

# Cut Out By The Middle Man

## The Free Speech Implications Of Social Network Blocking and Banning In The EU

by **Patrick Leerssen**, post-graduate student of Information Law at the University of Amsterdam, research assistant at the Institute for Information Law (IvIR)

**Abstract:** This article examines social network users' legal defences against content removal under the EU and ECHR frameworks, and their implications for the effective exercise of free speech online. A review of the Terms of Use and content moderation policies of two major social network services, Facebook and Twitter, shows that end users are unlikely to have a contractual defence against content removal. Under the EU and ECHR frameworks, they may demand the observance of free speech principles in state-issued blocking orders and their implementation by intermediaries, but cannot invoke this 'fair balance' test against the voluntary removal decisions by the social network service. Drawing on prac-

tical examples, this article explores the threat to free speech created by this lack of accountability: Firstly, a shift from legislative regulation and formal injunctions to public-private collaborations allows state authorities to influence these ostensibly voluntary policies, thereby circumventing constitutional safeguards. Secondly, even absent state interference, the commercial incentives of social media cannot be guaranteed to coincide with democratic ideals. In light of the blurring of public and private functions in the regulation of social media expression, this article calls for the increased accountability of the social media services towards end users regarding the observance of free speech principles.

Keywords: Social Media; Banning; Private Censorship; Removal Orders

© 2015 Patrick Leerssen

Everybody may disseminate this article by electronic means and make it available for download under the terms and conditions of the Digital Peer Publishing Licence (DPPL). A copy of the license text may be obtained at <http://nbn-resolving.de/urn:nbn:de:0009-dppl-v3-en8>.

Recommended citation: Patrik Leerssen, Cut Out By The Middle Man: The Free Speech Implications Of Social Network Blocking and Banning In The EU, 6 (2015) JIPITEC 99 para 1.

### A. Introduction<sup>1</sup>

1 Social media have taken up a central role in public discourse, and are often hailed as a boon to free speech. Social network services (SNS) such as Facebook and Twitter facilitate civic participation in numerous ways. Firstly, they can act as soapboxes for the 'average citizen' to voice his or her opinion, leading to high-profile expressions of political sentiment such as with the #jesuischarlie and #illridewithyou hashtags. Secondly, SNSs act as gateways for accessing external links and resources, with the average news website relying on Facebook and/or Twitter for over 25% of its traffic.<sup>2</sup>

Thirdly, they have also played an important role in the organisation of major 'real world' political manifestations such as the Arab Spring protests and the Occupy movement.<sup>3</sup> In comparison to the linear dissemination models of 'mass media' such as radio, press and television, SNSs have been praised for creating a more diverse and accessible public debate.<sup>4</sup> And yet, these networked systems also lead to a (re-)centralisation of power around a new set of privileged actors: the social network service providers themselves.

2 While we tend to view social media platforms as neutral carriers of information, their operators possess the technical means to remove information

and suspend accounts. As such, they are uniquely positioned to delimit the topics and set the tone of public debate. Increasingly, they have shown themselves prepared to apply these techniques in order to moderate their users and block undesirable information.<sup>5</sup> SNSs may take on this editorial role out of their own commercial interest, or as a matter of compliance with (perceived) legal duties or government orders. In both cases, this may lead them to stifle potentially legitimate forms of expression. This raises the question whether end users can legally contest SNS removal decisions, and protect themselves from such interference. To what extent are social network services required to observe free speech principles under the EU legal framework when removing end-user content from their services? Does this level of protection guarantee the effective exercise of the right to freedom of expression in practice?

- 3 This article will start by examining the Terms of Use and content moderation policies of two major SNS, Facebook and Twitter, in order to illustrate their handling of user-generated content and to examine whether end users can rely on contractual grounds to contest content removal. This will be followed by a review of European Convention of Human Rights (ECHR) case law regarding positive obligations to protect free speech and their application to intermediary content removal. Subsequently, it will review the EU's legal framework, focusing on its e-Commerce regime and the Court of Justice of the European Union's Charter-based case law. Finding that neither framework is likely to provide a defence against *voluntary* removal decisions by the SNS (as opposed to injunction-based measures), this article will explore the potential for abuse of this competence. Firstly, it will detail how EU governments have attempted to influence SNS content policies in the context of anti-terrorism efforts, allowing for the indirect exercise of state power and a 'privatisation' of censorship. Secondly, it will be argued that, even absent state interference, the commercial incentives of SNSs and their responsibilities towards end users and third parties do not guarantee the observance of free speech principles. In light of the blurring of public and private functions in the regulation of expression via social networks, this article calls for the increased accountability of the SNS towards end users as a means to protect online speech.
- 4 Depending on one's definition, the term 'social network service' can apply to a broad range of online services, from dating websites such as eHarmony to videohosting platforms such as Youtube. This article will focus on Twitter and Facebook as illustrations of this broader category, due to their unique popularity and global reach. Twitter currently serves over 250 million users and Facebook over 1 billion.<sup>6</sup> These numbers are rivalled only by the Chinese 'Weibo' and the Russian 'VKontakte', which, relatively

speaking, do not reach a major audience outside their country of origin.<sup>7</sup> Furthermore, Facebook and Twitter are not dedicated to one particular format or topic and often include highly political forms of discussion (in comparison to, say, eHarmony's online dating community or LinkedIn's professional networking model). Therefore, as gatekeepers to online political discourse, these two websites are especially deserving of scrutiny regarding the level of free speech protection provided to their users.<sup>8</sup>

## B. SNS Terms of Use

- 5 Specific rules on content removal can be found in social network Terms of Use (ToU), which govern the contractual relationship between end users and the service provider. Before delving into the general constraints imposed by fundamental rights and other public law sources, it is therefore worth examining the level of protection that has resulted from these private agreements. Contractual assurances can set conditions for content removal, and also contribute to its foreseeability by informing end users of their rights and responsibilities relating to their content. This section will examine the Facebook and Twitter ToU, assessing their level of protection against the removal of user-generated content ('blocking' or 'removing') and termination or suspension of service ('banning').
- 6 Article 5 of the Facebook ToU, titled 'Protecting Other People's Rights', starts with the following paragraphs:
  - *You will not post content or take any action on Facebook that infringes or violates someone else's rights or otherwise violates the law.*
  - *We can remove any content or information you post on Facebook if we believe that it violates this Statement or our policies.*<sup>9</sup>
- 7 Removal is thus permitted for those content categories prohibited by Facebook's policy, and for any content that infringes individual rights or violates the law. All forms of illegal content, such as criminal hate speech, child pornography or copyright infringement, are removable, but this also goes for breaches of Facebook's terms, which outline a large number of additional prohibitions. These include bullying, harassment, intimidation, and nudity, as well as the use of Facebook 'to do anything unlawful, misleading, malicious or discriminatory'.<sup>10</sup> Grounds for removal thus reach far further than the requirements of the law, and include some rather vague terms. What constitutes a misleading or malicious post is highly subjective, and could vindicate the removal of a broad range of content. To make matters worse, the Article 5 removal competence also covers cases where

Facebook *believes* that a violation of their statement has occurred. Strictly speaking, this would seem to relieve them of the burden of proving an actual infringement. Any slightly contentious content could easily be filed under the above prohibitions, and thus be susceptible to deletion.

- 8 The ‘banning’ of Facebook users is governed by article 14, ‘termination’, which states the following: ‘If you violate the letter or spirit of this Statement, or otherwise create risk or possible legal exposure for us, we can stop providing all or part of Facebook to you.’<sup>11</sup> Uploading any prohibited content as described above is thus grounds for account termination. Furthermore, end-user conduct is further subject to a broad range of prohibitions, from egregious behaviour such as the operation of pyramid schemes, to relatively innocuous acts such as the sharing of one’s password, or making multiple accounts. Facebook also prohibits users from doing ‘anything else which might jeopardise the security of your account’. Moreover, users must use their real name and ‘keep contact information up to date’. Breaching any of these rules can result in account termination. Given the broad range of ToU prohibitions, as well as the prohibition of violations *in spirit*, users are faced with a far-reaching possibility of intervention. In short, while Facebook’s right to remove content and ban users is conditional upon some form of illegality or contractual infringement, the contractual prohibitions are defined so broadly as to provide a large degree of interpretive discretion for the intermediary, and almost no protection or foreseeability for end users.
- 9 While Facebook requires at least some illegality or an infringement of their terms to permit account termination and content removal, Twitter’s powers of intervention are completely unconditional. Article 8 of the Twitter ToU reads: ‘(...) We reserve the right *at all times* (but will not have an obligation) to remove or refuse to distribute *any* Content on the Services, to suspend or terminate users, and to reclaim usernames without liability to you.’<sup>12</sup> Reference is made to ‘the Twitter Rules’, which outline Twitter’s policy as to prohibited content.<sup>13</sup> In comparison to Facebook’s content rules, they seem more protective of the user: besides rules on security and commercial communications, all they explicitly prohibit is impersonation, violent threats, and ‘unlawful use’. Nevertheless, these rules are merely policy guidelines within a contractual framework which permits any and all removals and bans. While contracts determine the intermediary’s legal right to removal and tend to be rather broad or even unconditional, supplementary documents such as the *Twitter Rules* and Facebook’s *Community Guidelines* outline how the intermediary intends to exercise their removal powers in practice. They serve to inform users as to the limits of acceptable conduct and content in a more detailed and understandable

manner, but do not affect the intermediary’s legal position. While these statements of policy might contain balanced and reasonable principles, they do not amount to a contractual guarantee included in the platform’s terms of service. Even if SNSs could generally be expected to adhere to their self-imposed rules and guidelines as a matter of policy, thereby granting users a *de facto* enjoyment of their right to impart and access information, their Terms of Use reveal a refusal to guarantee such rights *de jure*.

- 10 End users’ acceptance of social network Terms of Use should not be misconstrued as to mean that these consumers are familiar with, or condone the degree of free speech protection they afford. As a recent study has pointed out, there are many factors which challenge the traditional view that competition between services and consumer choice can determine the appropriate level of protection.<sup>14</sup> First of all, research has shown that a majority of the online public neglects to even read (a significant portion of) online Terms of Service.<sup>15</sup> Furthermore, the minority that do take this effort, may lack the legal expertise and resources to properly assess the content of the ToU provisions.<sup>16</sup> As a result of this informational asymmetry, SNSs are not forced to compete fully with one another as to their relative level of free speech protection.<sup>17</sup> Furthermore, even informed consumers may have little choice due to the strong network effects exhibited by SNSs; switching to a more protective service might be unfeasible if all one’s friends and connections stay with the incumbent.<sup>18</sup> Moreover, SNSs require time and effort from users to establish their profiles and networks, which discourages them from switching between platforms – a phenomenon described by Chiu as the ‘stickiness’ of social media.<sup>19</sup> In light of these considerations, unwarranted or unexpected content blocking should not be seen as a possibility which end users knowingly and willingly subject themselves to as a result of free and informed choice between various online service providers.<sup>20</sup> End users may not be fully aware of removal conditions, or may simply lack suitable alternative options.

## C. State interference and private censorship under the European Convention on Human Rights

- 11 The European Convention on Human Rights’ right to free speech, laid down in article 10, is solely enforceable against Council of Europe Member States. As an international treaty, it does not have a direct effect on the legal relationships between private parties (‘horizontal effect’).<sup>21</sup> Rather, it creates State duties to refrain from interfering with fundamental rights (negative obligations), and duties to undertake specific actions in order to safeguard the effective enjoyment of these

rights (positive obligations).<sup>22</sup> In the context of SNSs, blocking injunctions issued by a court or administrative body might therefore be contested as a form of state interference with the right to free speech. Such a ‘state interference’ does not occur when SNSs remove content voluntarily. However, it will be argued that such private interferences can potentially trigger the State’s positive obligations to protect the free speech of end users. This section will examine the ECHR’s case law on freedom of speech, paying particular attention to the State’s positive obligations, in order to assess its effect on SNS content moderation.

- 12 Article 10 ECHR protects the freedom of expression as follows. Paragraph 1 guarantees a right of freedom of expression to everyone, while paragraph 2 allows for limitations of this right based on a limitative list of grounds such as national security, public security, the prevention of crime and the protection of the rights of others. However, these restrictions must be ‘prescribed by law’ and be ‘necessary in a democratic society’.<sup>23</sup> The former requirement speaks to the rule of law principle and demands that the exercise of state power is accessible and foreseeable.<sup>24</sup> The latter can be said to comprise a more general proportionality test, which weighs the interest in upholding freedom of expression versus other competing interests – the interference must ‘answer to a pressing social need’ and apply ‘necessary and sufficient’ means to that end.<sup>25</sup> In assessing these criteria, the Court seeks to strike a ‘fair balance’ between the conflicting rights and interests involved. In *Soering v UK*, the Court stated that ‘inherent in the whole of the Convention is a search for a fair balance between the demands of the general interest of the community and the requirements of the protection of the individual’s fundamental rights’.<sup>26</sup>
- 13 The importance of the internet as a forum for public debate was recognised in *Yildirim v Turkey*. Subsequently, in *Animal Rights Defenders* the Court recognised that social media provide a platform for the exercise of free speech. Consequently, a blocking injunction compelling an SNS to remove content can be considered a form of state interference with the freedom of expression which would have to fulfil the requirements of Article 10(2). Affected social network users can thus contest orders which fail to adequately observe the right to freedom of expression as developed in the ECHR’s case law. In addition, such injunctions could also be considered an interference with the free speech of SNSs.<sup>27</sup>
- 14 When considering the necessity of these measures in a democratic society and the ‘fair balance’ between the interests involved, a number of factors must be taken into account, such as the nature of the speech affected; the public interest which the injunction serves; and the measure’s proportionality.<sup>28</sup> Concerning the nature of speech involved, a high level of protection is consistently awarded to political speech and ‘contributions to the public debate’.<sup>29</sup> This has led Jacob Rowbottom to conclude that much of social media speech enjoys a lower level of protection under the ECHR, as it commonly concerns more casual and amateur expression.<sup>30</sup> This may be true for a majority of content, but it also follows that injunctions affecting the social media activities of politicians, journalists and activists should be treated with greater scrutiny. The Court has also held that artistic expression is protected under the Convention.<sup>31</sup> Furthermore, while the right to free speech covers statements that ‘offend, shock or disturb’<sup>32</sup>, States have greater discretion in dealing with ‘gratuitously offensive’<sup>33</sup> speech acts such as holocaust denial.
- 15 State action can also fall foul of the ‘fair balance’ due to a lack of proportionality. This requirement could prohibit overly broad injunctions which go further than necessary to pursue their aim.<sup>35</sup> In the context of SNSs, the interests of copyright enforcement might require the removal of infringing images, but need not call for the suspension of accounts involved, or deletion of comments and responses associated with said images. Injunctions categorically prohibiting certain terms or files, or placing a duty of care such as a monitoring or filtering obligation on the intermediary, could also raise questions of proportionality (although this issue has been treated in greater detail under EU law, to be discussed in section E).
- 16 The above shows that the ECHR places limits on state orders compelling content removal by SNS companies. While the protection of end users against such interference is by no means absolute, the Convention can be relied upon to prohibit the most extreme and disproportionate interferences by public authorities. But what of *voluntary* content removal decisions, where the state is not directly involved? As stated, private actors are not directly bound by the Convention. Instead, the ECHR can also create positive State obligations, by which the State can be required to take action in order to ensure the effective exercise of Convention rights by its citizens.<sup>36</sup> Arbitrary or unwarranted removal by SNSs could therefore trigger such an obligation. After all, the Court has repeatedly stated that “[t]he Convention is intended to guarantee not rights that are theoretical or illusory but rights that are practical and effective.”<sup>37</sup> The remainder of this section will focus on positive obligations in the context of Article 10 ECHR, in order to determine whether they might provide protection against voluntary forms of content removal.
- 17 In ECHR case law, positive obligations to protect free speech have been identified under a broad range of circumstances. For example, in *Özgür Gündem* and *Dink*, the Turkish state was found to have breached

its positive obligation to protect journalists against harassment and assault by their fellow citizens (in the case of *Dink* even leading to the victim's death).<sup>38</sup> The positive dimension of the right to free speech would have required the Turkish government to actively investigate credible threats, and provided for the physical security of targets.<sup>39</sup> In *Fuentes Bobo*, the dismissal of a journalist for remarks made on television about his employer was also found to trigger positive obligations due to the chilling effects of such measures.<sup>40</sup> Accordingly, the Court found a positive obligation to apply labour law in such a way as to prevent its abuse for the limitation of free speech.<sup>41</sup>

- 18 By analogy, if a threat to free speech is identified in the removal of SNS content, a positive obligation for the state to protect end users' freedom of expression might entail the prevention of unfair content removal by the intermediary. Prohibiting some forms of content removal, would in effect require the uninterrupted provision of the social network's hosting service to the end user involved. In contrast to the cases described above, it does not concern the prevention of repercussions or retaliations following speech acts, but rather the facilitation or enabling of expression. In other words, it concerns the positive freedom *to* express rather than the negative freedom *from* interference with that right.<sup>42</sup> So far, the Court has been reluctant to identify such a positive obligation. For example, in the *Appleby* case, where activists were barred from protesting in a publicly accessible, yet privately owned shopping centre, the Court found that the right to free speech did not override the owner's property rights.<sup>43</sup> This outcome has been interpreted as a sign of reluctance on behalf of the European court of Human Rights to grant rights of access to specific private venues or forums.<sup>44</sup>
- 19 Despite this reluctance, a positive obligation to enable access to the media is not entirely without precedent in ECHR case law. In *Verein gegen Tierfabriken*, a Swiss private broadcaster refused to accept television commercials on animal rights due to a state-wide ban on political advertising, and the Court considered this a breach of the Swiss state's positive obligations.<sup>45</sup> Central to the Court's findings was the broadcaster's monopolistic position: it considered that, in order to reach the entire Swiss public, the NGO 'had no other means than the national television programmes of the Swiss Radio and Television Company at its disposal, since these programmes were the only ones broadcast throughout Switzerland'.<sup>46</sup> This decision shows that, at least for national television networks, the principle of media pluralism demands of States a positive obligation to ensure non-discriminatory and fair access to the audio-visual platform, even where this platform is operated by private actors. This is in line with previous case law from the

*Informationsverein Lentia* judgment, which found that a state monopoly on broadcasting, precluding access by private parties, is a disproportionate limitation on the freedom of expression.<sup>47</sup> When a private organisation dominates the market, a lack of viable alternatives may also trigger increased duties to guarantee access in the interest of media pluralism and public debate.<sup>48</sup> In such cases, the Court found, 'regard must be had to the "fair balance" that has to be struck' between the community interest and individual rights.<sup>49</sup>

- 20 It should be noted, however, that the precise scope of this positive obligation is unclear, since an almost identical policy on political advertising in the UK did not amount to a violation in *Animal Rights Defenders*.<sup>50</sup> In distinguishing the UK situation from the Swiss, the Court pointed to the availability of alternative media, and social media in particular, in order for NGOs to reach their audience: 'Importantly, the applicant has full access for its advertisement to non-broadcasting media including the print media, the internet (including social media) as well as to demonstrations, posters and flyers. Even if it has not been shown that the internet, with its social media, is more influential than the broadcast media in the respondent State, those new media remain powerful communication tools which can be of significant assistance to the applicant NGO in achieving its own objectives.'<sup>51</sup> While posters and flyers were not mentioned as viable alternatives to television advertising in *VgT*, they were proposed as such in *Animal Rights Defenders*. While it is not an outright reversal of the findings in *VgT*, the Court does seem to have steered in a different direction by de-emphasising the importance of equal access and highlighting the availability of alternative media.<sup>52</sup> To reconcile these judgments, one might conclude that the interchangeability of different media is to be determined on a case-by-case basis. Evidently, the strict monopoly on broadcasting in Switzerland triggered a positive obligation, but the more diverse media landscape in the UK did not necessitate positive state action.
- 21 In light of the above, a state obligation to prevent undue interference by SNSs could theoretically be construed in cases where such a platform provides the only viable expressive opportunity. There is an interesting discussion to be had as to the functional equivalences and differences between social media and other modes of expression<sup>53</sup>, which mostly falls outside the scope of this article. However, some unique characteristics of social media, which might rule out other alternatives, deserve mention. Firstly, social media are interactive forms of expression, allowing for community feedback, debate and organisation to a degree unseen in television or print. Secondly, social media can allow access to new audiences, both demographically and geographically. Thirdly, social media tend to have significantly lower financial

barriers; creating a user or organisation page, or contributing to the pages and groups of others, is generally free, whereas access to television or print media can involve substantial costs. These are but a few of the reasons why social media should not always be treated interchangeably with alternatives such as broadcasting or pamphleteering.<sup>54</sup>

- 22 It should be somewhat obvious that SNSs offer unique affordances absent in other media, but to construe a right to access one particular platform, one would also have to establish that SNSs are not interchangeable *inter se*. After all, end users wishing to contest the removal of their content from Facebook might find a suitable alternative in Twitter. An important element to consider here is that SNSs generally do not provide direct contact with an audience, but instead require users to build up a network of friends or followers who are interested in their activities.<sup>55</sup> If a committed Facebook user with thousands of friends finds himself suddenly constrained by content policies, other platforms on which he is less well established may not directly provide a viable alternative.<sup>56</sup> Furthermore, SNSs have varying purposes ranging from professional networking (LinkedIn) and online dating (eHarmony) to Facebook's social functions and Twitter's more public forms of exchange, with different audiences harbouring different expectations. It would follow that SNSs are not necessarily interchangeable, and that end users may lack viable alternatives if removed from a particular service.
- 23 Applied to social media, the test of 'viable alternatives' first mentioned in *Appleby* speaks to a very real concern regarding the position of online intermediaries. In the same way that duties and responsibilities were created for television monopolies in *Informationsverein Lentia* and *Verein Gegen Tierfabriken*, the remarkable influence and dominance of a few key players on the SNS market might justify a re-examination of their obligations towards the public. However, the more recent decision in *Animal Rights Defenders*, where pamphleteering and social media were proposed as alternatives to television advertising, suggests that the significance of functional differences between media has not (yet) had a strong effect on the Court's assessment of positive obligations to ensure access. Since this judgment, the importance of new media, and social networks in particular, has only increased, and yet we are left with no guiding principles to assess the unique (and diverse) characteristics of these platforms. If the Court is indeed inclined to treat such a broad range of communicative methods as functional or at least interchangeable equivalents, a right to access social network services would be difficult to support.
- 24 The 'living instrument' doctrine requires Convention rights to be interpreted dynamically in response to

societal and technological developments.<sup>57</sup> In this light, the rise of social media might warrant a re-assessment of case law on positive obligations which would extend its applicability beyond strict monopolies and place greater stock in the need for access to online forums. One problem with such an approach, however, is that it remains unclear exactly how this state obligation might be fulfilled in practice. The Court has always refrained from demanding specific forms of intervention. For instance, when the Swiss government argued in *VgT* that the claimant's demands were tantamount to claiming a 'right to broadcast', the Court simply responded as follows:

*"The Court recalls that its judgment is essentially declaratory. Its task is to determine whether the Contracting States have achieved the result called for by the Convention. Various possibilities are conceivable as regards the organisation of broadcasting television commercials; the Swiss authorities have entrusted the responsibility in respect of national programmes to one sole private company. It is not the Court's task to indicate which means a State should utilise in order to perform its obligations under the Convention."*<sup>58</sup>

- 25 Thus, while the Court could hypothetically conclude that certain content removal decisions trigger positive State obligations, it is unlikely to go so far as to outline specific remedies. The Council of Europe's standard-setting activities through the Parliamentary Assembly and Committee of Ministers could play an advisory role in setting out appropriate measures for the protection of social network users, but ultimately the responsibility for the implementation of effective safeguards, and the discretion as to their means, lies with the Member States themselves.
- 26 As stated above, the protection of end users against intermediary censorship would require the uninterrupted provision of these services, despite contractual grounds for termination or suspension thereof. As such, state action in this area would entail a limitation of the freedom of contract. Other fields where positive obligations for the protection of free speech have been identified, such as television broadcasting, or employment law, have traditionally had strong limits placed on contractual freedom in favour of sector-specific regulation: for instance, the labour law concept of 'unfair dismissal' at issue in *Fuentes Bobo* places limits on the conditions under which employers can terminate an employment relationship. The positive obligation at stake simply entailed the manner in which this existing rule was to be interpreted in light of free speech considerations. Similarly, the decisions by private broadcasters in *Tierfabriken* were also subject to *sui generis* regulation, and their refusal of certain content was due to a state prohibition on political advertising. Again, an altered interpretation or slight amendment of the existing rules would suffice to fulfil the State's positive

obligations. While the above cases could be treated within existing regulatory paradigms, it is unclear through what instrument a Member State might protect their citizens from intermediary censorship on the internet. Although this does not necessarily speak against the necessity for such intervention, the lack of tried and trusted regulatory tools for the protection of end user rights might discourage the Court from identifying a positive obligation in these scenarios.

- 27 In conclusion, a review of judgements reveals that the ECHR's level of protection for free speech rights in private relationships lacks clarity; the case law is limited to a handful of judgments, some of which were made prior to significant changes in the media landscape.<sup>59</sup> Cases such as *VgT* problematise the strict distinction between private and public forums found in *Appleby*. However, this more graduated approach based on viable alternatives has few other precedents and thus lacks sufficient clarity for a predictable application to social networks. Although an argument based on changes in the media landscape and the unique position of SNSs could support the protection of end users against arbitrary or unnecessary censorship, cases such as *Animal Rights Defenders* suggest that the Court has not yet arrived at such a nuanced treatment of functionally different forms of media, and is hesitant to call for free-speech based interventions in private relationships. Since the Convention cannot create horizontally enforceable rights, and rather operates through positive obligations, it is also unclear what precisely could be demanded of states in order to adequately safeguard end user rights.

## D. Blocking injunctions, third party notices and voluntary removal under the EU framework

- 28 European Union law plays an important role in regulating the content moderation policies of SNSs. The e-Commerce Directive sets out rules on content liability, which play an important role in shaping intermediary incentives for content moderation.<sup>60</sup> It also places limits on state interference through injunctive measures and (indirectly) provides the basis for notice and takedown procedures, which enable the enforcement of intellectual property rights and other third party interests online.<sup>61</sup> In addition, the Enforcement directive provides more specific guidance on IP-based injunctions. Through the CJEU's case law based on the Charter of Fundamental Rights of the European Union, this regime has also increasingly been made subject to fundamental rights considerations which can limit the competences of states and intermediaries in removing online content.<sup>62</sup> EU law can thus require, permit or proscribe content removal

by intermediaries. This section will explore this balance of rights and duties and evaluate the legal implications of content removal in the EU *acquis*.

- 29 Broadly speaking, when SNSs host content provided by their users, this activity qualifies as a 'hosting' service in the sense of Article 14 of the e-Commerce Directive. Recital 42 of the Directive indicates that intermediaries are eligible for this classification, so long as their activities are '*of a mere technical, automatic and passive nature, which implies that the information society service provider has neither knowledge of nor control over the information which is transmitted or stored*'.<sup>63</sup> The CJEU's decision in *SABAM / Netlog* shows that social media platforms are covered by this provision.<sup>64</sup> The most important consequence of this classification is that the intermediary cannot be held liable for illegal content uploaded by their end users.<sup>65</sup> Only once the intermediary obtains knowledge of illegal information, and then fails to expeditiously remove it, does he lose this protection for civil and criminal liability.<sup>66</sup> This rule applies to all forms of illegal information, whether it be copyright infringement, child pornography or extremist hate speech.<sup>67</sup> By protecting intermediaries from legal risks created by end user wrongdoing, this exemption can be seen as an important pre-condition for the viability of user-generated content business models, and by extension for the protection of online speech through SNSs.<sup>68</sup>
- 30 While the 'safe harbour' of Article 14 of the e-Commerce Directive protects intermediaries from liability for end-user content, this rule does not prevent the imposition of blocking injunctions through court orders or administrative orders. Recital 45 states that the safe harbours '*do not affect the possibility of injunctions of different kinds; such injunctions can in particular consist of orders by courts or administrative authorities requiring the termination or prevention of any infringement, including the removal of illegal information or the disabling of access to it.*' In some cases, EU law demands the creation of injunctive relief, such as in the interest of copyright enforcement and related rights under the Directive on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market (InfoSoc Directive) and the Directive on the enforcement of intellectual property rights (Enforcement Directive).<sup>69</sup> These directives also set more specific conditions for the application of such measures. However, the e-Commerce Directive's reference to 'any infringement' shows that national law may also allow for injunctions on other grounds, such as defamation or criminal hate speech. Recital 45 indicates that both courts<sup>70</sup> and administrative bodies may issue such orders.
- 31 These state-issued encroachments on the 'safe harbour' are in turn restricted by article 15 of the e-Commerce Directive. This provision prescribes

that States may not impose on hosting providers any ‘general obligation (...) to monitor the information which they transmit or store, nor a general obligation actively to seek facts or circumstances indicating illegal activity’. As such, blocking injunctions must be limited to specific content and cannot apply to the service as a whole. For instance, the Court ruled in *SABAM / Netlog* that Belgium’s imposition of a general filtering obligation on a social media service, aimed at detecting and preventing copyright infringements throughout the entire network, fell afoul of this provision.<sup>71</sup> The neutral position of intermediaries is thus protected under EU law, and intermediaries cannot be compelled to actively search for illegal content on their networks.<sup>72</sup>

- 32 More specifically, in the context of intellectual property claims, the Enforcement Directive shows that I.P. injunctions directed at third parties “must be ‘effective, proportionate and dissuasive’ and that the national rules governing them must be ‘designed in such a way that the objective pursued by the Directive may be achieved’.”<sup>73</sup> At the same time, however, the measures they impose must be ‘fair and proportionate and must not be excessively costly’ and must ‘not create barriers to legitimate trade’.<sup>74</sup>
- 33 The above provisions show that undue state interference with online intermediary services may be contrary to the EU’s aim of market competition. In addition, their preambles show that these requirements must be read in light of fundamental rights considerations. Recital 9 of the E-Commerce Directive states that ‘The free movement of information society services can in many cases be a specific reflection in Community law of a more general principle, namely freedom of expression as enshrined in Article 10(1) of the Convention for the Protection of Human Rights and Fundamental Freedoms, which has been ratified by all the Member States; for this reason, directives covering the supply of information society services must ensure that this activity may be engaged in freely in the light of that Article.’ Furthermore, Recital 2 of the Enforcement Directive states that ‘[the protection of intellectual property] should not hamper freedom of expression, the free movement of information, or the protection of personal data, including on the Internet.’ Thus, the EU law on intermediary injunctions calls for state orders to be grounded in a minimum level of free speech protection.
- 34 The need to weigh the purposes of blocking against competing fundamental rights has been further crystallised in the CJEU’s case law. Borrowing from ECHR’s terminology (see above, section III), they have made injunctions aimed at copyright enforcement subject to a ‘fair balance’ requirement: ‘where several fundamental rights are at issue, the Member States must, when transposing a directive, ensure that they rely on an interpretation of the directive

which allows a fair balance to be struck between the applicable fundamental rights protected by the European Union legal order.”<sup>75</sup> This concept was first introduced in the case of *Promusicae*, and later repeated in cases such as *Scarlet / SABAM*, *Netlog / SABAM* and *UPC Telekabel*.<sup>76</sup> In these cases, injunctions against internet service providers aimed at copyright enforcement were struck down for requiring general filtering measures by the intermediary, thereby impinging on end users’ right to privacy and freedom to receive or impart information, as well as the intermediary’s freedom to conduct a business. The Court found that, in order to strike a fair balance, injunctions must not “unnecessarily deprive internet users of the possibility of lawfully accessing the information available”.<sup>77</sup> While these cases all concerned the enforcement of intellectual property rights, the Court’s generalised formulation does not seem to suggest that this ‘fair balance’ test would be absent for injunctions on the basis of, for instance, defamation claims. When commanding removal by intermediary services, state authorities and courts must take end users’ fundamental rights into consideration, including their right to free speech.

- 35 The Court in *UPC Telekabel* decided that, insofar as state injunctions do not specify the technical measures required, the targeted intermediaries themselves must also attempt to strike a ‘fair balance’ between the public interest and their end users’ fundamental rights.<sup>78</sup> Furthermore, *locus standi* must be granted to affected end users wishing to have such implementing actions assessed as to this ‘fair balance’.<sup>79</sup> This judgment is somewhat groundbreaking in applying the ‘fair balance’ test not only to state actors, but also to private intermediaries.<sup>80</sup> However, the ‘fair balance’ requirement has not yet been applied to strictly voluntary acts of removal, where the intermediary acts independently to moderate content on its services.<sup>81</sup>
- 36 The free speech implications of content removal in the absence of direct state interference have not yet been examined by the Luxembourg court. However, such intermediary interventions are not uncommon on social media: they may occur as a matter of policy, such as Facebook’s ban on nudity, but may also be instigated by third parties through the process known as ‘notice and takedown’. Under the e-Commerce Directive, intermediaries incur liability for content through knowledge of its illegality, and third parties can trigger a removal obligation through notification: once informed of illegal content on their servers, hosting providers must then ‘expeditiously remove’ said information, or render themselves liable for its illegality.<sup>82</sup> The major social networks have put in place procedures to facilitate this process, with specific pages and forms intended for the reporting of illegal or infringing content.<sup>83</sup> Thus, the EU’s e-Commerce regime still forces SNSs to moderate

end-users for illegal content and infringements, even absent direct state interference. The notice and takedown procedure allows third parties to trigger content removal duties, backed up by the threat of subsequent litigation. If intermediaries must take into account their end users' fundamental rights when implementing court orders, does this duty also apply in the takedown phase preceding judicial enforcement? In other words does the 'fair balance' requirement also apply to the handling of takedown notices?

- 37 'Notice and takedown' can be seen a consequence of the e-Commerce Directive's conditions for liability, but is not directly mentioned or specified. The 'takedown notice' has no independent meaning under the EU *acquis communautaire*, and is not subject to any substantive or formal requirements. As such, the act of 'notification' does not necessarily provide evidence of actual knowledge. While notices can indeed be a factor in establishing the intermediary's knowledge they may lack sufficient (correct) information for the intermediary to identify an illegality: the CJEU decided in *L'Oréal v. Ebay* that notices 'cannot automatically preclude the exemption from liability provided for in Article 14 of Directive 2000/31, given that notifications of allegedly illegal activities or information may turn out to be insufficiently precise or inadequately substantiated'.<sup>84</sup> Thus, the intermediary is expected to make an independent assessment of the facts available to him in order to determine the legality of hosted information. The Directive permits Member States to set further rules or procedures for notification, but a majority have not implemented formal notice procedures for hosting intermediaries.<sup>85</sup>
- 38 Due to the non-formalised nature of takedown procedures, it is difficult to distinguish in practice between content removal instigated through third party notices, and voluntary interventions based on the platform's Terms of Use. SNSs may decide to comply with a meritless defamation claim because the flagged content breaches their ToU prohibition of bullying or harassment. Thus, notices may result in the removal of legal content.<sup>86</sup> Conversely, risk-avoidant intermediaries may decide to remove illegal content, even though the referring notice contains insufficient information to render them liable. In other words, SNS content removal instigated through the notice and takedown process does not always correspond to a legal duty. It cannot be distinguished *a priori* from the 'voluntary', policy-based conduct moderation. It would therefore be difficult to argue that implementations of takedown requests are categorically subject to a different fundamental rights standard than other forms of voluntary content moderation by SNSs.
- 39 Admittedly, recital 46 of the e-Commerce Directive requires that the removal of or disabling of access to

illegal content has to be undertaken in the observance of the principle of freedom of expression. However, in the absence of more concrete implementation in national procedures, or an operative EU provision to the same effect, this notion has not been translated into a directly effective safeguard. Currently, the identification of direct state interference remains key for an appeal by end users to their free speech rights under the EU framework.

- 40 I argue that this strict distinction between public and private forms of intervention is problematic. After all, while no state action is necessary to complete a notice and takedown process, its efficacy is dependent on the possibility of subsequent state enforcement through litigation. Furthermore, these possibilities are created by the limitations on content illegality and intermediary liability enacted by Member State legislators and interpreted by its judiciary. While state powers are not being *exercised* to achieve content removal in such instances, they are at least being *invoked*. Conversely, the discretionary powers of private SNSs can be, and have been, applied to further governmental interests (as will be further explored in section VI). Through this legal realist perspective a strict distinction between publicly enforced and 'voluntary' removal becomes incoherent.<sup>87</sup> State responsibility for private behaviour is not a binary distinction, but a matter of degree.
- 41 An adequate protection of online expression therefore requires a more nuanced approach than a strict public / private distinction can provide. The ECHR's case law on positive obligations could provide such a method: To the extent that the EU framework can result in the neglect of end users' free speech rights by intermediaries, one could argue that Member States are breaching their positive obligation to ensure the effective exercise of this right.<sup>88</sup> Safe harbour provisions – as a precondition for the viability of SNS business models – can then be seen as an indirect protection of online expression. Limitations and exceptions thereto, such as the 'actual knowledge'-based takedown regime, should then be limited in scope so as to curb third party influence on content moderation policies. The EU regime allows states to negotiate this balance between third party claimants and the protection of intermediaries. Where it may currently be lacking, however, is the protection of *end users* against voluntary forms of removal; lacking meaningful contractual safeguards, they have little ground to contest the decisions of SNSs and have them reviewed as to their compatibility with free speech principles.
- 42 The protection of end users against content removal is currently not significantly greater under the EU framework than under the Convention. Both place strict requirements on interferences by public

authority, but leave the freedom of private online services largely intact. A significant difference, however, is that the Charter elevates this freedom to the status of a fundamental right. The ‘freedom to conduct a business’ under Article 16 states that ‘*The freedom to conduct a business in accordance with Community Union law and national laws and practices is recognised*’. Voluntary content removal, as a business practice otherwise in accordance with Union law, therefore falls under the protective scope of this right.<sup>89</sup> Accordingly, the ‘fair balance’ to be achieved under EU law would have to take into account not only the interests of end users, and the interest in restricting certain forms of speech, but also the interest in protecting the SNS operator’s freedom to conduct its service. Although the Charter must provide an equivalent minimum level of protection to free speech as its Convention counterpart, this additional provision suggests that EU Member States must exercise restraint when limiting the content moderation competences of SNS companies.

## E. Voluntary removal and State Censorship

- 43 The above examination of intermediaries’ rights and obligations shows that there are few legal limitations placed on the moderation and removal of end user’s SNS content. While official state action aimed at restricting speech is subject to judicial review under a developed case law on freedom of speech, the private nature of SNSs allows them a far greater range of discretion. The current fundamental rights framework does not seem to imply a right for end users to have their speech carried by SNSs. Barring an even further expansion of the Charter’s ‘fair balance’ duties for intermediaries or the Convention’s positive obligations, end users are thus left with a lack of protection against unwarranted, unforeseeable or arbitrary censorship by intermediary actors. This lack of safeguards against voluntary content removal creates two distinct, but closely related threats: Firstly, it can allow state censors to circumvent existing safeguards through systems of informal pressure. Secondly, intermediaries may find it in their own interest to remove content and at times disregard free speech principles (to be discussed in section VII). These two explanations of content removal decisions may often coincide or overlap, and this article does not intend to provide a monocausal account for specific incidents. Nevertheless, they provide a useful heuristic to discuss the threats caused by the lack of accountability for voluntary removal, as well as possible regulatory solutions.
- 44 To start with a form of state interference which EU law prohibits: governments outside the EU have placed great pressure on SNSs through the threat of ISP-level blocking. Not to be confused with injunctions on end user content directed at the SNS itself, these orders are directed at local Internet Service Providers and block local audiences access to the website as a whole. The Turkish government has twice used this measure. Firstly, Twitter was blocked in 2014 when it was used to spread a torrent of audio recordings implicating the prime minister and his inner circle in an alleged corruption scandal.<sup>90</sup> Secondly, Twitter, Facebook and YouTube were all blocked in 2015 for hosting photos of a prosecutor who was taken hostage by militants in Istanbul.<sup>91</sup> The injunction was revoked eight hours later when the companies complied with the order.<sup>92</sup> These incidents show how, even where companies hold no local assets, ISP-level blocking can be applied to demand compliance with national laws or rules and alter their community policies.<sup>93</sup> In the EU, ISP-level blocks are common practice in the context of copyright enforcement, and have recently also been applied for anti-terrorist purposes.<sup>94</sup> However, no EU state has yet directed such measures against major SNS websites. Furthermore, EU law seems to preclude this option: the blocking of an entire website would affect a high volume of legitimate traffic, and is therefore unlikely to strike a ‘fair balance’ required under the Charter.<sup>95</sup>
- 45 Although EU states cannot strong-arm SNSs in the same way that other governments have done, they have found other ways to influence content policies. Described by Benkler as ‘regulation by raised eyebrow’, and by Birnhack and Elkin-Koren as ‘the invisible handshake’, approaches where governments aim to shape private policy indirectly, can achieve a largely similar result.<sup>96</sup> This soft power approach involves a combination of publicly-voiced appeals to corporate responsibility by senior officials, and close, intransparent ties between intermediaries and law enforcement agencies.<sup>97</sup> Ostensibly voluntary measures by intermediaries may then be applied to further government interests, without revealing state involvement or triggering constitutional safeguards.<sup>98</sup>
- 46 Examples of political pressure on SNS companies have been frequent, in particular in the context of anti-terrorist activities. In the wake of the Rigby murders, the UK Security Committee called on Facebook to proactively monitor its community for terroristic content in cooperation with law enforcement.<sup>99</sup> Following the *Charlie Hebdo* attacks, and in response to an increased online presence of extremist organisations such as the Islamic State, heads of government throughout Europe have made public statements calling on SNSs to contribute to anti-terrorism efforts. At the World Economic Forum in 2015, President Francois Hollande called on corporations to “fight terror,” stating: “The big operators, and we know who they are, can no longer close their eyes if they are considered accomplices

of what they host. We must act at the European and international level to define a legal framework so that Internet platforms which manage social media be considered responsible, and that sanctions can be taken.”<sup>100</sup> Broad governmental support for an increased responsibility of online intermediaries in the fight against terrorism was reaffirmed in a Joint Statement of EU ministers, which stated that: ‘the partnership of the major Internet providers is essential to create the conditions of a swift reporting of material that aims to incite hatred and terror and the condition of its removing, where appropriate/possible.’<sup>101</sup> These are but a few examples of government officials publicly calling for SNSs to take a specific course in their content moderation policies.

- 47 These calls have not fallen upon deaf ears. In 2015, Facebook officially expanded its content removal policy to a broader range of terrorist activities: in a statement to the BBC, Monika Bicket, Facebook’s global head of content policy, explained that ‘we now make clear that not only do we not allow terrorist organisations or their members within the Facebook community, but we also don’t permit praise or support for terror groups or their acts or their leaders, which wasn’t something that was detailed before.’<sup>102</sup> For its part, Twitter also undertook efforts to purge their network of ISIS beheading videos<sup>103</sup>: In 2015 they tripled the size of their content moderation team<sup>104</sup>, expanded their definition of prohibited ‘violent or threatening’ behaviour<sup>105</sup>, and began experimentation with automated algorithms for the filtering of abusive or inappropriate content.<sup>106</sup> While France’s ban on the glorification of terrorism is soon to be contested before courts by civil rights groups, and a Dutch proposal to the same effect failed to secure parliamentary support<sup>107</sup>, Facebook or Twitter can make such changes without any true accountability. For governments, the advantage of ‘raised eyebrow’ methods is clear: Why go to all the effort of passing a law – with all the constitutional, political and procedural hurdles this involves – when intermediaries can be persuaded to adopt such rules unilaterally and with no clear form of oversight?
- 48 The power of public statements made by government actors is also evident in the controversy surrounding Wikileaks’ diplomatic cables revelations. Yochai Benkler has convincingly documented how the US government made a concerted effort to ostracize Julien Assange as an enemy of the state, with vice-president Biden describing him as a ‘high tech terrorist’ and a senator publicly calling on other governments and corporations to distance themselves from their ‘illegal, outrageous and reckless acts’.<sup>108</sup> Not long after, a large number of intermediaries such as hosting providers and payment platforms withdrew their services to Wikileaks.<sup>109</sup> Rather than issuing binding orders, which would be subject to free speech and due process safeguards, states can

play on the responsibility of intermediaries such as SNSs to do the dirty work, which many will be willing and legally permitted to do ‘at the mere whiff of controversy’.<sup>110</sup>

- 49 Concurrent to these publicly voiced emphases on the responsibility of intermediaries and the editorialisation of online content policies, we also see initiatives for closer cooperation between law enforcement and SNS content moderators. Echoing the EU Ministers’ focus on intermediary responsibility and the need for swift reporting mechanisms, the EU’s counter-terrorism coordinator, Gilles de Kerchove, proposed to have experts from member states flagging terror-related content, stating “We have to help them, and refer to them, and signal content (...) Each member state should have a unit with people trained to do that.”<sup>111</sup> He cited the UK’s Counter-Terrorism Internet Referral Unit (CTIRU) as a best practice for other Member States to emulate: ‘Member States should consider establishing similar units to the UK CTIRU and replicate relationships with the main social media companies to refer terrorist and extremist content which breaches the platforms’ own terms and conditions (and not necessarily national legislation).’<sup>112</sup> Quite unambiguously, this statement shows how Terms of Use and community policies are coming to supplant legislation as a means to regulate online speech, and that governments are eager to make use of this method. Furthermore, the terminology of ‘helping’ and ‘referring’, rather than ordering or demanding, implies the intention to hold intermediaries responsible for such decisions, rather than law enforcement itself. This ‘voluntaristic’ approach championed by Kerchove and the CTIRU has proven to be highly effective in achieving content takedown. According to the same statement, SNSs have voluntarily removed 72,000 pieces of terrorist content following referrals from the CTIRU because they have agreed that the content represents a breach of their rules.<sup>113</sup> However, as the above has shown, it is precisely this non-statutory and voluntary quality, exacerbated by the politically-charged and non-specific aim of ‘counter-terrorism’<sup>114</sup>, which renders these systems susceptible to abuse.<sup>115</sup> Especially in the current political climate, intermediaries may find their demands difficult to refuse.
- 50 An important source of information regarding governmental pressure on SNSs has been the release of so-called ‘transparency reports’. Since 2012, Twitter has published data on the governmental information and removal requests they receive worldwide, and the manner in which they have been processed.<sup>116</sup> Facebook, LinkedIn and a range of other significant online intermediaries have since followed suit.<sup>117</sup> Twitter’s reports distinguish between court-ordered injunctions and other government requests, showing that, for example, Germany issued one court order and forty-two removal requests,

and that Twitter blocked content in 37% of these instances.<sup>118</sup> Twitter's reports confirm the trends of increased interference described above, with total annual requests worldwide increasing from 48 in 2012 to 1099 in 2014. While Russia and Turkey are chiefly responsible for this explosive increase, they are quickly followed by the EU member states. For 2014, Twitter reported 11 court orders and 249 government requests from EU member states.<sup>119</sup> Facebook does not make a categorical distinction between these interactions, and reports a total of 72 EU Member State removal requests for the second half of 2014.<sup>120</sup>

51 By drawing attention to government interference, these reports could encourage states to exercise restraint in policing online communities. However, they lack a clear definition of what constitutes an extra-judicial removal request, and seem to omit the work of referral units such as the CTIRU. While Facebook and Twitter report a total of 15 and 44 removal requests from the UK since 2012<sup>121</sup>, the CTIRU claims to have achieved 29,000 content removals by 'social media and other parts of the internet industry', which explicitly include Twitter and Facebook.<sup>122</sup> Are we to conclude that these referrals do not fall under the definition of government requests for the purpose of these reports? If so, an essential element of government influence on online speech is being omitted, illustrating a limited interpretation of what constitutes state interference in the attitudes of social media companies.<sup>123</sup>

52 Individual examples of undue government influence on SNS policies are inherently difficult to illustrate with concrete examples, since they are aimed at denying state involvement and deferring responsibility to private parties. However, there are certain incidents which can serve to illustrate the vulnerability of the current system. For instance, in the run-up to the UK royal wedding in 2011, controversy arose surrounding Facebook's decision to remove a number of pages dedicated to organising protest rallies.<sup>124</sup> These removals coincided with the arrest of involved activists, and some alleged that Facebook had suppressed legitimate political dissent in cooperation with law enforcement authorities.<sup>125</sup> Facebook denied these allegations, and claimed that the removal decision was due to the protestor's use of pseudonyms - a breach of the Facebook Terms of Use. The removal was simply part of a 'routine check' for such compliance, unrelated to coinciding law enforcement measures or the controversial, political nature of the content involved.<sup>126</sup> Admittedly, direct government ties have not been proven in the above case, but it highlights the potential for abuse present in the current system. So long as end users cannot establish that Facebook's removal is the direct consequence of state interference, rather than a voluntary decision based on community policy, governments could stifle political speech without

direct accountability. And if governmental influence on SNSs policy is informal, non-transparent and non-binding, establishing this link could be very difficult indeed.

53 Through the self-regulatory body known as the Global Network Initiative (GNI), some SNSs and other online intermediaries have taken it upon themselves to limit governmental influence and protect their end users' privacy and freedom of expression. Founded in 2008, this organisation, which counts Facebook and LinkedIn among its members, established principles and guidelines for the fair and transparent processing of removal requests.<sup>127</sup> Their principles include the following:

- Require that governments follow established domestic legal processes when they are seeking to restrict freedom of expression.
- Interpret government restrictions and demands so as to minimise the negative effect on freedom of expression.
- Interpret the governmental authority's jurisdiction so as to minimise the negative effect on freedom of expression.
- Seek clarification or modification from authorised officials when government restrictions appear overbroad, not required by domestic law or appear inconsistent with international human rights laws and standards on freedom of expression.<sup>128</sup>

54 The GNI demands that intermediaries minimise the harmful effect of government requests, and resist non-formalised and voluntary measures (i.e. those not required by law).

55 The visibility and transparency of these requests towards the end user are further preconditions for effective protection of their rights. After all, end users must be made aware of the fact that content removal has been caused by government intervention before they can consider appealing these actions. To quote Baudelaire, "the finest trick of the devil is to persuade you that he does not exist".<sup>129</sup> The GNI prescribes that end users must be given 'clear, prominent and timely notice to users when access to specific content has been removed or blocked by the participating company or when communications have been limited by the participating company due to government restrictions.'<sup>130</sup> Furthermore, 'Notice should include the reason for the action and state on whose authority the action was taken.'<sup>131</sup>

56 While the GNI principles, if fully observed, would go a long way in protecting SNS end users, the organisation's main shortcoming is its lack of enforcement competences. The GNI does carry

out compliance reviews of its members, yet it does not have the competence to sanction breaches or otherwise enforce their guidelines. So far, only one such assessment has been carried out, in which all three companies reviewed – Yahoo!, Microsoft and Google – were found to be in compliance. It is unclear, however, what the consequence of a negative outcome would have been. In the absence of binding self-regulatory safeguards, the GNI principles might therefore best be seen as a set of best practices.<sup>132</sup> In any case, while Facebook and Twitter have embraced the GNI's stance on end users' speech rights, section II has shown that this is not reflected in their end user contracts.

- 57 Another non-binding document which outlines guiding principles for the treatment of takedown requests is the recently published Manila Principles, drafted by a worldwide coalition of digital rights NGOs as a set of guidelines for the protection of freedom of expression in communications facilitated by internet intermediaries.<sup>133</sup> Its recommendations are similar to those of the GNI. However, where the GNI recommends that 'governments follow established domestic process when they are seeking to restrict freedom of expression', the Manila Principles go even further by prohibiting any extra-judicial measures: "Governments must not use extra-judicial measures to restrict content. This includes collateral pressures to force changes in Terms of service, to promote or enforce so-called "voluntary" practices and to secure agreements in restraint of trade or in restraint of public dissemination of content."<sup>134</sup> Even more so than the GNI, these principles take a hard stance on the 'regulation by raised eyebrow'. And even more so than the GNI, they are of a strictly normative nature and create no guarantees for end users in practice.
- 58 The principles embraced by SNSs and civil society in these documents, with their emphasis on codified, transparent government request procedures, is quite different from the government methods described above. Governments have proven themselves willing, and often able, to appeal to SNSs' corporate responsibility and their voluntary removal capacities in order to further such goals. Rather than operating through referral or assistance which would rely on SNSs' voluntary removal competences, interaction with law enforcement agencies should be formalised so that requests result in binding, specific and transparent removal *orders*. This approach places the responsibility with law enforcement authorities to follow national legal processes, and with SNSs to resist non-binding demands.
- 59 An emphasis on rule of law principles and transparency requirements could go some way in curbing the circumvention of vertical free speech safeguards. However, it seems impossible to rule out this risk completely. What motivates intermediary removal decisions is fundamentally

difficult to ascertain, and government demands may inevitably play a role in these deliberations. Can we really prohibit figures such as prime ministers from publicly mentioning their dislike of violent content online? This seems neither feasible nor desirable (it would sooner be a restriction of the politicians' free speech). Comments such as these play out in a broader context of public discourse, where the perceived preferences of the social network community, or the public as a whole, can be as much a motivating factor as the demands of state authorities. Although states undoubtedly have powerful voices in affecting these deliberations, they must be seen in a broader context of a multiple actors who can influence SNSs through the mercurial process known as a 'public debate'. Instead of focusing solely on the role of government in SNS content moderation, the online speech would also benefit from increased accountability for the SNSs themselves.

## F. Voluntary Removal and Private Censorship

- 60 The previous section has focused on states and their governments as the principal drivers of censorship, and on SNS operators as their occasional partner in these endeavours. However, it would be a mistake to assume that the SNS services are otherwise neutral parties that will moderate content only to the extent that governments can persuade them to. Their own commercial goals and incentives may also lead them to stray from strict non-intervention. Indeed, many common forms of content moderation do not correspond to a legal duty, such as the prohibition of pornography or bullying. And yet, these rules are largely accepted - even expected - by social network users. Voluntary content removal is often legally and commercially viable for intermediaries, and is linked to the community's needs and demands as well as the leverage of other private parties. This section will consider how the commercial incentives of SNS operators relate to the free speech rights of their users.
- 61 In some cases, such as online extremism, governments' political aims and intermediaries' commercial incentives can be seen to overlap. To start with an example: what motivated Twitter's large-scale effort to fight IS-related content on its network? From the previous section, it appears that government demands may have influenced Twitter's behaviour. However, it can also simply be seen as a response to user demand. This seems to have been the Guardian's interpretation, when it reported that '*people do not want to see this imagery, and media platforms are responding.*'<sup>135</sup> The low popularity of ISIS' political programme and the shocking nature of their propaganda content contributed to an environment where the Twitter community readily accepted, and

in many cases actively called for, their exclusion from the network.<sup>136</sup> In this light, voluntary content removal might be explained primarily as a part of their service, or as a means of PR, rather than as government-backed suppression.

- 62 End users' demands and expectations should not necessarily forestall free speech concerns. After all, the fundamental right of free speech serves to protect not only majoritarian consensus, but also minority positions.<sup>137</sup> These might not be as favourably treated where intermediaries simply respond to the whims of 'the public'. The freedom of SNSs to determine their own content policies has allowed them to respond decisively to a deluge of shocking and anti-democratic terrorist content, but it has also permitted, for example, Facebook's hard-line sexual conservatism, which has resulted in questionable removal decisions regarding displays of gay and lesbian affection<sup>138</sup> and images of breastfeeding and artistic nudity.<sup>139</sup> For example, Facebook's refusal to host an image of Courbet's nude 'L'Origine Du Monde' has sparked controversy in France.<sup>140</sup> Regardless of whether one approves of Facebook's sexually conservative approach, or Twitter's tough stance on terrorist content, these policies reflect the preferences of a dominant community culture which can marginalise diverging practices and attitudes. The SNS operator's power to remove content at the behest of the community's (perceived) wishes, subject to the shifting political and moral attitudes, introduces an unforeseeable threat to dissent and pluralism.
- 63 Another aspect to consider is the susceptibility of SNSs to private-sector demands. Through takedown notices, third parties can exert pressure on SNS removal policies. Numerous factors contribute to the likelihood of SNSs underrating their end users' free speech interests and unnecessarily complying with third party demands: the high volume of requests, particularly in the context of copyright enforcement<sup>141</sup>, and the correspondingly high transaction costs involved in adequately evaluating their claims;<sup>142</sup> legal uncertainty as to the conditions for 'actual knowledge' and intermediary liability; and end users' lack of contractual protection against content removal.<sup>143</sup> In many cases, risk-avoidance may then weigh heavier than the free speech interests of their (non-paying) customers. This imbalance of incentives speaks against an overreliance on the SNS operator as an arbiter between users and other private interests.<sup>144</sup>
- 64 These pressures from within and without the SNS may diverge from democratic considerations on the limits of acceptable speech. Even with sufficient constitutional safeguards against state co-optation, SNSs cannot necessarily be relied upon to observe free speech principles autonomously. A more complete protection from both public and private

interference must therefore focus on increasing the direct accountability of SNSs towards individual end users. Users should have some means to contest the unwarranted removal of their content - that is to say, removal which does not strike a balance between the removal ground and the end user's free speech rights. As described in section II, this balance requires a different calculus than that which is applied to state interference, with greater leniency as to the aims of intervention; for instance, while a national ban on pornography might be deemed excessive, an SNS could have legitimate reasons for maintaining such a policy. An important factor in determining the intermediary's discretion should be their degree of dominance, as reflected in the ECHR's case law.<sup>145</sup> As SNSs become more popular and influential as quasi-public forums, their removal policies should be held to stricter requirements.

- 65 Here, the GNI could provide some guidance as to best practices for SNSs. Although their rules are formulated so as to apply exclusively to the treatment of state orders, certain principles could also be extended to voluntary content moderation: content policies should be formulated clearly and interpreted restrictively, and content deletion should be limited to the infringing elements (such that an entire discussion thread, profile or page is not removed due to one prohibited comment or image).<sup>146</sup> For instance, the use of a false name may be prohibited, but it need not provide a ground for the complete deletion of an entire political protest page for all its followers.<sup>147</sup> Similarly, region-specific blocking should be preferred to total deletion (as the lesser of two evils).<sup>148</sup> Where possible, users are to be notified of their policy breaches before removal and be given the opportunity to respond and rectify policy breaches, as well as the ability to appeal decisions internally. These medium-specific factors could be taken into account when assessing the removal decisions in light of the freedom of expression.
- 66 These principles require a higher level of protection for end users than their contracts currently provide. If self-regulatory efforts such as the GNI cannot achieve this result, there is a need for the law to ensure such safeguards. Under the Charter, this could be achieved through the extension of the 'fair balance' duty on SNSs to include voluntary removal decisions. Under the ECHR, it would involve an appeal to positive state obligations to ensure an appropriate level of protection for end users. These fundamental rights-based approaches would have to be dealt with through end user litigation, and might require a significant re-examination of current doctrines. They are further problematised by the slow pace of jurisprudence in a rapidly changing media landscape, and, in the case of the ECHR, by a lack of horizontal effect. A more effective response would therefore involve active state regulation,

which might involve the prohibition of ToU clauses with overly broad content removal competences, or reversals of the burden of proof (an approach similar to EU consumer protection<sup>149</sup>). As Wauters *et al.* have argued, EU regulators could also improve contractual safeguards for SNS safeguards through the facilitation and encouragement of collective redress mechanisms.<sup>150</sup> It falls outside the scope of this article to determine whether these actions should occur at the national or EU level; whether some form of co-regulation is possible; and to which online services it should apply. However, the slow pace of fundamental rights case law and a lack of commercial incentives towards genuinely effective self-regulation may necessitate an operationalisation of such norms through public rulemaking.

## G. Conclusion

- 67 This article has revealed some chinks in the constitutional armour, and assessed the risks that they create. The current EU framework places a large degree of trust and responsibility in the hands of a few SNS companies, who are uniquely positioned to place boundaries on the tone and topic of public debate. This article has reviewed contractual provisions of two major SNSs, Facebook and Twitter, and found that they provide the operator with a broad, if not unlimited competence to remove content and terminate accounts. Many end users will therefore be unable to rely on their contractual relationship to contest disproportionate interferences with their online expression. Furthermore, network effects and information asymmetry hinder meaningful competition between services as to the level of speech protection granted.
- 68 The ECHR provides safeguards against the abuse and misuse of blocking injunctions and other coercive state measures directed at SNSs. Though the Convention's obligations are directed towards states, the theory of positive obligations also allows for the weighing of conflicting fundamental rights in private relationships. However, a review of such decisions reveals the Court's hesitance to acknowledge such obligations in the context of free speech, particularly in the context of access to privately-owned media. The unique affordances of SNSs and the dominant position of a few services, could lend credence to the invocation of positive obligations, but precedents such as *Appleby* and *Animal Rights Defenders* indicate the Court's reluctance to impose access duties on private parties.
- 69 The limitations on state injunctions are made more specific under EU law proper. The e-Commerce Directive sets out rules as to the foreseeability and proportionality of such interventions, and through the CJEU's case law the 'fair balance' requirements of
- the Charter have been concretely applied to IP-based blocking injunctions so as to protect the end users' free speech rights. However, the case law falls short of treating the free speech implications of voluntary removal by the intermediary. These 'voluntary' measures include removal achieved through notice and takedown procedures, which, due to their non-formalised nature, cannot be treated as a separate category *a priori*. For end users, the identification of direct state action currently remains essential for an appeal for their free speech rights against online services. It is submitted that the CJEU's jurisprudence should further explore the horizontal dimension of free speech rights online, and depart from a strict distinction between public and private responsibilities for the observance of these claims.
- 70 A more nuanced approach is required to protect users against unaccountable SNS content moderation. One concern is the susceptibility of such voluntary competences to state influence. This article has given practical examples of how EU governments have used a combination of political pressure and informal cooperation to effectuate change in SNS content policy, particularly in the context of anti-terrorist efforts. A shift from legislative regulation and formal injunctions to public-private collaborations allows state authorities to circumvent the traditional constitutional safeguards in place for their interferences with public discourse. While an adherence to rule of law principles and transparency requirements could go some way in curbing these developments, it is argued that these informal and indirect interactions based on 'raised eyebrows' and 'invisible handshakes' are inherently difficult to regulate, and that reform must therefore focus on increasing the platform's accountability towards the end user.
- 71 States are but one of many actors seeking to influence SNSs' content policies. To the extent that content removal is inspired by the demands and expectations of the end users, SNSs may be encouraged to sideline ideals of pluralism and dissent in favour of a majoritarian approach. To the extent that it is inspired by third party claims, transaction costs and the risk of liability may also discourage an adequate evaluation of free speech considerations. Therefore, even absent state interference, the commercial incentives of SNSs cannot be guaranteed to coincide with democratic ideals. While SNSs must retain some freedom to determine their own content policies, the platform's degree of dominance should contribute to stricter requirements of proportionality and subsidiarity for blocking and banning interventions.
- 72 In the context of social media content moderation, the public is private and the private is public; governments have been able to operationalise private moderation powers to further regulatory goals, and SNS companies are increasingly taking

on a quasi-judicial role in determining the limits of public discourse, with their Terms of Use and content policies coming to supplant legislative prohibitions. As these lines begin to blur, so should the distinction between public and private censorship, between horizontal and vertical free speech safeguards. SNS operators are well-positioned to act as agents of censorship in the online environment, at once highly influential and scarcely accountable. In order to address the structural threat to free speech posed by the powers of these middle men, end users' rights should be made to incorporate and reflect their fundamental rights.

- 1 The author extends his thanks to Britt van Breda, Christina Angelopoulous, Rade Obradović and Youssef Fouad for their comments and insights. Special thanks are owed to Tarlach McGonagle for his assistance throughout the research project.
- 2 A study by the Pew Research Center found that over 35% of youths find their news through SNS, and news websites receive 25% of their traffic on average from referrals by Facebook and Twitter. A. Lenhart et al., *Teens, Social Media & Technology Overview 2015*, Pew Research Center 9 April 2015. Available at: [http://www.pewinternet.org/files/2015/04/PI-TeensandTech\\_Update2015\\_0409151.pdf](http://www.pewinternet.org/files/2015/04/PI-TeensandTech_Update2015_0409151.pdf)
- 3 Z. Tufekci & C. Wilson, 'Social media and the decision to participate in political protest: Observations from Tahrir Square', *Journal of Communication* 62:2 (2012). C. Fuchs, *Social media: A Critical Introduction*, (SAGE Publications London 2012) p. 83. An interesting academic discussion, falling outside the scope of this article, has arisen as to the political efficacy of these 'born digital' protests. Authors such as Zeynep Tufekci argue that, although SNS allow for faster mobilisation and manifestation, and thus media exposure, they do not seem to support creation of unified, sustained protest movements. This is in part due to their potential for action without organisational hierarchies and a resulting inability to engage with institutional politics. Whether the increase in large-scale protests has also improved the ability to effectuate real change, is therefore subject to debate. However, regardless of whether these trends can be seen as a net gain for democracy, it is clear that SNS are now an essential instrument for political involvement and are therefore worth examining as to the risk of censorship. Z. Tufekci, 'The Medium and the Movement: Digital Tools, Social Movement Politics, and the End of the Free Rider Problem', *Policy & Internet* 6:2. E. Zuckerman 'New Media, New Civics?', *Policy & Internet* 6:2 (2014). More optimistically: C. Shirky, 'The political power of social media: technology, the public sphere and political change', *Foreign Affairs* January 2011. Available online at: <https://www.foreignaffairs.com/articles/2010-12-20/political-power-social-media> More sceptically: M Gladwell, 'Small Change: why the revolution will not be tweeted', *The New Yorker* 4 October 2010. Available online at: <http://www.newyorker.com/magazine/2010/10/04/small-change-3>
- 4 D. McGoldrick, 'The Limits of Freedom of Expression on Facebook and Social Networking Sites: A UK Perspective', *Human Rights Law Review* 13:1 (2013). O'Flaherty, 'Freedom of Expression: Article 19 of the ICCPR and the Human Rights Committee's General Comment No 34' (2012) *Human Rights Law Review* 12:4.
- 5 R. MacKinnon et al., *Fostering Freedom Online*, UNESCO Series on Internet Freedom (UNESCO Publishing 2014).
- 6 <http://newsroom.fb.com/company-info/>  
<https://about.twitter.com/company>
- 7 R. MacKinnon et al., *Fostering Freedom Online*, UNESCO Series on Internet Freedom (UNESCO Publishing 2014).
- 8 J. Zittrain, *A History of Online Gatekeeping*, *Harvard Journal of Law and Technology* 9:2 (2006), p. 255.
- 9 <https://www.facebook.com/legal/terms>
- 10 <https://www.facebook.com/legal/terms>
- 11 <https://www.facebook.com/legal/terms>
- 12 <https://twitter.com/ToS> (Emphasis added).
- 13 <https://support.twitter.com/articles/18311-the-twitter-rules>
- 14 E. Wauters et. al., 'Towards a better protection of social media users: a legal perspective on the terms of use of social networking sites', *International Journal of Law and Information Technology* 22:3 (2014).
- 15 E. Wauters et. al., 'Social Networking Sites' Terms of Use: Addressing Imbalances in the User-Provider Relationship through Ex Ante and Ex Post Mechanisms', *Journal of Intellectual Property, Information Technology and e-Commerce Law* 5:2 (2015).
- 16 E. Wauters et. al., 'Towards a better protection of social media users: a legal perspective on the terms of use of social networking sites', *International Journal of Law and Information Technology* 22:3 (2014).
- 17 On informational asymmetry as an obstacle to competition: G. Akerlof, 'The Market for Lemons: Quality Uncertainty and the Market Mechanism', *The Quarterly Journal of Economics*, 84:3 (1970).
- 18 Chiu, Aaron, T., 'Note. Irrationally bound: Terms of Use licenses and the breakdown of consumer rationality in the market for social network sites' *Southern California Interdisciplinary Law Journal* 21 (2011).
- 19 Ibid.
- 20 E. Wauters et. al., 'Social Networking Sites' Terms of Use: Addressing Imbalances in the User-Provider Relationship through Ex Ante and Ex Post Mechanisms', *Journal of Intellectual Property, Information Technology and e-Commerce Law* 5:2 (2015).
- 21 D. Harris et al., *Law of the European Convention on Human Rights*, (Oxford University Press 2014), p. 21-24.
- 22 Ibid.
- 23 *Convention on Human Rights*, Article 10(2).
- 24 *Sunday Times v UK* (No 1), app. no. 6538/74, ECHR 26 April 1979.; *Gaweda v Poland*, app. no. 26229/95 ECHR 14 March 2002. D. Harris et al., *Law of the European Convention on Human Rights*, (Oxford University Press 2014, Third Edition), p. 21-24.
- 25 D. Harris et al., *Law of the European Convention on Human Rights*, (Oxford University Press 2014, Third Edition), p. 21 - 24 .  
*Sunday Times v UK* (No 1), app. no. 6538/74, ECHR 26 April 1979. *Handyside v UK*, app. no. 5493/72, ECHR 7 December 1976.
- 26 D. Harris et al., *Law of the European Convention on Human Rights*, (Oxford University Press 2014, Third Edition), p. 21 - 24 .  
*Soering v. UK*, app. no. 14038/88, ECHR 7 July 1989.
- 27 *Delfi AS v Estonia*, app. no. 64569/09, ECHR 16 June 2015.
- 28 D. Harris et al., *Law of the European Convention on Human Rights*, (Oxford University Press 2014, Third Edition).
- 29 This increased level of protection extends to politicians and journalists, as well as activists and NGOs. See, respectively: *Feret v Belgium*, app. no. 15615/07 ECHR 16 July 2009. *Goodwin v UK*, app no. 17488/90 ECHR 27 March 1996. *Steel & Morris v. UK*, app. no. 68416/01 ECHR 05 February 2005. *TASZ v Hungary*, app. no. 37374/05 ECHR 14 April 2009.

- 30 J. Rowbottom, 'To Rant, Vent and Converse: protecting low level digital speech', *The Cambridge Law Journal* 71:02, p. 357. This article also argues for the increased protection of such 'low level' speech against disproportionate civil and criminal sanctions. Rowbottom argues that laws which were originally aimed at regulating professional mass media can place overly strict requirements on the mostly informal, albeit publicly visible, communications of SNS users.
- 31 *Müller and Others v. Switzerland*, app. no. 10737/84, ECHR 24 May 1988, para. 33.
- 32 *Handyside v UK*, app. no. 5493/72, ECHR 7 December 1976.
- 33 *Otto-Preminger-Institut v. Austria*, app. no. 13470/87, 20 September 1994, para 49.
- 34 *Garaudy v France*, app. no. 65831/01, ECHR 24 June 2003.
- 35 D. Harris et al., *Law of the European Convention on Human Rights*, (Oxford University Press 2014, Third Edition).
- 36 *Airey v. Ireland*, app. no. 6289/73, ECHR 9 October 1979, para 24.
- 37 *Airey v Ireland*, App. no. 6289/73, ECHR 9 October 1979.
- 38 *Özgür Gündem v Turkey*, app no. 23144/93, 16 March 2000. *Dink v Turkey*, app. no. 2668/07, ECHR 14 September 2010.
- 39 *Ibid.*
- 40 *Fuentes Bobo v. Spain*, app. no. [39293/98](#), ECHR 29 February 2000. In contrast, in *Palomo Sanchez*, the dismissal of labourers due to an insulting satire of their employer in a union magazine did not trigger a positive obligation for the State to intervene. Rather, the Court concluded that dismissal was 'not a manifestly disproportionate or excessive sanction capable of requiring the State to afford redress by annulling it or by replacing it with a more lenient measure'. *Palomo Sánchez And Others v. Spain*, app. nos. 28955/06, 28957/06, 28959/06 and 28964/06, ECHR 12 September 2011.
- 41 *Fuentes Bobo v. Spain*, app. no. [39293/98](#), ECHR 29 February 2000.
- 42 D. Tambini et al., 'The Privatisation of Censorship: self regulation and freedom of expression'. In: D. Tambini et. al. (ed.), *Codifying Cyberspace: communications self-regulation in the age of internet convergence* (Routledge / UCL Press 2008), pp. 269-289. This distinction between positive and negative rights was originally made by Isaiah Berlin: I. Berlin, 'Two Concepts of Liberty', in I. Berlin, *Four Essays on Liberty*, (Clarendon Press 1969).
- 43 *Appleby And Others v. The United Kingdom*, app. no. 44306/98, ECHR 6 May 2003.
- 44 D. Mac Síthigh, 'From freedom of speech to the right to communicate' in M. Price et al. (eds.), *Routledge Handbook of Media Law*, (Routledge 2013), pp. 175-191.
- 45 *Verein Gegen Tierfabriken v. Switzerland*, app. no. 24699/94, ECHR 28 June 2001, para. 63
- 46 *Verein Gegen Tierfabriken v. Switzerland*, app. no. 24699/94, ECHR 28 June 2001, para. 63
- 47 *Informationsverein Lentia*, Application no. 13914/88, 24 November 1993.
- 48 The presence of viable alternatives was also referenced in the *Appleby* case – in the context of public manifestations, the Court found that public spaces such as squares, formed a viable alternative to protests in a privately-owned shopping mall. T. McGonagle, 'The Council of Europe against online hate speech: Conundrums and challenges', Council of Europe Expert Paper MCM (2013)005.
- 49 D. Harris et al., *Law of the European Convention on Human Rights*, (Oxford University Press 2014), p. 21-24.
- 50 *Animal Defenders International v. UK*, app no. 48876/08, ECHR 22 April 2013.
- 51 *Ibid.*
- 52 T. Lewis, 'Reasserting the Primacy of Broadcast Political Speech after *Animal Defenders International*?—Rogaland Pensioners Party v Norway', *Journal of Media Law*, 1:1, (2009), pp. 37-48. T. Lewis, '*Animal Defenders International v United Kingdom*: Sensible Dialogue or a Bad Case of Strasbourg Jitters?' *The Modern Law Review* 77: 3, (2014) pp. 460-474.
- 53 The relevance of functional differences between media has been recognised by the ECHR in cases such as *Khurshid Mustafa*, which concerned tenants' access to satellite television broadcasts and their clash with landlords' property rights. *Khurshid Mustafa and Tarzibachi v. Sweden*, app no. 23883/06, ECHR 16 December 2008. See: T. McGonagle, 'The Council of Europe's standards on access to the media for minorities: A tale of near misses and staggered successes', in: M. Amos et al. (eds.), *Freedom of Expression and the Media* (Nijhoff 2012), pp. 111-140.
- 54 On the unique affordances of SNS, see, generally: C. Fuchs, *Social Media: A Critical Introduction* (SAGE Publications London 2012). D. Boyd & T. Ellison, 'Social network sites: Definition, history, and scholarship', *Journal of Computer Mediated Communication*, vol. 13, no. 1 (2007). D. Boyd, (2010). 'Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications.' In Z. Papacharissi (ed.), *Networked Self: Identity, Community, and Culture on Social Network Sites* (Routledge 2010), pp. 39-58.
- 55 Chiu, Aaron, T., 'Note. Irrationally bound: Terms of Use licenses and the breakdown of consumer rationality in the market for social network sites' *Southern California Interdisciplinary Law Journal* 21 (2011).
- 56 *Ibid.* Chiu describes consumers' difficulty in changing platforms as the 'stickiness' of SNS.
- 57 *Tyrer v. United Kingdom*, app. no. 5856/72, ECHR 25 April 1978. G. Letsas, 'Strasbourg's Interpretive Ethic: Lessons for the International Lawyer', *European Journal of International Law* 21:3 (2010), pp. 509-541.
- 58 *Verein Gegen Tierfabriken v. Switzerland*, app. no. 24699/94, ECHR 28 June 2001, para. 63.
- 59 D. Tambini et al., 'The Privatisation of Censorship: self regulation and freedom of expression'. In: D. Tambini et. al. (ed.), *Codifying Cyberspace: communications self-regulation in the age of internet convergence* (Routledge 2008).
- 60 Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market [E-Commerce Directive].
- 61 Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market, Arts. 12-15.
- 62 E. Dommering, *De Zaak Scarlet/SABAM: Naar een horizontale integratie van het auteursrecht* AMI 2011/2.
- 63 See also: Case C-324/09. *L'Oréal SA and Others v eBay International AG and Others*, (2011), 2011 ECR I-06011. Case C-236/08. *Google France SARL and others v. LVMH* (2010), ECR I-02417.
- 64 Case C-360/10, *SABAM v. Netlog NV*, (2011) ECLI:EU:C:2012:85 Para. 27: 'In that regard, first, it is not in dispute that the owner of an online social networking platform - such as Netlog - stores information provided by the users of that platform, relating to their profile, on its servers, and that it is thus a hosting service provider within the meaning of Article 14 of Directive 2000/31.'
- 65 Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market [E-Commerce Directive], article 14.
- 66 *Ibid.*

- 67 E. Dommering, *De Zaak Scarlet/SABAM: Naar een horizontale integratie van het auteursrecht* AMI 2011/2. J van Hoboken, "Legal Space for Innovative Ordering: On the Need to Update Selection Intermediary Liability in the EU" (2009) *International Journal of Communications Law & Policy* 13:1.
- 68 C. Omer, 'Intermediary Liability for Harmful Speech: Lessons from Abroad', *Harvard Journal of Law and Technology* 28:1 (2014). R. MacKinnon et al., *Fostering Freedom Online*, UNESCO Series on Internet Freedom (UNESCO Publishing 2014). An alternative perspective is found in the ECHR's recent decision in *Delfi AS v. Estonia*, where the liability of a news website for defamatory end user comments was not found to breach the freedom of expression, despite expeditious removal efforts by the intermediary. While the Court acknowledged the contribution of safe harbours to online expression, article 10 ECHR evidently does not require an absolute level of protection. *Delfi AS v Estonia*, app. no. 64569/09, ECHR 16 June 2015. See: M. Husovec, 'ECTHR Rules on Liability of ISPs as a Restriction of Freedom of Speech', *Journal of Intellectual Property Law & Practice*, 2014, Vol. 9, No. 2. The free speech concerns inherent in the viability of online intermediary services is recognised in recital 9 of the e-Commerce Directive: Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market [E-Commerce Directive].
- 69 Directive 2004/48/EC of the European Parliament and of the Council of 29 April 2004 on the enforcement of intellectual property rights, Article 11. Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society, Article 8(3).
- 70 "The limitations of the liability of intermediary service providers established in this Directive do not affect the possibility of injunctions of different kinds; such injunctions can in particular consist of orders by courts or administrative authorities requiring the termination or prevention of any infringement, including the removal of illegal information or the disabling of access to it." Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market, Recital 45.
- 71 Case C-360/10, *SABAM v. Netlog NV*, (2011) ECLI:EU:C:2012:85.
- 72 S. Kulk & F. Borgesius, 'Filtering for Copyright Enforcement in Europe after the Sabam Cases', *European Intellectual Property Review* 34:11 (2012).
- 73 Directive 2004/48/EC of the European Parliament and of the Council of 29 April 2004 on the enforcement of intellectual property rights, Article 3(2). *L'Oréal SA and Others v eBay International AG and Others*, (2011), 2011 I-06011, para. 135-136.
- 74 *L'Oréal SA and Others v eBay International AG and Others*, (2011), 2011 I-06011, para. 139. Directive 2004/48/EC of the European Parliament and of the Council of 29 April 2004 on the enforcement of intellectual property rights, Article 3(2).
- 75 Case C-314/12. *UPC Telekabel Wien GmbH v Constantin Film Verleih GmbH* (2014), not yet published.
- 76 Case C-275/06, *Promusicae v. Telefónica de España SAU* (2008), ECR I-00271. Case C-70/10, *Scarlet Extended SA v. SABAM* (2011), ECR I-11959. Case C-314/12. *UPC Telekabel Wien GmbH v Constantin Film Verleih GmbH* (2014), not yet published. Case C-360/10, *Netlog NV v SABAM* (2012), ECR I-11354. See also: C. Angelopoulos, *Are blocking injunctions against ISPs allowed in Europe? Copyright enforcement in the post-Telekabel EU legal landscape*, *Journal of Intellectual Property Law & Practice* 9:10 (2014).
- 77 *Ibid.*
- 78 *Ibid.* See M. Husovec, 'CJEU Allowed Website Blocking Injunctions with Some Reservations' (2014) *Journal of Intellectual Property Law & Practice* 9:8.
- 79 *Ibid.*
- 80 C. Angelopoulos, *Are blocking injunctions against ISPs allowed in Europe? Copyright enforcement in the post-Telekabel EU legal landscape*, *Journal of Intellectual Property Law & Practice* 9:10.
- 81 *Ibid.*
- 82 Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market, Article 14.
- 83 User-generated content is generally accompanied by buttons allowing users to 'flag' or 'report' content to SNS moderating teams, often with drop down menus allowing the flagger to specify his objection. Depending on the amount of information provided, these may or may not create 'actual knowledge'. The presence of these systems does not preclude notification by other means. E. Wauters et al., 'Towards a better protection of social media users: a legal perspective on the terms of use of social networking sites', *International Journal of Law and Information Technology* 22:1 (2014). R. MacKinnon et al., *Fostering Freedom Online*, UNESCO Series on Internet Freedom (UNESCO Publishing 2014).
- 84 *L'Oréal SA and Others v eBay International AG and Others*, (2011), 2011 I-06011, para. 122.
- 85 A majority of Member States have not implemented formal notice procedures for hosting intermediaries, and leave courts to determine 'actual knowledge' according to national legal standards on knowledge. T. Verbiest et al., *Study on the liability of Internet Intermediaries*, EC Markt/2006/09/E (2007) p.15, available at: [http://ec.europa.eu/internal\\_market/e-commerce/docs/study/liability/final\\_report\\_en.pdf](http://ec.europa.eu/internal_market/e-commerce/docs/study/liability/final_report_en.pdf)
- 86 A majority of Member States have not implemented formal notice procedures for hosting intermediaries, and leave courts to determine 'actual knowledge' according to national legal standards on knowledge. T. Verbiest et al., *Study on the liability of Internet Intermediaries*, EC Markt/2006/09/E (2007) p.15, available at: [http://ec.europa.eu/internal\\_market/e-commerce/docs/study/liability/final\\_report\\_en.pdf](http://ec.europa.eu/internal_market/e-commerce/docs/study/liability/final_report_en.pdf).
- 87 On legal realism and the public / private distinction in the context of free speech, see: J. Balkin, 'Some Realism About Pluralism: Legal Realist Approaches to the First Amendment', *Duke Law Journal* 1990:3, pp. 375 (1990). D. Kennedy, 'The Stages of the Decline of the Public/Private Distinction', *University of Pennsylvania Law Review* 130:6, pp. 1349-1357 (1982). M. Radin & R. Polk Wagner, 'The Myth of Private Ordering: Rediscovering Legal Realism in Cyberspace', *Chicago - Kent Law Review* 73 (1998).
- 88 The doctrine of positive obligations allows the ECHR to weigh the enforcement of private rights (property, intellectual property, contracts, etc.) against free speech considerations. However, it should be noted that, to the extent that the distinction between negative and positive obligations has led the ECHR to refuse free speech protection under the latter, such as in deference to property rights in *Appleby*, this approach can also be criticised from a legal realist perspective: M. Sanderson, 'Free Speech in Public Spaces: The Privatisation

- of Human Rights in *Appleby v. UK*, 15 Kings College Law Journal 159 (2010).
- 89 W. Hins, 'The Freedom to Conduct Business and the Right to Receive Information for Free: Sky Österreich', 51 Common Market Law Review 51:2 (2014).
  - 90 K. Rawlinson, 'Turkey blocks use of Twitter after prime minister attacks social media site', The Guardian 21 March 2014. Available online at: <http://www.theguardian.com/world/2014/mar/21/turkey-blocks-twitter-prime-minister>
  - 91 R. Akkoc, 'Turkey blocks access to social media and YouTube over hostage photo', The Telegraph 6 April 2015. Available online at: <http://www.telegraph.co.uk/news/worldnews/europe/turkey/11518004/Turkey-blocks-access-to-Facebook-Twitter-and-YouTube.html>
  - 92 Ibid.
  - 93 Gawker's Adrian Chen reports that a leaked internal memo from 2012 reveals that Facebook as a rule blocked depictions of the maps of Kurdistan for Turkish users, as well as 'all attacks on Ataturk': A. Chen 'Inside Facebook's Outsourced Anti-Porn and Gore Brigade, Where 'Camel Toes' are More Offensive Than 'Crushed Heads'', Gawker 16 February 2012. Available online at: <http://gawker.com/5885714/inside-facebooks-outsourced-anti-porn-and-gore-brigade-where-camel-toes-are-more-offensive-than-crushed-heads>
  - 94 M. Husovec, 'Injunctions against innocent third parties: case of website blocking', Journal of Intellectual Property, Information Technology and e-Commerce Law 4:2 (2012). J. Garside, 'Ministers will order ISPs to block terrorist and extremist websites', The Guardian 27 November 2013.
  - 95 That is not to say, of course, that the EU regime for ISP-level blocking is wholly impervious to abuse or misuse. In the context of copyright enforcement, scholars have identified rising problems in the practice of website blocking, especially tensions with the right to a fair trial, legality and costs of injunctions. See: M. Husovec, Injunctions against innocent third parties: case of website blocking, Journal of Intellectual Property, Information Technology and e-Commerce Law 4:2 (2012).
  - 96 Y. Benkler, 'A Free, Irresponsible Press: Wikileaks and the battle over the soul of the networked fourth estate', Harvard Civil Rights-Civil Liberties Law Review 46. M. Birnhack & N. Elkin-Koren, 'The Invisible Handshake: The Reemergence of the State in the Digital Environment', Virginia Journal of Law and Technology 8:6 (2003).
  - 97 "Examples of the symbolic manoeuvring and dialogue between governments and media are numerous; state involvement in ISPs ranges from exhortation in ministerial speeches, to direct involvement in setting up task forces. In Italy, the 'voluntary' ISP code was drafted by the Ministry of Communications." D. Tambini et al., 'The Privatisation of Censorship: self regulation and freedom of expression'. In: D. Tambini et al. (eds.), Codifying Cyberspace: communications self-regulation in the age of internet convergence (Routledge 2008), p 416.
  - 98 Y. Benkler, 'A Free, Irresponsible Press: Wikileaks and the battle over the soul of the networked fourth estate', Harvard Civil Rights-Civil Liberties Law Review 46. M. Birnhack & N. Elkin-Koren, 'The Invisible Handshake: The Reemergence of the State in the Digital Environment', Virginia Journal of Law and Technology 8:6 (2003).
  - 99 P. Dominiczak et al., 'Facebook "could have prevented Lee Rigby murder"', The Telegraph 26 November 2014. Available online at: <http://www.telegraph.co.uk/news/uknews/terrorism-in-the-uk/11253518/Facebook-could-have-prevented-Lee-Rigby-murder.html>
  - 100 P. Sawers, 'Facebook and Twitter could be hate-speech 'accomplices' if new French law passes', Venturebeat 27 January 2015. Available online at: <http://venturebeat.com/2015/01/27/facebook-and-twitter-could-be-hate-speech-accomplices-if-new-french-law-passes/>
  - J. York, 'Unpacking France's Chilling Proposal to Hold Companies Accountable for Speech', The Electronic Frontier Foundation 6 February 2015. Available online at: <https://www.eff.org/deeplinks/2015/02/unpacking-frances-chilling-proposal-hold-companies-accountable-speech>
  - 101 EU ministers of interior, Joint Statement of 11 January 2015. Available online at: [http://ec.europa.eu/dgs/home-affairs/what-is-new/news/news/docs/20150111\\_joint\\_statement\\_of\\_ministers\\_for\\_interior\\_en.pdf](http://ec.europa.eu/dgs/home-affairs/what-is-new/news/news/docs/20150111_joint_statement_of_ministers_for_interior_en.pdf)
  - 102 L. Kelion, 'Facebook revamps its takedown guidelines', BBC 16 March 2015. Available online at: <http://www.bbc.com/news/technology-31890521>
  - 103 H. Parkinson, 'James Foley: How social media is fighting back against Isis propaganda' The Guardian 20 August 2014. Available online at: <http://www.theguardian.com/technology/2014/aug/20/james-foley-how-social-media-is-fighting-back-against-isis-propaganda>
  - 104 S. Dredge, 'Twitter has tripled the size of its team handling abuse reports', The Guardian 27 February 2015. Available online at: <http://www.theguardian.com/technology/2015/feb/27/twitter-tripled-team-abuse-reports>
  - 105 A. Hern, 'Twitter announces crackdown on abuse with new filter and tighter rules' The Guardian 21 April 2015. <http://www.theguardian.com/technology/2015/apr/21/twitter-filter-notifications-for-all-accounts-abuse>
  - 106 V. Gadde, 'Twitter executive: Here's how we're trying to stop abuse while preserving free speech', The Washington Post 16 April 2015. <https://www.washingtonpost.com/posteverything/wp/2015/04/16/twitter-executive-heres-how-were-trying-to-stop-abuse-while-preserving-free-speech/>
  - 107 <http://www.constitutionnet.org/news/striking-balance-hate-speech-glorification-terrorism-freedom-speech-post-charlie-hebdo>
  - 108 Y. Benkler, A Free, Irresponsible Press: Wikileaks and the battle over the soul of the networked fourth estate' Harvard Civil Rights-Civil Liberties Law Review 46.
  - 109 Y. Benkler, A Free, Irresponsible Press: Wikileaks and the battle over the soul of the networked fourth estate' Harvard Civil Rights-Civil Liberties Law Review 46.
  - 110 Ibid.
  - 111 B. Quinn, 'YouTube staff too swamped to filter out all terror-related content', The Guardian 28 January 2015. Available online at: <http://www.theguardian.com/technology/2015/jan/28/youtube-too-swamped-to-filter-terror-content>
  - 112 DS 1035/15, EU Counter-terrorism Coordinator input for the preparation of the informal meeting of Justice and Home Affairs Ministers in Riga on 29 January 2015, Council of the European Union 17 January 2015. Available online at: <http://www.statewatch.org/news/2015/jan/eu-council-ct-ds-1035-15.pdf> (emphasis mine)
  - 113 DS 1035/15, EU Counter-terrorism Coordinator input for the preparation of the informal meeting of Justice and Home Affairs Ministers in Riga on 29 January 2015, Council of the European Union 17 January 2015. Available online at: <http://www.statewatch.org/news/2015/jan/eu-council-ct-ds-1035-15.pdf>
  - 114 S. Zeidan, 'Agreeing to Disagree: Cultural Relativism and the Difficulty of Defining Terrorism in a Post-9/11 World', 29 Hastings International & Comparative Law Review 215 (2005). In addition, a report by Quilliam Foundation found that these repressive online measures are often ineffective at countering real terrorism: 'Negative measures, including Government-backed censorship and filtering initiatives, are ineffective

- in tackling online extremism, tackling the symptoms rather than the causes of radicalisation,' [...] 'Motivated extremists and terrorist affiliates can evade such measures easily through the dark net and virtual private networks. [...] 'Blocked materials consistently reappear online and there is no effective way for ISPs (Internet service providers) or social media companies to filter extremist content.' G. Hussein & E. Saltmann, *Jihad Trending: A Comprehensive Analysis of Online Extremism and How to Counter it*, Quilliam 2014. Available online at: <http://www.quilliamfoundation.org/publications/>
- 115 E. Wauters et. al., 'Towards a better protection of social media users: a legal perspective on the terms of use of social networking sites', *International Journal of Law and Information Technology* 22:1 (2014).
- 116 <https://transparency.twitter.com/>
- 117 [govtrequests.facebook.com  
linkedin.com/legal/transparency](http://govtrequests.facebook.com/linkedin.com/legal/transparency)
- 118 <https://transparency.twitter.com/>
- 119 <https://transparency.twitter.com/>
- 120 For a more detailed analysis of the Twitter and Facebook transparency reports, see: R. MacKinnon et al., *Fostering Freedom Online*, UNESCO Series on Internet Freedom (UNESCO Publishing 2014).
- 121 <https://transparency.twitter.com/govtrequests.facebook.com>
- 122 UK Minister of State James Brokenshire Speaking to the House of Commons: „I can update the House that since 2010 the counter terrorism internet referral unit has taken down more than 29,000 pieces of illegal terrorist material from the internet. I underline the fact that any online activity by the three groups under consideration, including Facebook pages and Twitter accounts, has been referred to CTIRU. If it is assessed as illegal—there is a legal test that has to be met—CTIRU will flag it directly to Facebook and Twitter for removal.” Hansard, HC Deb, 2 April 2014, c957, 2014-04-02. Available online at: [http://www.publications.parliament.uk/pa/cm201314/cmhansrd/cm140402/debtext/140402-0003.htm#140402-0003.htm\\_spnew12](http://www.publications.parliament.uk/pa/cm201314/cmhansrd/cm140402/debtext/140402-0003.htm#140402-0003.htm_spnew12) Similar numbers were confirmed by UK Home Secretary Theresa May in 2014: “Since the start of this government, the Counter-Terrorism Internet Referral Unit has secured the removal of 65,000 items from the internet that encouraged or glorified acts of terrorism. More than 46,000 of these have been removed since December last year.” T. May, Home Secretary’s Speech on Counter-Terrorism, Gov.UK 24 November 2014. Available online at: <https://www.gov.uk/government/speeches/home-secretary-theresa-may-on-counter-terrorism>
- 123 This conclusion was also drawn by MacKinnon et al. in a report for UNESCO: “Industry sources confirmed to this report’s researchers that authorities of some governments sometimes also seek to have content restricted via the companies’ own self-regulatory mechanisms, rather than make formal requests through official channels following legally specified process. As of this writing, neither company has included in its transparency reports any data about the extent to which governments in various countries use the companies’ own self-regulatory mechanisms designed for individuals to report content that violates the Terms of Use.” R. MacKinnon et al., *Fostering Freedom Online*, UNESCO Series on Internet Freedom (UNESCO Publishing 2014), p. 143. The author has contacted the CTIRU, as well as Facebook and Twitter, asking for further clarification as to the figures and definitions involved. He did not receive any response.
- 124 J. Preston. ‘Facebook Deactivates Protest Pages in Britain’, *The New York Times Blog* 29 April 2011. <http://mediadecoder.blogs.nytimes.com/2011/04/29/facebook-deactivates-protest-pages-in-britain/?partner=rss&emc=rss>
- S. Malik., ‘Activists claim purge of Facebook pages’, *The Guardian* 29 April 2011. <http://www.theguardian.com/uk/2011/apr/29/facebook-activist-pages-purged>
- 125 Ibid.
- 126 Ibid.
- 127 <https://globalnetworkinitiative.org/participants/index.php>  
<https://globalnetworkinitiative.org/corecommitments/index.php>
- 128 <https://globalnetworkinitiative.org/implementationguidelines/index.php>
- 129 C. Baudelaire, *Le Joueur généreux* (1864) Available online at: <http://baudelaire.litteratura.com/pop/texte/167-le-joueur-generoux.html>
- 130 <https://globalnetworkinitiative.org/implementationguidelines/index.php>
- 131 <https://globalnetworkinitiative.org/implementationguidelines/index.php>
- 132 Institute for Information Law (IViR), ‘Study of Fundamental Rights Limitations for Online Enforcement Through Self-Regulation’, (forthcoming, July 2015, pre-publication version on file with author).
- 133 <https://www.manilaprinciples.org/>
- 134 <https://www.manilaprinciples.org/principles> Manila Principles, Article 6(b).
- 135 H. Parkinson, ‘James Foley: How social media is fighting back against Isis propaganda’ *The Guardian* 20 August 2014. Available online at: <http://www.theguardian.com/technology/2014/aug/20/james-foley-how-social-media-is-fighting-back-against-isis-propaganda>
- 136 H. Parkinson, ‘James Foley: How social media is fighting back against Isis propaganda’ *The Guardian* 20 August 2014. Available online at: <http://www.theguardian.com/technology/2014/aug/20/james-foley-how-social-media-is-fighting-back-against-isis-propaganda>
- 137 Informationsverein Lentia, Application no. 13914/88, 24 November 1993. *Verein Gegen Tierfabriken v. Switzerland*, app. no. 24699/94, ECHR 28 June 2001. *Garaudy v France*, app. no. 65831/01, ECHR 24 June 2003., *Feret v Belgium*, app. no. 15615/07 ECHR 16 July 2009, *Özgür Gündem v Turkey*, app no. 23144/93, 16 March 2000. See also: T. McGonagle, ‘The State and beyond: activating (non-)media voices’, in: H. Sousa et al. (eds), *Media Policy and Regulation: Activating Voices, Illuminating Silences*, Communications and Society Research Centre, University of Minho, (2014), p. 187-198.
- 138 A. Withnall, ‘Facebook suspends Italian woman’s account after she posts image of two women kissing in support of LGBT rights’ *The Independent*, 19 May, 2014. Available at <http://www.independent.co.uk/news/world/europe/facebook-suspends-italian-womans-account-after-she-posts-image-of-two-women-kissing-in-support-of-lgbt-rights-9398750.html>
- J. Hudson, ‘The Controversy Over Facebook’s Gay Kissing Ban Isn’t Over’, *The Wire* 22 April 2011. Available online at: <http://www.thewire.com/technology/2011/04/controversy-over-facebooks-gay-kissing-ban-isnt-over/36954/>
- 139 M. Sweney, ‘Mums furious as Facebook removes breastfeeding photo’, *The Guardian* 30 December 2008. Available online at: <http://www.theguardian.com/media/2008/dec/30/facebook-breastfeeding-ban>
- 140 J. Bartlett, ‘Facebook’s Nudity Policy Laid Bare’, *The Telegraph* 10 March 2015. Available online at: <http://www.telegraph.co.uk/technology/facebook/11461004/Facebooks-nudity-policy-laid-bare.html>

- 141 Twitter reports to have received 17,809 copyright-based takedown notices in the second half of 2014. 66 of these notices led to content removal. <https://transparency.twitter.com/copyright-notices/2014/jul-dec>
- 142 E. Bonadio, 'File Sharing, Copyright and Freedom of Speech', 33 European Intellectual Property Review 10 (2011), p. 12-13.
- 143 Verbiest et al., Study on the liability of Internet Intermediaries, EC Markt/2006/09/E (2007), p. 14-15. Available online at: [http://ec.europa.eu/internal\\_market/e-commerce/docs/study/liability/final\\_report\\_en.pdf](http://ec.europa.eu/internal_market/e-commerce/docs/study/liability/final_report_en.pdf)
- 144 T. Verbiest et al., Study on the liability of Internet Intermediaries, EC Markt/2006/09/E (2007), p. 14-15. Available online at: [http://ec.europa.eu/internal\\_market/e-commerce/docs/study/liability/final\\_report\\_en.pdf](http://ec.europa.eu/internal_market/e-commerce/docs/study/liability/final_report_en.pdf)
- 145 Institute for Information Law (IViR), 'Study of Fundamental Rights Limitations for Online Enforcement Through Self-Regulation', report for Open Society Foundation (forthcoming).
- 146 <https://globalnetworkinitiative.org/implementationguidelines/index.php>
- 147 S. Malik, 'Activists claim purge of Facebook pages', The Guardian 29 April 2011. <http://www.theguardian.com/uk/2011/apr/29/facebook-activist-pages-purged>
- 148 The decision to withhold content on a regional basis, rather than delete it entirely, was introduced by Twitter in 2011. It was first applied for Germany's prohibitions on Holocaust denial, and has since become their default response to valid state orders. These tailored measures allow for compliance with national rules, without directly compromising communications globally. However, these methods have also been criticised for creating a 'slippery slope' by lowering the barrier for interference by intermediaries. It should also be considered that the chilling effect caused by local restrictions can be detrimental to discourse on a global scale. With this in mind, the possibility for region-specific moderation is generally preferable to outright deletion, but must be applied with caution and restraint. K. Rawlinson, 'Twitter uses new "country-withheld content" rule to block neo-Nazi group tweets in Germany', The Independent 18 October 2012. Available online at: <http://www.independent.co.uk/life-style/gadgets-and-tech/news/twitter-uses-new-countrywithheld-content-rule-to-block-neonazi-group-tweets-in-germany-8216260.html>
- 149 Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market.
- 150 E. Wauters et. al., 'Social Networking Sites' Terms of Use: Addressing Imbalances in the User-Provider Relationship through Ex Ante and Ex Post Mechanisms', Journal of Intellectual Property, Information Technology and e-Commerce Law 5:2 (2014).