’.

The Court therefore found in favour of the plaintiffs. Although Yahoo! protested at first, it surrendered and ‘pulled all Nazi materials from its auction sites’ on 2 January 2001.

While the Yahoo! case raised concerns about the extra-territorial effectiveness of national judgments, it represented a paradigmatic shift in the conception of the Internet: where the Internet was deemed to be free of national constraints at the dawn of the digital age, Yahoo! showed that national laws of censorship could still apply.

Although filtering was not necessarily a new phenomenon, the Yahoo! decision cleared the path for the territorial filtering of Internet content.

2. Internet filtering typology

Filters come in many shapes and sizes, with different characteristics and consequences. In their report of the filtering of child pornography on the Internet in the Netherlands, commissioned by the Dutch government, Wouter Stol and fellow researchers crafted a comprehensive typology of Internet filters that will form the basis for our own.

In this typology, filters vary according to method, level of deployment, resources targeted and interaction with Internet traffic.

2.1. Filtering methods

Internet filtering software generally depends on two criteria:

---

8 Supra n 6.
9 Goldsmith & Wu (n 3), 8.
10 See Edwards (n 5), 626-627.
11 Ibid., 625.
'(a) whether the program detects banned keywords and other prohibited content on the page contents
(“dynamic filtering”) and (b) whether the site is on a database of sites to be blocked […] (“blacklist
filtering”)'.

As pointed out by Wouter Stol et al., blacklist filtering has the disadvantage of being incapable of
adapting automatically to new information appearing on the Internet, as blacklist filtering is based on
human review of websites. This entails a second disadvantage:

‘composing a blacklist on the basis of human review is labour-intensive, because the supply of information
on the internet changes continually, and […] that same information may be supplied from varying places’.

Dynamic filtering, on the other hand, is automated, based on keywords or phrases (in the case of
text), or on visual characteristics (to encompass images, for instance). The primary disadvantage of
dynamic filtering is the higher risk of blocking content that should not be blocked.

Blacklist filtering thus risks ‘underblocking’ to a greater extent than dynamic filtering, while the
reverse is true for the risk of ‘overblocking’. This shall be analysed in more detail when we examine
the effectiveness of filters.

It is worth mentioning that a combination of both methods is possible, through automated blacklist
management. Such blacklist filtering will present the same risk characteristics as dynamic filtering.

2.2. Level of deployment

Filters are generally deployed at one of four levels: at national level, at Internet service provider (ISP)
level, at server level or at end-user computer level.

The choice of the level may depend on the structure of the Internet backbone within a given
territory. For instance, national-level filters are a possibility only if one party (be it a country or
another kind of agent) has effective control over the architectural backbone of the Internet within a
national territory; ISP-level filters are more likely in the absence of concentrated power over the
underlying architecture.

In our typology, the concept of ‘ISP’ covers several parties, by reference to the E-Commerce
Directive of the European Union (EU), in which ‘service provider’ is defined as ‘any natural or legal
person providing an information society service’.

---

15 Haselton (n 12), 1.
16 Stol & others (n 14), 253.
services, in particular electronic commerce, in the Internal Market (‘Directive on electronic commerce’) (‘E-Commerce
Activities covered include the simple transmission of information or provision of access ('mere conduit'\textsuperscript{19}), the ‘automatic, intermediate and temporary storage’ of such information ('caching'\textsuperscript{20}) and its longer-term storage ('hosting'\textsuperscript{21}). While the linking to information (as search engines do) was not listed in the E-Commerce Directive, we consider such an activity to be included in the concept of ‘ISP’. Filtering at ISP-level may therefore concern filters placed by a number of different Internet operators, from access providers to website hosting providers (even website operators, as illustrated by the Yahoo! case).

Server-level filtering denotes filtering at the level of a server shared by several users to access the Internet, such as workplace or university networks.

Finally, computer-level filtering refers to filters that are specific to users of one individual computer. Parental control software on family computers is one such filter.

### 2.3. Resources targeted

Wouter Stol \textit{et al.} state that blacklists may block content on the basis of four resources: Internet Protocol addresses (IP addresses), domain names, Uniform Resource Locators (URLs) and hash codes\textsuperscript{22}. This distinction indicates the range of blocking.

IP addresses are the addresses of servers on which the content is hosted. As web hosting providers generally host several (thousand) websites on one server, blocking IP addresses may lead to the blocking of a broad range of Internet content.

Domain names are a more precise means of targeting content, as these resources concern one website.

Several websites, however, merely act as hosts for many ‘sub-websites’, notably since the advent of social networking websites and blogging platforms\textsuperscript{23}. URL blocking therefore allows the filter to target specific items on a specific domain.

Hash codes are the result of the conversion of data, by means of a mathematical function, to short integers or strings of characters\textsuperscript{24}. When processed through a ‘perfect’ hash function, an image will produce a specific string of characters that will be unique. Targeting hash codes thus enables filters to find images that are the exact reproduction of a blocked image, regardless of URL.

\textsuperscript{19} E-Commerce Directive (n 18), art 12.
\textsuperscript{20} E-Commerce Directive (n 18), art 13.
\textsuperscript{21} E-Commerce Directive (n 18), art 14.
\textsuperscript{22} Stol \& others (n 14), 253.
\textsuperscript{23} Thus, domains such as \url{wordpress.com}, \url{blogger.com}, \url{facebook.com} and \url{youtube.com} host a myriad of types of content.
While we agree with Wouter Stol et al. that blacklists operate in such a way, dynamic filtering also targets some of these resources (notably hash codes). This distinction is therefore important for the typology as a whole.

2.4. Interaction with Internet traffic

The final element of the typology, the nature of the filter, concerns the ‘ways in which filters are embedded in internet traffic’.

Filters are said to interact with Internet traffic mainly in two ways: they can act as part of the Domain Name System (DNS), or they can act as proxy servers.

In the case of a DNS filter, usually linked to a blacklist, the filter acts as a domain name server, an intermediary between the web browser and websites. When a user attempts to reach a website, the filter finds the IP address for the website’s domain name. The filter checks whether one of the resources is on the blacklist. Should it find a match, it will give an incorrect IP address to the browser (such as one of a ‘stop page’, or one that does not exist).

Proxy servers, also known as proxies, play a different role: a proxy is a separate computer that acts as an intermediary between a (local network of) computer(s) and the Internet, and all Internet traffic to and from the (local network of) computer(s) therefore goes through the proxy. A filter may be based on a simple proxy system, and thus directly check every Internet-related request as it goes through the proxy server, or may be based on a two-step proxy:

‘All internet traffic is led through a proxy first onto which a filter has been installed that basically discerns, i.e. on the basis of an [IP address or domain name], between suspected and unsuspected information flows. The suspected flows are diverted to a second proxy with a filter that checks on a detailed level ([i.e. on the basis of the other resources]) what can be admitted or not.’

3. Internet filtering worldwide

Internet filtering has become a widespread phenomenon, and we shall sketch the map of worldwide Internet filtering on the basis of the agent ordering the placing of filters.

3.1. Filters imposed by the State

On 12 March 2010, ‘World Day Against Cyber Censorship’, Reporters Without Borders (Reporters sans frontières) published their updated list of ‘Enemies of the Internet’ and of ‘Countries under
surveillance. The ‘enemies of the Internet’, namely the ‘worst violators of freedom of expression on the Net’, are said to be ‘Saudi Arabia, Burma, China, North Korea, Cuba, Egypt, Iran, Uzbekistan, Syria, Tunisia, Turkmenistan, and Vietnam’.

China, whose censorship was put in the spotlight when Google declared on 12 January 2010 that it would stop filtering results on www.google.cn, operates ‘one of the largest and most sophisticated filtering systems in the world’, combining all levels of filtering. The government justifies this ‘Great Firewall of China’ on grounds of the ‘preservation of social order and stability’.

The Chinese system uses national-level filtering, to effectively separate the Chinese Internet from the World Wide Web via a combination of blacklisting and dynamic filtering. It also uses ISP-level filtering, by enacting ‘self-censorship’ laws, based on the ‘voluntary’ Public Pledge on Self-Discipline for the China Internet Industry of 2002. Self-monitoring instructions are also given to employers, which entails server-level filtering (and may lead to job dismissals). Finally, filtering also occurs at the level of the end-user, as was illustrated by the plan to require computer manufacturers to include the ‘Green Dam’ filtering software. While China declared that it would not enforce this plan for consumers and would only require filter installation on computers in public places, some manufacturers did include the software in computers sold to consumers.

---


34 Wang & Hong (n 32), 73.

35 OpenNet Initiative (n 31), 9.


Other States on the ‘Enemies of the Internet’ list are equipped with filtering mechanisms, although they do not necessarily reach the same level of filtering. Cuba, for instance, mimics the Chinese separation of national Intranet and international network. National-level filtering applies to the ‘international Internet’, but this is limited to blacklist filtering\(^38\).

Iran has developed extensive national-level filtering, combining blacklists and dynamic filtering, and ‘authorities claim to have blocked hundreds of thousands of sites’\(^39\), a claim that Saudi authorities have also made\(^40\) regarding their blacklists\(^41\).

These so-called ‘Enemies of the Internet’ are, however, not alone in employing filtering techniques.

Reporters Without Borders start their section on ‘Countries under surveillance’ with Australia\(^42\), which announced in December 2007 that:

‘[ISPs] will be required to provide a “clean feed” of Internet material to schools and households, with the ability for adults to opt-out of the clean feed by notifying the relevant ISP’\(^43\)

The details of the ISP-level blacklist plan, however, have changed since this announcement, and recent information suggests that the government will

‘require all ISPs in Australia to use ISP-level filtering to block overseas hosted Refused Classification (RC) material on the Australian Communications and Media Authority’s (ACMA) R.C Content list.’\(^44\)

Additional filtering would be optional for users wishing to subscribe to other lists of content, but the scheme has already been criticised for being too broad, after a leak of the blacklist to WikiLeaks\(^45\). In July 2010, the Australian government announced that the implementation of the filtering system would be delayed for the purposes of review\(^46\).

---

\(^38\) Reporters sans frontières (n 28), 13.

\(^39\) Ibid., 18.

\(^40\) Ibid., 24.


\(^42\) Reporters sans frontières (n 28), 39.


Several countries in Europe have enacted laws requiring ISP-level filtering, most notably in the context of the fight against child-pornography.

Germany passed in February 2010 a law for the fight against child pornography in communication networks\(^\text{47}\), requiring that ISPs adopt a blacklist maintained by a governmental organisation. There are, however, concerns that the list might be expanded to include other content deemed undesirable\(^\text{48}\), without democratic control of such changes. The state of enforcement of this law is currently uncertain, as the German government announced that it would renounce to filter only days before the President (then Horst Köhler) signed the legislation into law\(^\text{49}\).

The French government followed the German example by proposing the ‘LOPPSI 2’ law, which was adopted by the Assemblée Nationale in February 2010 and which the Sénat is currently examining\(^\text{50}\). While the outcome of the vote to come in the Senate is impossible to determine, there has already been much international criticism of the proposed legislation, considered by some to ‘make even Australia look liberal when it comes to state powers of internet censorship’\(^\text{51}\).

The Irish government was also recently reported to have held discussions on the introduction of Internet filtering\(^\text{52}\).


This appears to be a trend in Europe, as the European Commission submitted in March 2010 a Proposal for a Directive containing the following provisions:

1. Member States shall take the necessary measures to obtain the blocking of access by Internet users in their territory to Internet pages containing or disseminating child pornography. The blocking of access shall be subject to adequate safeguards, in particular to ensure that the blocking is limited to what is necessary, that users are informed of the reason for the blocking and that content providers, as far as possible, are informed of the possibility of challenging it.

2. Without prejudice to the above, Member States shall take the necessary measures to obtain the removal of internet pages containing or disseminating child pornography.

In the United States, ISPs have not been the main recipients of filtering legislation. Instead, there has been a tendency to link funding for schools and libraries to mandatory filtering. The constitutionality of such filtering was recently examined by a Washington State Supreme Court in a case opposing library users to a library:

‘[a] public library has never been required to include all constitutionally protected speech in its collection and has traditionally had the authority, for example, to legitimately decline to include adult-oriented material such as pornography in its collection. This same discretion continues to exist with respect to Internet materials.’

This filtering is a direct consequence of legislation: the Children’s Internet Protection Act (CIPA) contained provisions making ‘federal funding of libraries conditional upon their use of blocking software filters’, and was unsuccessfully challenged before the US Supreme Court. As a result of this legislation and the Supreme Court’s decision, all libraries obtaining ‘e-rate’ funding are required to filter certain kinds of content on the Internet. According to recent reports, 66.9% of urban public libraries applied for e-rate funding, while 17.4% of urban libraries chose not to apply for such funding precisely ‘because of the need to comply with CIPA’s filtering requirements’. As such, one

---


can expect that software filtering at the individual computer level is required by US law in 70% of urban libraries.

As far as public schools are concerned, the percentage of schools filtering Internet content seems to be close to 100%\[^{59}\].

Filters need not, however, always target a large amount of content. There are many instances in which national courts or national governments have ordered the blocking of one particular website.

An illustration was provided by Belgium in April 2009, when the judicial authorities requested ISPs to block access to the controversial ‘Stopkinderporno’ website\[^{60}\], which showed the location of individuals previously convicted of child molestation, because it presented the risk of leading citizens to take justice in their own hands\[^{61}\].

The Yahoo! case is another illustration, requesting that the website owner, another form of ISP, remove access to certain items on its website.

3.2. Voluntary or State-encouraged ISP-level filters

In the United Kingdom, there has been a system of self-imposed blacklist filtering by ISPs since 2004, when BT, one of the UK’s largest ISPs, provided other ISPs with software to implement a blacklist compiled by the Internet Watch Foundation (IWF)\[^{62}\], an ‘independent self-regulatory body’\[^{63}\]. This system, while voluntary in its inception, resembles State-imposed blacklisting, as the government stated in 2007 that all ISPs were required to implement such a system\[^{64}\], although high


adoption rates (98.6% of consumer broadband lines covered\(^65\)) have limited the need for such legislation.

The aim of the list is to fight child pornography, according to the IWF:

‘Every URL on the list depicts indecent images of children, advertisements for or links to such content. [...] As well as making the internet a safer place for everyone, this initiative can help to diminish the re-victimisation of children by restricting opportunities to view their sexual abuse and may disrupt the accessibility and supply of images to those who seek them out.’\(^66\)

As in the German and Australian cases, there are concerns regarding the reach of the list, which was shown to go beyond child pornography and to censor lawful Internet content\(^67\).

Several countries have followed the British example, notably Scandinavian countries, whose blacklists have been leaked to WikiLeaks\(^68\), and the Netherlands\(^69\). Most of these countries, however, chose to entrust a governmental body with the task of maintaining a blacklist, although Ernst Hirsch Ballin, Dutch Minister of Justice, stated after the report by Wouter Stol et al.\(^70\) that the Dutch police would no longer maintain the blacklist, leaving ISPs to do it on their own\(^71\).

Since the Yahoo! case, several ISPs, mostly web hosting providers and operators of websites with user-generated content, have begun to supervise hosted content. Some have created systems for the rapid response to complaints regarding hosted content, while those of a second category, fewer in number, actively monitor their content and employ filtering techniques.


\(^{69}\) Stol \& others (n 14), 251 \& 257.

\(^{70}\) Stol \& others (n 13).

eBay provides an illustration of systems of the first category with its anti-counterfeiting scheme, the ‘Verified Rights Owner (VeRO) Program’\(^{72}\). Under such a scheme, rights owners may notify eBay of alleged infringements\(^{73}\).

This does not, however, amount to filtering, and L’Oréal brought an action against eBay with the aim of forcing eBay to implement dynamic filtering, which would place eBay in the second category. While L’Oréal lost its case in a number of European countries\(^{74}\), the High Court of London\(^{75}\) referred questions to the European Court of Justice (ECJ)\(^{76}\), and the outcome of the case may become relevant to our legal analysis\(^{77}\).

3.3. Company, library, school and home filtering

A number of organisations and persons may decide to use filtering techniques on their own servers or computers, generally a combination of blacklist and dynamic filtering through the use of commercial or open-source filtering software.

Companies, for instance, may view the Internet as both enhancing and reducing productivity, and may therefore seek to limit the risk of employees spending hours on websites unrelated to their work. Although no figures are available to assess the scale of such filtering today, a recent US survey showed that 54% of 1400 companies with 100 or more employees wholly prohibited the use of social networking websites such as Facebook and Twitter\(^{78}\). This suggests that figures of 2001, which showed that 14% of companies in Italy used filters and 7% carried out ‘active system monitoring’\(^{79}\), may no longer reflect the reality of 2010.


\(^{73}\) Ibidem.


\(^{75}\) L’Oréal v eBay [2009] EWHC 1094 (Ch).

\(^{76}\) European Court of Justice (ECJ), Case C-324/09 L’Oréal & others, Reference, 7 November 2009, OJ C 267/40.

\(^{77}\) Infra, 29–30.


Libraries and schools, which we have briefly examined from the point of view of US legislation, may also choose to use filtering software to avoid situations where patrons or students inadvertently stumble across, or see someone else access, offensive or obscene material. In the UK, 80% of 444 schools and colleges reported having filters in place in 2005\textsuperscript{80}, although 45% of these filters were required by the Local Education Authority and 40% of the filters ‘came preloaded on equipment’\textsuperscript{81}.

Individuals may also choose to implement filters. This is generally the case of parents wishing to limit their children’s Internet access. Ofcom, the UK telecommunications regulator, reported in March 2010 that 43% of surveyed parents of children aged 5 to 15 used filtering software\textsuperscript{82}. As awareness of filters is high\textsuperscript{83}, this figure suggests that parents enjoy greater freedom of choice than schools regarding the installation of filters.

\textsuperscript{80} C Barrow & G Heywood-Everett, ‘E-safety: the experience in English educational establishments’ (British Educational Communications and Technology Agency 2006), 41, \url{http://partners.becta.org.uk/page_documents/research/esafety.pdf}, accessed 7 July 2010.

\textsuperscript{81} Ibid., 41–42.

\textsuperscript{82} Ofcom, ‘UK children’s media literacy’ (26 March 2010), \url{http://stakeholders.ofcom.org.uk/market-data-research/media-literacy/medlitpub/medlitpubrss/ukchildrensml/}, accessed 7 July 2010.

\textsuperscript{83} Ofcom reported in 2007 that 82% of parents were aware of the existence of filtering software: Ofcom, ‘Ofcom’s Submission to the Byron Review - Annex 5: The Evidence Base - The Views of Children, Young People and Parents’ (30 November 2007), 49, \url{http://stakeholders.ofcom.org.uk/binaries/research/telecoms-research/annex5.pdf}, accessed 7 July 2010.
II. Legal analysis of filters

Although Internet filtering is practiced worldwide, the use of filters raises many questions with respect to laws and rights. Sacrificing scope for depth, we have limited our legal analysis both in space and in terms of the level of filtering. Reference shall therefore be made only to rules applicable in (parts of) Europe, and our analysis is focussed mainly on national- and ISP-level filtering, although parts of our analysis equally apply to server- and computer-level filtering.

1. Freedom of expression

Fundamental rights and freedoms have played a central role in European legislation since the signing of the European Convention on Human Rights (ECHR) in Rome in 1950, and this was reaffirmed by the EU in its Charter of Fundamental Rights.

While these instruments stem from two different legal systems, namely the Council of Europe and the EU, there is much interaction and mutual influence between these two systems. The ECJ, for instance, held as follows in Ter Voort:

‘As the Court has consistently held […], fundamental rights form an integral part of the general principles of law, the observance of which the Court observes. For that purpose the Court draws inspiration from the constitutional traditions common to the Member States and from the guidelines supplied by international treaties for the protection of human rights on which the Member States have collaborated or of which they are signatories. The [ECHR] has special significance in that respect. It follows that the [EU] cannot accept measures which are incompatible with observance of human rights thus recognized and guaranteed.’

One of the freedoms protected by both instruments is highly relevant to the existence of filters, as it pertains to the creation of and access to information: freedom of expression.

Article 10 ECHR provides as follows:

1. Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. […]

2. The exercise of these freedoms, since it carries with it duties and responsibilities, may be subject to such formalities, conditions, restrictions or penalties as are prescribed by law and are necessary in a democratic

---


85 Since the entry into force of the new EU treaties (TEU and TFEU), all references to the ‘Community’ are to be read as ‘European Union’.

86 ECJ, Case C-219/91 Ter Voort [1992] ECR I-05485 [34].
society, [...] for the prevention of disorder or crime, for the protection of health or morals, for the protection of the reputation or rights of others, [...]’

Article 11 of the Charter of Fundamental Rights, to be read in conjunction with Article 52, provides for similar rules.

The importance of these freedoms in the digital realm was underlined in Directive 2009/140/EC and its amendments to Directive 2002/21/EC, of the EU’s Telecons Package. Moreover, in a Declaration adopted on 28 May 2003 by the Committee of Ministers of the Council of Europe (‘Declaration on freedom of communication on the Internet’)89, European States reiterated their commitment to this fundamental freedom.

The following question arises: does a filter hinder freedom of access to information and freedom of speech?

1.1. Applicability of the right to freedom of expression

As Internet filters block access to (parts of) websites, they constitute an interference with the reception and imparting of information embodied in the relevant websites. The issue of the applicability of Article 10 ECHR to content blocked by a filter, however, is not resolved by the mere classification of filters as ‘interference’: many objections to applicability may arise.

We have observed that both private and public persons or entities resort to filters. One might object that Article 10 ECHR speaks only of interference by a public authority and is thus inapplicable to filtering by private parties. The European Court of Human Rights (hereinafter ECtHR or ‘the Court’) has nevertheless held that Article 10 ECHR has horizontal effect: everyone has the right to freedom of expression without interference by private parties90.

---


90 European Court of Human Rights (ECtHR), Fuentes Boho v Spain (App no 39293/98) ECHR 29 February 2000 [38]: ‘A cet égard, la Cour rappelle que l'article 10 […] peut également s’appliquer lorsque ces relations relèvent du droit privé […] En outre, dans certains cas, l’Etat a l’obligation positive de protéger le droit à la liberté d’expression contre des atteintes provenant même de personnes privées’ (the text of this judgment is available in French only).
The question whether private parties such as ISPs should nonetheless be regarded as ‘public authorities’ may also be approached from the angle of the function fulfilled: in the UK, courts held that the Advertising Standards Authority, a self-regulatory body, was reviewable, as it exercised a function that would otherwise have been exercised by a public authority91. This classification as ‘public authority’ is of great importance regarding the legal basis of lawful interference and regarding transparency requirements92.

Another objection to the applicability of Article 10 ECHR would be that filters generally target content that does not deserve protection by freedom of expression. It is, however, settled case-law of the ECtHR that Article 10 ECHR covers all information, regardless of whether it is private, commercial, public or even offensive93. The Court has nevertheless been ‘less tolerant of restrictions of political expression’ than it has been of interference with other forms of speech94.

A third objection might be that filtering does not interfere with the content, applying only to the means of transmitting and receiving information. In _Autronic_, however, the ECtHR stated that:

‘Article 10 […] applies not only to the content of information but also to the means of transmission or reception since any restriction imposed on the means necessarily interferes with the right to receive and impart information.’95

Internet filters therefore constitute an interference with the transmission and reception of content protected by Article 10 ECHR, whether the content be pedo-pornographic, politically undesirable or simply objectionable from the point of view of the filter operator or creator.

The protection offered by Article 10 ECHR is, however, not absolute.

As explicitly stated in Article 10(2) ECHR, there may be ‘formalities, conditions, restrictions or penalties’ if they are ‘prescribed by law’ and if they are ‘necessary in a democratic society’.

---

92 Infra, 31-33.
94 Tambini & others (n 91), 272.
95 ECtHR, _Autronic AG v Switzerland_ (App no 12726/87) (1990) Series A no 178 [47].
Such limits to the precedence of freedom of expression over the possibility of filtering were enshrined in Principle 3 of the aforementioned Declaration on Freedom of Communication on the Internet:

‘Provided that the safeguards of Article 10, paragraph 2, of the [ECHR] are respected, measures may be taken to enforce the removal of clearly identifiable Internet content or, alternatively, the blockage of access to it, if the competent national authorities have taken a provisional or final decision on its illegality.’

In the light of the fact that such Declarations of the Committee of Ministers are used for the interpretation of the ECHR, one must assess whether, in the event of applicability of Article 10 ECHR, the interference to freedom of expression caused by filtering techniques is a justifiable interference in accordance with Article 10(2) ECHR.

1.2. ‘Prescribed by law’

The first requirement of Article 10(2) ECHR, namely that an interference be ‘prescribed by law’, was examined by the ECtHR in Sunday Times (No. 1):

‘[T]he word “law” in the expression “prescribed by law” covers not only statute but also unwritten law. […] First, the law must be adequately accessible: the citizen must be able to have an indication that is adequate in the circumstances of the legal rules applicable to a given case. Secondly, a norm cannot be regarded as a “law” unless it is formulated with sufficient precision to enable the citizen to regulate his conduct: he must be able - if need be with appropriate advice - to foresee, to a degree that is reasonable in the circumstances, the consequences which a given action may entail.’

With such a definition in mind, are filtering techniques prescribed by law?

(i) Law

We have observed that in Germany and France, legislation making filtering mandatory has already been passed or may soon be enacted.

For other countries, the answer to the question is not so evident, as national assemblies appear not to have enacted legislation pertaining to Internet filtering specifically.

We have seen, however, that ‘law’ must not be understood as being confined to published acts of a legislative body. As per Sunday Times, an interference with freedom of expression shall be held to be

96 DCFI (n 89), Principle 3.

97 ECtHR, Demir and Baykara v Turkey (App no 34503/97) ECHR 12 November 2008 [74].

98 ECtHR, Sunday Times v United Kingdom (No. 1) (App no 6538/74) (1979) Series A no 30 [47]-[49].
‘prescribed by law’ if the law is ‘adequately accessible’ and if there is foreseeability of the legal consequences of actions99. Such an understanding of the concept leaves room for manoeuvre:

‘[T]here is clearly a scope for discussion in many cases as to whether certain aspects of the self- and co-regulatory regime constitute rules that are “prescribed by law”. At one end of a continuum, purely voluntary ethics codes of single companies are clearly not law, but at the other, codes that are encouraged through a legislative framework but administered by an industry association may be considered for these purposes to be law.’100

The requirement of a legislative framework was further exemplified in Barthold, where the ECtHR held that rules adopted by a professional association were law because the professional association enjoyed an ‘independent rule-making power’ by virtue of parliamentary delegation101.

May one therefore consider that Internet filtering in countries such as Denmark and the UK is encouraged through a legislative framework?

In Scandinavia and the Netherlands, ISP-level filtering is not mandatory, but the body in charge of the blacklist is of governmental character. For instance, in the Netherlands, ISPs enter into a bilateral agreement with the police, which maintains a blacklist:

‘The ISP takes the obligation to apply the blacklist without change, and run the software that delivers the factual filtering and the blocking on the basis of the blacklist […] As a consequence the ISP in question is filtering on direct instruction from the police.’102

The role undertaken by the police might lead one to assume that filtering is at the very least permitted by law, based notably on the fact that the fight against child pornography is at the heart of such filtering regimes.

In the EU, however, proponents of the view that the criminalisation of child pornography encourages ISPs to filter Internet content (e.g. Cybercrime Convention of 2001103), have to reckon with the E-Commerce Directive, which provides that ISPs are not liable ‘for the traffic of data they have not initiated or cannot influence with respect to content’104. The criminalisation of child pornography does therefore not suffice as a legal framework in such circumstances.

In their analysis of Dutch filtering practice, Wouter Stol et al. found that the assumption of legality did not necessarily hold true in reality:

99 ECtHR, Sunday Times (n 98) [49].
100 Tambini & others (n 91), 282.
102 Stol & others (n 14), 257.
104 Stol & others (n 14), 256.
We shall examine the issue of liability in further detail infra, 27-30.
enforcement authorities, in particular the police, do not avail over legal powers to filter and to block internet traffic in general, including traffic to and from internet sites that hold child pornographic material.\textsuperscript{105}

In the UK, one may not even begin to assume that the interference is ‘prescribed by law’ as we did in the Dutch case, as the authority in charge of the blacklist is a non-governmental organisation, a quasi-industry body to which neither the UK government nor Parliament have delegated any task regarding Internet filtering\textsuperscript{106}.

We shall nevertheless press on with our analysis, operating under the questionable assumption that all filtering systems are encouraged through a legislative framework, so as to better assess the degree of conformity of such systems with freedom of expression. Should the aforementioned EU Directive on combating child pornography\textsuperscript{107} be adopted, this assumption would likely be verified.

A crucial question, therefore, is whether these blacklists and related rules meet the two requirements of \textit{Sunday Times}, namely accessibility and foreseeability of law.

\textit{(ii) Accessibility}

Accessibility of law depends on the publication of norms:

‘[T]he relevant legal principles which are at the origin of interferences with the enjoyment of human rights must have been in some way published so as to be available to interested parties.’\textsuperscript{108}

For blacklist filtering to meet this requirement, there must be a clear indication of the principles guiding the creation and maintenance of the blacklist. While this can easily be achieved by the filtering authorities, another issue arises, namely that of awareness of the regulatory role of the entity in question. Citizens know that their government and legislative assemblies make laws, and turn to them for access to ‘law’; are they aware, however, that their access to the Internet is filtered and that this is not necessarily done by the State?

Because many citizens will be unaware of these facts, States with national-level filtering have a duty to ensure that citizens have access to the principles underlying the act of filtering, even if filtering is

\textsuperscript{105} Ibidem.


\textsuperscript{107} Supra n 53.

carried out by non-governmental entities. In this respect, these principles need not be enshrined ‘in the very text which authorises the imposition of restrictions’\(^{109}\).

As such, current practices of national-level blacklist filtering, of which citizens are mostly unaware, seem problematic with respect to accessibility of law.

(iii) Foreseeability

The requirement of foreseeability is inherently linked to transparency, as the extent to which citizens are able to foresee the consequences of their actions depends on their capacity to assess the content of a legal norm. It is intrinsically linked to accessibility as well: without access to the general principles of a legal norm, a citizen will be unable to determine which principles should guide his actions.

Generally, however, blacklists are kept confidential, which raises concerns as to the mere possibility of foreseeability: given that the contents of a blacklist are not disclosed, it is difficult for citizens to assess which behaviour to adopt.

Blacklists are nonetheless often compiled following specific guidelines, the ‘general principles’ to which we referred earlier. If the authority maintaining a blacklist adheres strictly to these guidelines, a citizen may, upon learning of the guidelines, foresee how he should act. In such a situation, there would be sufficient legal foreseeability. In practice, however, we have observed that several blacklists go well beyond their professed scope\(^{110}\), which negates foreseeability.

We observe therefore that the conformity of filtering techniques with the ECHR appears to be questionable at best in a number of European States, merely from examining the requirement that the interference with freedom of expression be prescribed by law.

1.3. ‘Necessary in a democratic society’

Any restriction to freedom of expression must not only be ‘prescribed by law’, it must also be ‘necessary in a democratic society’ (the second requirement of Article 10(2) ECHR), a notion defined by the ECtHR in *Sunday Times* as implying that the measure respond to a ‘pressing social need’, be ‘proportionate to the legitimate aim pursued’ and be justified by ‘relevant and sufficient’ grounds\(^{111}\).

Before assessing whether Internet filtering conforms with these requirements, one must assess whether the measure pursues an aim that is legitimate under Article 10(2) ECHR.

---


\(^{110}\) Supra nn 45 & 67.

\(^{111}\) ECtHR, *Sunday Times* (n 98) [62].
Currently, Internet filtering systems in Europe profess having the aim to fight child pornography. The IWF illustrates this, stating that the blocking of websites is ‘one element in a wider effort to combat the making and distribution of images of child sexual abuse via the internet’\textsuperscript{112}.

Such an aim directly echoes the ‘protection of health or morals’ and the ‘protection of the […] rights of others’ explicitly mentioned in Article 10(2) ECHR, and it is difficult to imagine how anyone could contend otherwise, especially in the light of the recent entry into force of the Convention on the Protection of Children against Sexual Exploitation and Sexual Abuse\textsuperscript{113}.

Another aim that the IWF professes to have is to ‘protect people from inadvertent access to potentially criminal images of child sexual abuse’\textsuperscript{114}, echoing the aim of crime prevention. It is worth questioning the value of this aim, merely by pondering whether inadvertent access to such images is possible in a World Wide Web which we access by means of oft updated search engine results\textsuperscript{115}, with the knowledge that search engines remove links to illegal content rapidly\textsuperscript{116}. Casual users cannot be deemed to be ‘ostensibly the real targets of the efforts involved’\textsuperscript{117}, as they have never been the focus of the fight against child pornography. Were it the case, the regulatory legitimacy of the fight against child pornography might come into question\textsuperscript{118}.

\textit{(i) ’Pressing social need’}

The Wingrove case concerned a film, ‘Visions of Ecstasy’, which mingled religious ecstasy and pornography in a manner that was criticised as blasphemous in the UK. While the ECtHR did not examine the question of the ‘pressing social need’, Judge Lohmus, in a dissenting opinion, stated the following:

‘In cases of prior restraint (censorship) there is interference by the authorities with freedom of expression even though the members of the society whose feelings they seek to protect have not called for such interference. The interference is based on the opinion of the authorities that they understand correctly the

---

\textsuperscript{112} IWF (n 66).

\textsuperscript{113} Council of Europe Convention on the Protection of Children against Sexual Exploitation and Sexual Abuse (adopted 27 October 2007, entered into force 1 July 2010), CETS 201.

\textsuperscript{114} IWF (n 66).

\textsuperscript{115} ‘On the web, search engines clearly dominate. Other approaches such as web directories […] only play an underpart.’ D Lewandowski, ‘Search engine user behaviour: How can users be guided to quality content?’ (2008) 28 Information Services \& Use 261, 262.


\textsuperscript{117} Edwards (n 5), 664.

\textsuperscript{118} For a more in-depth analysis of regulatory legitimacy, see infra, 31–33.
feelings they claim to protect. The actual opinion of believers remains unknown. I think that this is why we cannot conclude that the interference corresponded to a “pressing social need”.

It is fair to presume that citizens throughout Europe have negative feelings about child abuse, given the large body of legislation on the fight against child abuse and child pornography and in the light of the social response to paedophile cases, such as the ‘White March’ during the Marc Dutroux case in Belgium. In support of this idea, the ECtHR has evoked the ‘seriousness of child abuse as a social problem’ in a case on the reporting of suspicions of child abuse. These elements seem to suggest that the fight against child abuse is indeed a pressing social need.

(ii) ‘Relevant and sufficient’

The requirement that the reasons for interference be ‘relevant and sufficient’, while oft ignored by the ECtHR in its analysis of national measures, suggests that an authority may not justify its measure by reference to general principles. In one particular case, the Court held that:

‘the domestic authorities have not demonstrated in a “relevant and sufficient” manner why the grounds generally advanced in support of the prohibition of political advertising also served to justify the interference in the particular circumstances of the applicant association’s case.’

Pursuant to such rules, one may not simply state that the interference created by Internet filtering is justified by the fight against child pornography, the protection of health and morals or the protection of the rights of the child. There are, however, specific justifications for the filtering of Internet content. For instance, the IWF states that blocking can ‘disrupt the accessibility and supply of images [of child sexual abuse] to those who seek them out’. It is likely that such reasons will be deemed both relevant and sufficient in the framework of the ECHR.

---

119 ECtHR, Wingrove v United Kingdom (App no 17419/90) ECHR 25 November 1996 (Judge Lohmus).
120 ECtHR, Juppala v Finland (App no 18620/03) ECHR 2 December 2008 [42].
122 ECtHR, VgT Verein gegen Tierfabriken v Switzerland (App no 24699/94) ECHR 28 June 2001 [75]. See also ECtHR, Bykov v Russia (App no 4378/02) ECHR 10 March 2009 [65]-[67].
123 IWF (n 66).
A greater source of concern for filters with regard to the second requirement of Article 10(2) ECHR comes from the test of proportionality of the measure. While the ECtHR never defined the test of proportionality, case-law of the ECJ provides such a definition:

‘The [ECJ] has consistently held that the principle of proportionality is one of the general principles of [EU] law. By virtue of that principle, the lawfulness of the prohibition of an economic activity is subject to the condition that the prohibitory measures are appropriate and necessary in order to achieve the objectives legitimately pursued by the legislation in question; when there is a choice between several appropriate measures recourse must be had to the least onerous, and the disadvantages caused must not be disproportionate to the aims pursued.’

While the latter part of this definition, the ‘least onerous’ requirement, is not explicitly recognised as an aspect of proportionality in ECtHR case-law, the Court has nevertheless resorted to such a criterion in some cases.

The aims pursued by filtering in European countries are generally the fight against child pornography in general and the disruption of access to pedo-pornographic material in particular. To assess proportionality, one must first determine whether filtering is appropriate and necessary in order to achieve such objectives.

a) ‘Appropriate and necessary’

The question of the appropriateness and necessity of the measure relates to the practical relationship between aims and effects of the measure. It is therefore based on the adequacy of this measure:

‘[I]t must be possible to show […] not only that the risk is real (even if uncertain), but also that the response has a rational relationship to that risk’

In this context, effectiveness becomes paramount.

In theory, internet content filtering blocks access to such content for all users to whom the filter applies. When applied to filtering, the test of effectiveness is twofold, encompassing both appropriateness and necessity: how well does filtering technology block targeted content (the

---

124 Supra n 85.

125 ECJ, Case C-331/88 The Queen v Minister of Agriculture, Fisheries and Food and Secretary of State for Health, ex parte Fedesa and others [1990] ECR I-4023 [13].


question of ‘underblocking’, or not blocking all the targeted content), and how well does it avoid blocking content that should remain accessible (the question of ‘overblocking’)?

In the case of blacklists used for national- and ISP-level filtering in Europe, their inherent risk is that of underblocking, as one cannot manually examine all the content on the Internet and as ‘new information appearing on the internet is untouched by the filter’\(^{128}\). Furthermore, if these blacklists are compiled automatically, they will necessarily carry the risk of overblocking as well\(^{129}\). Even without such an automated dimension, prohibited content may disappear from a website after it has been blocked\(^{30}\).

These risks are not limited to national- and ISP-level filtering, nor to blacklists. Much research has been carried out on the effectiveness of filtering software installed at server- and computer-level, and results appear to be controversial. Two important studies were carried out in 2008, for the Australian government\(^{131}\) and the European Commission\(^{132}\) respectively. While the former argues that ‘the selected filter products [exhibit] high degrees of accuracy in identifying and blocking prohibited and potentially prohibited content and low rates of overblocking’\(^{133}\), the latter states the following:

‘While […] the tested filters detect more harmful content, they also increasingly unduly block harmless content. The score of the best performing tool was 2.5 on a scale of 4 in 2006, and stayed at the same level in 2007 and 2008 due to overblocking good content notwithstanding improvements in the detection of bad content.’\(^{134}\)

Filtering systems are therefore potentially neither appropriate (due to underblocking) nor necessary (due to overblocking), as clearly demonstrated by Messrs Kamenev and Clayton\(^{135}\).

\(^{128}\) Stol & others (n 14), 252.

\(^{129}\) Supra, 4.


\(^{133}\) ACMA (n 131), 46.

\(^{134}\) European Commission (n 132), 27.

\(^{135}\) Supra nn 45 & 67.
Moreover, while filtering is ‘code’ and therefore harder to go against (‘code is law’), anti-censorship activists and those who are targeted by filters have, thus far, always succeeded in creating code-based means of circumventing filters.

Therefore, as concluded by Wouter Stol et al., filtering may not be effective in halting access to these images:

> ‘Our findings indicate that filters are not effective against “enthusiasts” who mutually exchange pornographic material. They know how and where to make contact with each other anyway. [...] repressive measures against the commercial spread of child pornography will cause the suppliers to explore other, less risky ways to offer their material’

Mike Galvin, who helped create the IWF system, admitted as much in an interview, stating that he had helped build ‘a system that won’t stop the hardened paedophile’.

Moreover, it seems that paedophile rings regularly use non-website technology to access illicit material, from encrypted and limited-access peer-to-peer networks to instant messaging. Filtering technologies, however, are rooted in the concept of the website. Thus, the techniques used appear to be not only ineffective but also technologically outdated.

A last element to take into consideration regarding the effectiveness of a measure is its degree of legitimacy, as legitimacy and effectiveness are correlated. We have briefly examined the legitimacy of the objectives, but not that of the means, which implies an assessment of the measure’s transparency. Because transparency is not explicitly mentioned in ECtHR or ECJ case-law for proportionality, this shall be examined in more detail later.

b) ‘Least onerous’

As per ECJ case-law, measures adopted must be the ‘least onerous’ when alternatives exist.

---


138 *Stol & others* (n 14), 254.

139 *Mathieson* (n 116).

140 *Edwards* (n 5), 630.

141 See e.g. the comparison of ICANN and VCR regulatory effectiveness: AD Murray, ‘Conceptualising the Post-Regulatory (Cyber–state)’, in R Brownswod & K Yeung (eds), *Regulating Technologies: Legal Futures, Regulatory Frames and Technological Fixes* (Hart Publishing, Oxford 2008), 302-309.

142 *Infra*, 31–33.

143 *Supra* n 125.
Proponents of Internet filtering will be quick to state that there are no alternatives, but this is something that the IWF refutes:

‘We consider removal at source to be the most effective way of combating child sexual abuse images online and other criminal content within our remit which has been almost eradicated from UK networks.’\(^{144}\)

The issue that arises with this alternative to blocking is that many websites containing such images are hosted on servers outside the jurisdictional remit of the country concerned. Yet, even in such cases, there is often judicial cooperation enabling two countries to work together to remove illegal material from the Internet. While some countries are often cited as ‘not taking a tough enough stance on child pornography’ and as not having effective means of legal cooperation, such as Russia and the Ukraine\(^{145}\), Wouter Stol \textit{et al.} found that most of the content blocked by the Dutch blacklist was hosted in the UK and in the United States\(^{146}\), two countries with which legal cooperation was wholly possible.

The effectiveness of this alternative was demonstrated by the Internet community during the weeks leading to the German parliamentary vote on Internet filtering:

‘The working group on censorship demonstrated the alternatives for instance by actually removing over 60 websites containing child pornographic content in 12 hours, simply by emailing the international providers who then removed this content from the net. The sites were identified through the black lists of other countries documented on Wikileaks. This demonstration underlines the protesters main arguments: instead of effectively investing time and efforts to have illegal content removed from the internet, the German government is choosing censorship and blocking – an easy and dangerous way out.’\(^{147}\)

This alternative is potentially more onerous in terms of law enforcement, but far less harmful to freedom of expression (although there is also the risk of website hosts removing content too hastily when notified of its apparent illegality\(^{148}\)). In the framework of fundamental rights and freedoms, it is therefore difficult to justify filtering as being the ‘least onerous’ alternative.

2. Liability

2.1. Exclusion of liability

The fundamental principle underlying the liability provisions of the E-Commerce Directive\(^{149}\) is that an Internet intermediary (an ISP) should not be held liable for the transmission, hosting or caching of

\(^{144}\) IWF (n 66).

\(^{145}\) Stol \textit{et al.} (n 14), 260.

\(^{146}\) \textit{Ibid.}, 258.

\(^{147}\) Beckedahl (n 48).

\(^{148}\) \textit{Infra}, 28-29.

\(^{149}\) E-Commerce Directive (n 18), art 12-14.
information. Without such an exclusion of liability, ISPs would for instance be criminally liable for the transmission of images of child sexual abuse, as this transmission would involve possession (albeit temporary) of these images.

This exclusion of liability is nevertheless qualified by additional requirements, such as for example the fact that the ISP may have no actual knowledge of illegal activity or information or that it may proceed to no modification or selection of the information. These additional requirements, however, are limited in scope and in number: both debates in the European Parliament\(^{150}\) and a report of the Commission\(^{151}\) have shown that one may not add new conditions.

These exclusions are important for every form of ISP, be it the Internet access provider or the web host of Internet content. In the framework of fighting illegal content on the Internet, however, they become paramount for the protection of freedom of speech.

As observed previously\(^{152}\), removal at source is an alternative or a complement to filtering, and it is often perceived as more legitimate and more desirable. Much of the time, removal at source is achieved by contacting web hosts\(^{153}\), making them aware of the illegal nature of the content hosted and asking them to remove it. A mere notification may, however, not be sufficient to qualify as ‘actual knowledge’ of the illegal character of the activity or information\(^ {154}\), as was stated in the Declaration on Freedom of Communication on the Internet:

'It is to be expected that Member States will define in more detail what level of knowledge is required of service providers before they become liable. In this respect, so-called “notice and take down” procedures are very important. Member States should, however, exercise caution imposing liability on service providers for not reacting to such a notice. Questions about whether certain material is illegal are often complicated and best dealt with by the courts. If service providers act too quickly to remove content after a complaint is


\(^{152}\)Supra, 27.

\(^{153}\)Many free tools exist to identify the web host of any website. See e.g. www.whoishostingthis.com or who-hosts.com.

\(^{154}\)If one were to interpret ‘actual knowledge of the illegal activity or information’ as indicating that ‘actual knowledge’ of the likelihood of illegal activity or information is required, rather than ‘actual knowledge’ of the illegal character of one clearly identifiable activity or information, it would render this prohibition devoid of substance, as hosting providers know that there is a likelihood of illegal activity or information appearing on their servers.
received, this might be dangerous from the point of view of freedom of expression and information. Perfectly legitimate content might thus be suppressed out of fear of legal liability.

The risk of removing legitimate content by ‘notice and take down’ was shown to be real by two studies, one in the UK and one in the Netherlands, where researchers posted out-of-copyright material online before contacting web hosts under the guise of a fake organisation and stating that the website violated the organisation’s copyright. Rather than examining whether the content was indeed illegal, most ISPs removed the content immediately.

Therefore, while filtering itself raises issues with regard to freedom of expression, the prospect of liability is also a source of concern.

2.2. Prohibition of any general obligation to monitor

Liability is also important in the context of filtering, as the E-Commerce Directive does not limit itself to the mere exclusion thereof. It further provides as follows in its Article 15(1):

‘Member States shall not impose a general obligation on providers, when providing the services covered by Articles 12, 13 and 14, to monitor the information which they transmit or store, nor a general obligation actively to seek facts or circumstances indicating illegal activity.’

Such a prohibition begs the question whether State-imposed filtering is at all permissible. Indeed, filtering requires the monitoring of information that is transmitted and it involves actively seeking facts indicating illegal activity, namely the access to filtered content (which we presume to be illegal).

This provision was undoubtedly crafted with filtering software in mind, given that ‘monitoring is possible only through technical means in the digital environment’.

If general obligations to monitor are prohibited, specific obligations to terminate or prevent an infringement are nevertheless allowed if they meet three conditions: they must target ‘clearly

---

155 DCFI (n 89), 10.
157 E-Commerce Directive (n 18), art 15.
160 E-Commerce Directive (n 18), artt 12(3), 13(2) & 14(3); DCFI (n 89), Principle 6.
161 DCFI (n 89), Principle 3.
identifiable Internet content’, ‘the competent national authorities [must] have taken a provisional or final decision on its illegality’ and they must respect the safeguards of Article 10(2) ECHR.

The question of the scope of permissible injunctions may be decided by the ECJ in the *L’Oréal* case\(^{162}\).

It follows from the analysis of Article 15 of the E-Commerce Directive that mandatory dynamic filtering is prohibited, as it does not target ‘clearly identifiable’ content. As such, the competent national authorities are not in a position to determine whether the content is illegal or not.

Are blacklists permissible, if they are deemed to be specific injunctions? Automatically generated blacklists are likely to be prohibited, because they create new exclusions dynamically, presumably without the necessary prior decision of competent authorities on the illegal character of each new exclusion.

Blacklists that are crafted by hand, however, do not present this risk. There is nevertheless cause for concern in the way that blacklists are currently in use: few of the bodies compiling the list are ‘competent national authorities’, and the safeguards of Article 10(2) are, as we have seen previously\(^{163}\), not always respected.

### 2.3. Liability for removal of legitimate content

It is worth noting that nothing in legislation or case-law suggests any exclusion of liability of either ISPs or filtering authorities for the removal or blocking of legitimate content. Indeed, the wrongful blocking of a website may constitute a tort or delict in accordance with national laws. While no such case has yet been brought forth in Europe\(^{164}\), it is not unlikely to happen.

---

\(^{162}\) ECJ, Case C-324/09 *L’Oréal & others*, Reference, 7 November 2009, OJ C 267/40.

\(^{163}\) *Supra*, 18–27.

\(^{164}\) In Tunisia, one such case was brought before the courts, but was later dismissed.

See H Noman, ‘Tunisian journalist sues government agency for blocking Facebook, claims damage for the use of 404 error message instead of 403’, on *OpenNet Initiative* (12 September 2008), [http://opennet.net/node/950](http://opennet.net/node/950), accessed 28 July 2010;

3. Privacy and data processing

A third set of rules by which to assess filtering is that of data processing, in particular the provisions of the Data Protection Directive.\(^{165}\)

In deploying filters, instead of merely transmitting data between two points (e.g. a user and a website), ISPs are actively seeking to identify the content of telecommunications in order to allow or block the further transmission of data. As the European Data Protection Supervisor (EDPS), Peter Hustinx, stated in June 2008,

\[\begin{align*}
\text{\'The issue raises the question of the intervention of a commercial actor, offering a specific (telecommunication) service, in a sphere where it is in principle not supposed to intervene, that is, the control of the content of the telecommunications. The EDPS recalls that such control should in principle not be done by service providers, and certainly not in a systematic way. When it is necessary in specific circumstances, it should in principle be the task of law enforcement authorities.\text{\'}}^{166} \\
\end{align*}\]

There are therefore concerns as to the legality of this processing of data by ISPs, notably with regard to the fundamental right of privacy (Article 8 ECHR), which in turn requires that the measures be made ‘in accordance with the law and [be] necessary in a democratic society’. As we saw in our analysis of the conformity of filtering with freedom of expression,\(^{168}\) such requirements are not easily met by filtering systems.

4. Legitimacy in governance: transparency

For history to judge in a positive light any form of regulatory intervention, for regulation to be viewed as ‘right’, regulators need to demonstrate several key elements:

\[\begin{align*}
\text{\'they must show that their regulatory interventions are backed by legitimate regulatory purposes, that the regulatory means employed are legitimate, and that the interventions are actually effective.'}^{169} \\
\end{align*}\]

In examining the compatibility of filtering with freedom of expression, we have already touched upon ‘the “legitimation” of regulatory purposes and practices\(^{170}\) and the effectiveness of these practices.
Legitimacy of regulation, however, goes beyond the mere examination of the purported purposes: the real objectives may differ from the purported aims. Legitimacy necessarily entails the assessment of transparency:

‘Good (active) citizenship and political freedom imply an engagement between regulators and regulatees as to the purposes of the regulation. Unless the regulators’ purposes are transparent, there can be no meaningful debate about the acceptability of the measures taken.’

Is transparency ‘a general principle of EU law or, indeed, itself a fundamental right’? Although the question remains unanswered, ‘transparency has been endorsed (very clearly) as being a desirable and necessary objective in a democratic society’. We shall therefore talk of the ‘principle of transparency’, with the caveat that the ECJ may rule that it is or is not so in the coming months.

While self-regulation by the private sector is not subject to the principle of transparency to the same degree as regulation by a public authority, transparency is nevertheless desirable for any form of regulation, notably in the framework of EU-wide regulation:

‘The Commission will ensure that any use of […] self-regulation is always consistent with [EU] law and that it meets the criteria of transparency […] and representativeness of the parties involved.’

When we briefly examined whether filtering techniques presented enough transparency to generate foreseeability of the legal norm, we observed that blacklists often lack transparency, as they are kept confidential, and although there are generally guidelines for the compilation thereof, blacklists go beyond their professed scope.

The assessment of transparency of filters, however, does not end there.


\[172\] Ibid. [68].


\[175\] Supra n 85.


\[177\] Supra, 21.

When a user visits a website that is on the blacklist, how is he or she notified of the blocking? When the IWF blocked access to Wikipedia, users were presented with a 404 HTTP error\textsuperscript{179}, which signifies ‘Page not found’, rather than a 403 HTTP error, which means ‘Access forbidden’\textsuperscript{180}. While the 404 error may be used if the server does not wish ‘to make public why the request has not been fulfilled’\textsuperscript{181}, indicating ‘Page not found’ denotes a severe lack of transparency, as the user will believe that the problem lies with the web server rather than with the filtering system.

Moreover, when a web page is blocked, is any website operator notified, be it the website owner or the website host thereof? The IWF does not notify any entity or person of the addition of the web page, leaving it to the ISP implementing the blacklist to notify the website owner, which does not necessarily happen\textsuperscript{182}.

While one might be justified in advocating such techniques for illegal content, one must bear in mind that some legitimate content is also blocked. The comparison between blocking without justification and press censorship without justification is easily made.

\textsuperscript{179} Edwards (n 5), 655–656.


\textsuperscript{181} Ibid., section 10.4.4.

\textsuperscript{182} Edwards (n 5), 656.
III. Filters mindful of the law

Although regulatory connection, the review of regulation as technology progresses, is desirable, technological changes must also be mindful of the law. We observed in part II that filters and legislation still have enormous gaps to bridge: despite an increase in worldwide use, notably in Europe, filters are seldom respectful of freedom of expression, of provisions on liability of intermediaries and of privacy.

In this third and final part of our work, we shall attempt to reconcile filtering with legal norms and principles, by proposing a filtering system (hereinafter ‘Plan B filters’) that would meet the legal objections arising from our analysis.

1. Freedom of expression

We have seen that the compatibility of filters with freedom of expression is problematic, because of requirements that the filter be prescribed by law (filters are not always required by a specific law), foreseeable (guidelines are not always followed), appropriate and necessary (overblocking occurs), and the least onerous means (removal at source is less onerous).

1.1. ‘Prescribed by law’

To fulfil the first criterion of Article 10(2) ECHR, regardless of who implements the filter, it is important for the legislator to enact law stating that Internet content may be filtered according to a public list of criteria, with a delegation of authority, regarding the list of criteria and/or regarding the filtering system itself, to a reviewable body. The authority in charge of the filtering system or criteria would thus be a public authority, and it would enjoy greater legitimacy in the eyes of the public.

By enacting a clear list of criteria, a State would offer Internet users and website owners the possibility not only to foresee the consequences of the law, but also to hold filter operators accountable for any deviation from these criteria. Greater transparency would therefore benefit everyone.

1.2. ‘Necessary in a democratic society’

Effectiveness is the main barrier to compliance of filters with the second criterion of Article 10(2) ECHR. As overblocking was observed, there is a need to improve the judging of the illegal character
of Internet content and to implement measures for periodical checks of the illegality of blocked content.

To help remedy these flaws, we shall venture to suggest a solution that would frighten most cyber-paternalists183: involve the Internet user. By displaying a 403 ‘Access forbidden’ page with an embedded, easy to use ‘Filter removal request’ form, Internet users would gain the possibility to show that there is overblocking. Alexander, from our introduction, could for instance write that last week he read an interesting article about weedkillers on the blocked website, and that it had nothing to do with the promotion of illegal drugs. If the filtering authority is under the legal obligation to return a reasoned opinion, Plan B filters will not be deemed disproportionate, despite the existence of an alternative (removal at source).

Moreover, law enforcement agencies would welcome the data obtained about a user who submitted the form with false allegations regarding the legality of some websites.

2. Liability

We observed that, to be compatible with Article 15 of the E-Commerce Directive, filters must ‘target “clearly identifiable Internet content”, “the competent national authorities [must] have taken a provisional or final decision on its illegality” and they must respect the safeguards of Article 10(2) ECHR.’184

By ensuring that filtering is limited to manually compiled blacklists, authorities would guarantee that Plan B filters comply with these provisions, subject to the condition that they be deemed to be specific injunctions and not a general obligation to monitor.

As removal at source is desirable, Plan B filters would likely be accompanied by such measures. However, modifications to the ISP liability regime would be required to limit the risk of removal of legitimate content, for instance clarifying the notion of ‘actual knowledge’ with respect to illegality and requiring sufficient procedural guarantees for the respect of rights of the information contributor or website owner.

3. Privacy and data processing

Concerns about the legality of the processing of data by ISPs would need to be addressed by provisions expressly giving ISPs the authority to control the content of telecommunications in a limited and highly regulated manner. With strict requirements regarding the processing of such data,

---

183 Ibid., 625.
184 Supra, 29–30.
objections to Plan B filters would be limited, notably because these filters are designed with the respect of fundamental rights in mind.

4. Legitimacy and concluding remarks

Filters currently in use in Europe lack legitimacy as a result of their opacity. Plan B filters, however, have transparency at their core, and are thus more likely to be viewed as having legitimacy.

Although such filters are more mindful of the law, one might argue that their visibility and transparency reduces their effectiveness, because it becomes obvious that one must circumvent filters to access information, and means of circumvention are readily available. Moreover, the content provider is immediately aware that the content is filtered, and is likely to migrate the data to a new website or server.

These objections, however, negate the idea of having filters in place, as they tend to show that filters in their current form have no effect on the availability of unlawful content on the Internet and are thus devoid of legitimacy.

This leads us to the choice of the term ‘Plan B’. A filtered Internet may be a necessary evil that is more respectful of the rights of others than an unfiltered Internet. Until the former is sufficiently adequate and lawful, however, the latter should be our ‘Plan A’.

From our analysis, we therefore reach the following conclusion: until legislators and industry work together, as proposed for example in this section, to create an acceptable technical and legal mix, one that is mindful of fundamental rights and freedoms and that takes into account other important legal instruments, an unfiltered Internet appears more legitimate and less questionable than a filtered Internet, although the latter understandably responds to a pressing social need.

**Epilogue**

*Alexander looks at the browser window. The same old error has appeared, but he doesn’t trust it. He has seen too many of them to believe that all those websites are having problems. He has done some reading, and feels empowered by it. He doesn’t like to see his freedom limited arbitrarily. Some say that the Internet routes around censorship. Alexander feels like doing the same. Maybe he even has the law on his side.*
1. Articles & studies

American Library Association, ‘Libraries Connect Communities: Public Library Funding & Technology Access Study 2009-2010’ (2010),

N Anderson, ‘Move over, Australia: France taking ‘Net censorship lead’, on Ars Technica (17 February 2010),

Australian Communications and Media Authority, ‘Closed Environment Testing of ISP-Level Internet Content Filters’ (June 2008),

Australian Government – Department of Broadband, Communications and the Digital Economy, ‘Internet Service Provider (ISP) filtering’ (10 March 2010),

C Barrow & G Heywood-Everett, ‘E-safety: the experience in English educational establishments’ (British Educational Communications and Technology Agency 2006),

M Beckedahl, ‘The Dawning of Internet Censorship in Germany’, on netzpolitik.org (16 June 2009),


JC Bertot & others, ‘Public Libraries and the Internet 2009: Study Results and Findings’ (2009),


Bibliography

C Callanan & others, ‘Internet Blocking: Balancing Cybercrime Responses in Democratic Societies’ (October 2009),

A Campbell, Hansard HC vol 497 col 1546W (21 October 2009),


‘The IWF Blocking List, Recent UK Experiences’ (Dublin, 30 June 2009),


Crown Prosecution Service & Association of Chief Police Officers, Memorandum of Understanding concerning Section 46 Sexual Offences Act 2003,

D Deckmyn, ‘Primeur: Belgie blokkeert website’, De Standaard (22 April 2009),

D Drummond, ‘A new approach to China’, on The Official Google Blog (12 January 2010),
‘A new approach to China: an update’, on The Official Google Blog (22 March 2010),


M Ermert, ‘Germany Builds Infrastructure To Block The Internet’, on Intellectual Property Watch (19 June 2009),
European Commission, ‘Safer Internet plus Programme - 2008 Benchmark of tools to filter potentially harmful Internet content’ (November 2008),

R Fielding & others, ‘RFC 2616: Hypertext Transfer Protocol -- HTTP/1.1’ (June 1999), section 10.4,

O Fletcher, ‘China will not enforce Green Dam porn filter plan’, on MIS-Asia (13 August 2009),


S Gaudin, ‘Business use of Twitter, Facebook exploding’, on Computerworld (9 November 2009),


SB Gharbia, ‘Tunisia: Blogger's Home Raided, Laptop and CDs Robbed’, on Global Voices Advocacy (11 April 2009),


B Haselton, ‘Report on Accuracy Rate of FortiGuard Filter’ (2007),

Human Rights In China, ‘Respect the Consumer's Freedom of Choice: Computers Will Not Be Forced To Have “Green Dam” Installed’ (August 2009),

Human Rights Watch, “‘Race to the Bottom” - Corporate Complicity in Chinese Internet Censorship’ (August 2006) 18 Human Rights Watch 8(C).

J Insley, ‘L’Oréal loses British court battle with eBay’, The Guardian (22 May 2009),

Internet Services Unit - King Abdul Aziz City for Science and Technology, ‘Introduction to Content Filtering’ (2006),

Internet Watch Foundation, ‘About the Internet Watch Foundation (IWF)’,
‘IWF Facilitation of the Blocking Initiative’,

ISPA Belgium, ‘Press Release – Reaction on the Blocking of Stopkinderporno Website’ (22 April 2010),


M Kamenev, ‘First, China. Next: the Great Firewall of...Australia?’, Time (Sydney, 16 June 2010),


L Lessig, Code: version 2.0 (Basic Books, New York 2006),


K Lillington, ‘Putting up barriers to a free and open internet’, The Irish Times (16 April 2010),


SA Mathieson, ‘Back door to the black list’, The Guardian (26 May 2005),


AD Murray, ‘Conceptualising the Post-Regulatory (Cyber-state)’, in R Brownsword & K Yeung (eds), Regulating Technologies: Legal Futures, Regulatory Frames and Technological Fixes (Hart Publishing, Oxford 2008).


‘797 domains on Finnish Internet censorship list, including censorship critic, 2008’ (20 March 2009), <http://www.wikileaks.org/wiki/797_domains_on_Finnish_Internet_censorship_list,_including_censorship_critic,_2008>, accessed 3 July 2010;

‘Norwegian secret internet censorship blacklist, 3518 domains, 18 Mar 2009’,
JC York, ‘Germany Passes Legislation to Block Child Pornography’, on OpenNet Initiative (22 June 2009),


2. Cases

2.1. Supra-/International courts

European Court of Justice (ECJ), Case C-331/88 The Queen v Minister of Agriculture, Fisheries and Food and Secretary of State for Health, ex parte Fedesa and others [1990] ECR I-4023;
Case C-219/91 Ter Voort [1992] ECR I-05485;
Case C-112/00 Schmidberger [2003] ECR I-05659;
Joined Cases C-92/09 & C-93/09 Völker und Markus Schecke GbR & Hartmut Eifert, Reference, 6 March 2009, OJ C 129/04;
Case C-324/09, L’Oréal & others, Reference, 7 November 2009, OJ C 267/40.

European Court of Human Rights (ECtHR), Handyside v United Kingdom (App no 5493/72) Series A no 24;
Sunday Times v United Kingdom (No. 1) (App no 6538/74) (1979) Series A no 30;
Silver and others v United Kingdom (App no 5947/72) (1983) Series A no 61;
Barthold v Germany (App no 8734/79) (1985) Series A no 90;
Groppera Radio AG & others v Switzerland (App no 10890/84) (1990) Series A no 173;
Autronic AG v Switzerland (App no 12726/87) (1990) Series A no 178;
Casado Coca v Spain (App no 15450/89) (1994) Series A no 285-A;
Wingrove v United Kingdom (App no 17419/90) ECHR 25 November 1996;
De Haes and Gijsele v Belgium (App no 19983/92) ECHR 24 February 1997;
Fuentes Bobo v Spain (App no 39293/98) ECHR 29 February 2000;
VgT Verein gegen Tierfabriken v Switzerland (App no 24699/94) ECHR 28 June 2001;
Demir & Baykara v Turkey (App no 34503/97) ECHR 12 November 2008;
2.2. National courts

**France**


**United Kingdom**

*L’Oréal v eBay* [2009] EWHC 1094 (Ch)

**United States of America**

*Bradburn v N. Cent. Reg’l Library Dist*, Supreme Court of the State of Washington (WA 2010)(No. 82200-0),


*US v ALA*, 539 US 194, No. 02-361 (2003),


3. Legislation, treaties & declarations

3.1. Council of Europe


3.2. European Union


3.3. National legislators

France

Germany