Chapter 31

Accountability and Responsibility of Online Intermediaries

Giancarlo Frosio and Martin Husovec

in

Giancarlo Frosio (ed), *The Oxford Handbook of Online Intermediary Liability* (Oxford University Press, forthcoming 2019)

Legal theory is increasingly shifting the discourse from liability to enhanced 'responsibilities' for intermediaries under the assumption that OSPs' role is unprecedented for their capacity to influence the informational environment and users' interactions within it.[1] Hence, academia, policy-makers and society increasingly ascribe a public role to online intermediaries.[2] According to Shapiro, 'in democratic societies, those who control the access to information have a responsibility to support the public interest. […] these gatekeepers must assume an obligation as trustees of the greater good.'[3] The discourse focuses more and more on moral responsibilities of OSPs in contemporary societies and aims at building ethical frameworks for the understanding of OSPs responsibilities, eg corporate social responsibilities or human rights. Responsible behaviour beyond the law finds justification in intermediaries' corporate social responsibilities and their role in implementing and fostering human rights.[4] In the introduction to *The Responsibilities of Online Service Providers*, Mariarosaria Taddeo and Luciano Floridi noted that—given their prominent role in the present society—online intermediaries are increasingly expected to act according to current social and cultural values, which rises 'questions as to what kind of responsibilities OSPs should bear, and which ethical principles should guide their actions'.[5] Policy approaches might be returning to implement moral theories of intermediary liability, rather than utilitarian or

[1] See eg European Commission, 'Tackling Illegal Content Online. Towards an enhanced responsibility of online platforms' COM(2017) 555 final, s 6 (noting 'the constantly rising influence of online platforms in society, which flows from their role as gatekeepers to content and information, increases their responsibilities towards their users and society at large').

[2] Pressure comes increasingly from users as well as recent lawsuits against platforms supposedly liable of fomenting extremism, radicalism—and related terroristic actions—might prove. See 'Orlando nightclub victims' families sue Twitter, Google, Facebook' (*CNBC*, December 21, 2016) <http://www.cnbc.com/2016/12/21/orlando-nightclub-victims-families-sue-twitter-google-facebook.html>.

[3] Andrew Shapiro, *The Control Revolution: How the Internet is Putting Individuals in Charge and Changing the World We Know* (Public Affairs 2000) 225.

[4] See Emily Laidlaw, *Regulating Speech in Cyberspace: Gatekeepers, Human Rights and Corporate Responsibility* (CUP 2015); Dennis Broeders and others, 'Does Great Power Come with Great Responsibility? The Need to Talk About Corporate Responsibility' in Mariarosaria Taddeo and Luciano Floridi, The Responsibilities of Online Service Providers (Springer 2017) 315-323.

[5] Mariarosaria Taddeo and Luciano Floridi, 'New Civic Responsibilities for Online Service Providers' in Mariarosaria Taddeo and Luciano Floridi, *The Responsibilities of Online Service Providers* (Springer 2017) 1.

welfare theories. In this case, justification for policy intervention would be based on responsibility for the actions of users as opposed to efficiency or balance innovation vs harm.[6]

In Europe, the *Communication on Online Platforms and the Digital Single Market* puts forward the idea that 'the responsibility of online platforms is a key and cross-cutting issue.'[7] Again, in another Communication, the Commission made this goal even clearer by openly pursuing 'enhanced responsibility of online platforms' on a voluntary basis.[8] Online platforms would be invested by a duty to 'ensure a safe online environment' against illegal activities.[9] As the Commission puts it, the goal is 'to engage with platforms in setting up and applying voluntary cooperation mechanisms aimed at depriving those engaging in commercial infringements of intellectual property rights (IPRs) of the revenue streams emanating from their illegal activities, in line with a 'follow the money' approach'.[10] Hosting providers—especially platforms—would be called to actively and swiftly remove illegal materials, instead of reacting to complaints. They would be called to adopt effective voluntary '*proactive* measures to detect and remove illegal content online'[11] and are encouraged to do so by using automatic detection and filtering technologies.[12] Again, 'online platforms must be encouraged to take more effective voluntary action to curtail exposure to illegal or harmful content' such as incitement to terrorism, child sexual abuse and hate speech'.[13]

Looking at the legal liability rules always tells only half of a story. Legal rules are often only basic expectations which are further developed through market transactions, business decisions, and political pressure. Therefore, the real responsibility landscape is equally determined by a mixture of voluntary agreements, self-regulation, corporate social responsibility, and ad-hoc deal-making. Accountability schemes can differ significantly, ranging from legal entitlements to request assistance in enforcement to entirely voluntary private-ordering schemes. In this chapter, we try to provide a mapping of these approaches in order to illustrate the richness and trade-offs associated with such measures. Miscellaneous policy and enforcement tools, such as monitoring and filtering, graduated response, payment blockades and follow-the-money strategies, private DNS content regulation, online search manipulation, are discussed to complement the typical legal liability view of the regulation of intermediaries.

---

[6] For further discussion on justifications for intermediary liability see Chapter 3, s 4.

[7] European Commission, 'Online Platforms and the Digital Single Market: Opportunities and Challenges for Europe' (Communication) COM(2016) 288 Final, 9.

[8] See Communication (n 1).

[9] ibid s 3.

[10] See Communication (n 7) 8.

[11] Communication (n 1) s 3.3.1 (noting that adopting such voluntary proactive measures does not lead the online platform to automatically lose the hosting liability exemption provided by the eCommerce Directive.

[12] ibid s 3.3.2.

[13] Communication (n 7) 9.

1. Tools for Increasing Responsibility

Miscellaneous forms of 'responsible' behaviours beyond the law—such as codes of conducts, three-strikes schemes, voluntary filtering, online search manipulation, follow-the-money strategies, and private DNS content regulation—reflect a globalized, ongoing move towards privatization of law enforcement online through proactive actions and algorithmic tools that spans all subject matters relevant to intermediary liability online. Their common denominator is that that they go beyond the baseline legal expectations created by the legal liability framework. Inherently, their common trajectory is towards more proactive tackling of illegal or otherwise objectionable content.

However, these policies often differ in the way in which they come about. Even the same type of enforcement arrangements, such as graduated response, can be the result of a private ordering scheme, ad-hoc governmental policy administered by agencies, or of application of the legal claims to assistance in enforcement. In this section, we provide a brief mapping of different arrangements which were developed over the years. In the next section then, we highlight how the *mechanisms* behind these arrangements have significant consequences for parameters of rule of law, due process, transparency or costs allocation.

1.1. Graduate Response

So-called 'graduated response' or 'three-strike' regulations—seeking to block out household Internet connections of repeat infringers—has emerged as an early form of 'responsible' behaviour of OSPs. In some instances, graduate response arrangements have been judicially or legislatively mandated. Often, they result from voluntary arrangements. Although little is known regarding the specifics of these agreements, rightsholders might attempt to leverage their content and would only license if the providers implemented a disconnection strategy. French Hadopi Law and other countries such as New Zealand, South Korea, Taiwan, and the United Kingdom, have mandated gradual response schemes, actually managed by administrative agencies, rather than intermediaries.[14] However, industry-led self-regulation makes up the largest part of graduated response schemes as in the case of the 'six strikes' Copyright Alert System (CAS), discontinued in January 2017.[15] CAS implemented a system of multiple alerts. After a fifth alert, ISPs were allowed to take 'mitigation measures' to prevent future infringement.[16] Mitigation measures included 'temporary reductions of Internet speeds, temporary downgrade in Internet service tier or redirection to a landing page until the

[14] See Law no 2009-669 of 12 June 2009, promoting the dissemination and protection of creative works on the Internet (a.k.a. HADOPI law) (FR); Law no 2009-1311 of 28 October 2009, on the criminal protection of literary and artistic property on the Internet, Art 7 and 10 (FR) (providing internet suspension sanctions for persons using the Internet to commit infringement and obligations for owner of internet access to secure their internet access); Copyright (Infringing File Sharing) Regulations 2011 (NZ); Copyright Act as amended on 22 January 2014, Art 90-4(2) (TW); Digital Economy Act 2010 C 24 (UK) (however, the 'obligations to limit Internet access' have not been implemented yet).

[15] David Kravets, 'RIP, 'Six Strikes' Copyright Alert System' (*ArsTechnica*, 30 January 2017)

[16] See Center for Copyright Information, Copyright Alert System (CAS) <http://web.archive.org/web/20130113051248/http://www.copyrightinformation.org/alerts>.

subscriber contacts the ISP to discuss the matter or reviews and responds to some educational information about copyright, or other measures (as specified in published policies) that the ISP may deem necessary to help resolve the matter'.[17] In Australia, an industry-negotiated graduated response Code was submitted to the Australian Communications and Media Authority (ACMA) for registration as an industry code, requiring Internet Service Providers to pass on warnings to residential fixed account holders who are alleged to have infringed copyright.[18]

In Europe, Eircom was one of the first European ISPs to implement a voluntary Graduated Response Protocol under which Eircom would issue copyright infringement notices to customer, after a settlement had been reached between record companies and Eircom.[19] The Irish Supreme Court later upheld the validity of the scheme against an Irish Data Protection Commissioner's enforcement notice requiring Eircom to cease its operation of the Protocol.[20] Recently, an agreement has been negotiated between major British ISPs and rights-holders—with the support of the UK Government—under the name of Creative Content UK. [21] This voluntary scheme would implement four educational-only notices or alerts sent by the ISPs to their subscribers based on IP addresses supplied by the rights-holders, where the IP address is alleged to have been used to transmit infringing content.[22]

1.2. Changes to Online Search Results

Online search manipulation—and so-called demotion—enforce sanitization of presumptively illicit activities online through voluntary measures and private ordering. Demotion spans multiple subject matters and online allegedly illicit activities. Starting from the most recent effort of this kind, under the aegis of the UK Intellectual Property Office, representatives from the creative industries and leading UK search engines developed a Voluntary Code of Practice dedicated to the removal of links to infringing content from the first page of search results.[23] However, Google have been demoting allegedly pirate sites for some time now. In 2012, Google altered its PageRank search algorithm taking into account the number of DMCA-compliant notices for each

[17]Ibid

[18] See Communications Alliance Ltd, C653:2015 – Copyright Notice Scheme Industry Code (April 2015) <http://www.commsalliance.com.au/__data/assets/pdf_file/0005/48551/C653-Copyright-Notice-Scheme-Industry-Code-FINAL.pdf>.

[19] See EIR, Legal Music - Frequently Asked Questions <www.eir.ie/notification/legalmusic/faqs>.

[20] See *EMI v Data Protection Commissioner* [2013] IESC 34 (IR).

[21] Creative Content UK <http://www.creativecontentuk.org>.

[22] ibid

[23] See Intellectual Property Office, 'Press Release: Search Engines and Creative Industries Sign Anti-Piracy Agreement' (20 February 2017) <https://www.gov.uk/government/news/search-engines-and-creative-industries-sign-anti-piracy-agreement>.

website.[24] Shortly thereafter, in 2014, Google started to demote autocomplete predictions returning search results containing DMCA-demoted sites.[25]

Voluntary measures have been traditionally implemented with regard to manifestly illegal content, such as child pornography.[26] In addition, Google has adopted specific self-regulatory measures for revenge porn, which Google delists from Internet searches.[27] Other major platforms followed Google's lead. After being ordered by a Dutch court to identify revenge porn publishers in the past,[28] Facebook decided to introduce photo-matching technology to stop revenge porn and proactively filter its reappearance.[29] Finally, search manipulation and demotion begun to be applied to curb extremism and radicalization. Plans have been also revealed of a pilot scheme to tweak search to make counter-radicalisation videos and links more prominent.[30]

1.3. Payment Blockades and Follow the Money

Payment blockades—notice-and-termination agreement between major right holders and online payment processors—and 'voluntary best practices agreements' have been applied widely as part of 'a long-term, evolving strategy on the part of corporate copyright and trademark owners'.[31] In its Joint Strategic Plan for Intellectual Property Enforcement, the US Government backed-up voluntary measures and fully endorsed a 'follow the money' strategy.[32] Similarly, in the Communication *Towards a Modern, More European Copyright Framework*, the European Commission endorses a similar 'follow-the-money' approach.[33] According to the Commission, 'follow-the-money' mechanisms should be based on a self-regulatory approach through the implementation of Code of Conducts, such as the Guiding Principles for a *Stakeholders' Voluntary*

---

[24] See Annemarie Bridy, 'Copyright's Digital Deputies: DMCA-plus Enforcement by Internet Intermediaries' in John Rothchild (ed), *Research Handbook on Electronic Commerce Law* (Edward Elgar Publishing 2016) 200.

[25] ibid

[26] See Anchayil Anjali, and Arun Mattamana, 'Intermediary Liability and Child Pornography A Comparative Analysis' (2010) 5 J Int'l Comm L Tech 48.

[27] See Joanna Walters, 'Google to Exclude Revenge Porn from Internet Searches? (*The Guardian*, 21 June 2015) <https://www.theguardian.com/technology/2015/jun/20/google-excludes-revenge-porn-internet-searches>

[28] See Agence France-Press, 'Facebook Ordered by Dutch Court to Identify Revenge Porn Publisher' (*The Guardian*, 26 June 2015) <https://www.theguardian.com/technology/2015/jun/26/facebook-ordered-by-dutch-court-to-identify-revenge-porn-publisher>.

[29] See Emma Grey Ellis, 'Facebook's New Plan May Curb Revenge Porn, but Won't Kill It' (*Wired*, 6 April 2017)

[30] See Ben Quinn, 'Google to Point Extremist Searches Towards Anti-radicalization Websites' (*The Guardian*, 2 February 2016) <http://buff.ly/20J3pFi>.

[31] See Annemarie Bridy, 'Internet Payment Blockades' (2015) 67 Florida L Rev 1523.

[32] See Office of the Intellectual Property Enforcement Coordinator, *Supporting Innovation, Creativity & Enterprise: Charting a Path Ahead (US Joint Strategic Plan for Intellectual Property Enforcement FY 2017-2019)* (2017).

[33] See Commission, 'Towards a Modern More European Copyright Framework' (Communication) COM (2015) 260 final 10-11.

Agreement on Online Advertising and IPR.[34] As stated by the principles, 'the purpose of the agreement is to dissuade the placement of advertising on commercial scale IP infringing websites and apps (eg on mobile, tablets, or set-top boxes), thereby minimising the funding of IP infringement through advertising revenue.'[35] Payment processors like MasterCard and Visa have been pressured to act as IP enforcers, extending the reach of IP law to websites operating from servers and physical facilities located abroad.[36] Across 2011 and 2012, American Express, Discover, MasterCard, Visa, and PayPal, PULSE, and Diners Club entered into a best practice agreement with thirty-one major right holders.[37] The voluntary agreement was implemented by the launch of the Payment Processor Initiative run by the International AntiCounterfeiting Coalition (IACC).[38]

There are a number of instances where payment intermediaries' terms of service have been used as pressure points against protected speech. *Inter alia*, payment blockades crippled Wikileaks of 95% of its revenues, when PayPal, Moneybookers, Visa and MasterCard stopped accepting public donations. No legal proceeding was ever actually initiated against Wikileaks. In *Backpage v. Dart*, Tom Dart—the Sheriff for Cook County, Illinois—sent letters to Visa and MasterCard demanding that they cease doing business with Backpage.com due to content in the 'adult services' section of the classified ads site.[39] The credit card companies both complied, cutting off services to the entire site's worldwide operations.[40] Backpage claimed that the sheriff's informal censorship pressure amounted to a prior restraint on speech.[41] The case was finally appealed to the Seventh Circuit, which reversed the previous decision and upheld a prior restraint on speech defence.[42] As it turned out, however, a single action from a governmental official—lacking any due process scrutiny—was potentially capable to put under jeopardy an online business operating worldwide.

The *Backpage* case highlights how intermediaries often face business incentives that make them more likely to yield to pressure. In the United States, a prior restraint on speech defence would not be directly actionable against intermediaries—if governmental pressures cannot be proved—as it does apply only against public parties. Under European

---

[34] Commission, 'The Follow the Money Approach to IPR Enforcement – Stakeholders' Voluntary Agreement on Online Advertising and IPR: Guiding Principles' http://ec.europa.eu/docsroom/documents/19462/attachments/1/translations/en/renditions/native.

[35] Ibid

[36] Ibid 1523.

[37] See Best Practices to Address Copyright Infringement and the Sale of Counterfeit Products on the Internet (16 May 2011).

[38] See Bridy (n 96) 1549.

[39] See Letter from Sheriff Thomas J. Dart to Mr. Charles W. Scharf, Chief Executive, Visa Inc. (29 June 2015) <http://cdn.arstechnica.net/wp-content/uploads/2015/07/backpageexhibit.pdf>; Letter from Sheriff Thomas J. Dart to Mr. Ajaypal Banga, President and Chief Executive Officer, MasterCard Inc. (29 June 2015)

[40] See Rainey Reitman, 'Caving to Government Pressure, Visa and MasterCard Shut Down Payments to Backpage.com? (*EFF*, 6 July 2015) <https://www.eff.org/deeplinks/2015/07/caving-government-pressure-visa-and-mastercard-shut-down-payments-backpagecom>.

[41] See Backpage.com v. Sheriff Thomas J. Dart, 1:15-cv-06340 (N.D. Ill. 2015) (US) (Complaint for Injunctive and Declaratory Relief and Damages)

[42] See Backpage.com v. Sheriff Thomas J. Dart, 807 F3d 229 (7th Cir. 2015) (US).

law, there might be room to claim interference with online providers' freedom to conduct a business,[43] still this is costly and uncertain to prove.

Other tools which were developed to tackle money flow are information disclosures against payment providers. The European system of information disclosures against third parties were used in litigation to unveil identity of potential infringers by invoking these measures against banking institutions. Although the courts recognized the interest in secrecy and data protection as important, their application has to be balanced with the right to effective remedy of IP right holders.[44]

1.4. Private DNS Content Regulation

Domain hopping would evade law enforcement by moving from one ccTLDs (country code Top-Level Domains) or gTLDs (Generic Top-Level Domains) registrar to another, thus driving up time and resources spent on protecting IP right. In this context, responsible behaviour would rely on stewards of the Internet's core technical functions, such as ICANN, and implicates Internet infrastructure and governance.[45] Apparently, ICANN might be increasingly directly involved with online content regulation through its contractual facilitation of a 'trusted notifier' copyright enforcement program. ICANN contractual architecture for the new gTLDs embeds support for private, DNS-based content regulation on behalf of copyright holders—and, potentially, other 'trusted' parties—imposing on registry operators and registrars an express prohibition on IP infringement and obligations including suspension of the domain name.[46]

Through this contractual framework, ICANN facilitated voluntary enforcement agreements between DNS intermediaries and rightholders, such as the DNA's Healthy Domains Initiative. The registry operators agrees, if 'the domain clearly is devoted to abusive behaviour [. . .] in its discretion [to] suspend, terminate, or place the domain on registry lock, hold, or similar status' within ten business days form the complaint.[47] In general, as Bridy explained, 'in creating that architecture, ICANN did nothing to secure any procedural protections or uniform substantive standards for domain name registrants who find themselves subject to this new form of DNS regulation'.[48]

1.5. Standardization

The European Commission also increased pressure through its soft-law by creating a set of expectations that should be followed by the intermediaries to avoid further

---

[43] *See* Charter of Fundamental Rights of the European Union [2000] 2000/C OJ 364/1, Art 16.

[44] See C-580/13 - Coty Germany.

[45] See, for further in-depth review of this enforcement strategy, *Chapter 32*.

[46] See ICANN-Registry Agreement (2013) s 2.17 Specification 11; ICANN Registrar Accreditation Agreement (2013) s 3.18.

[47] See Donuts.Domains, Characteristics of a Trusted Notifier Program <http://www.donuts.domains/images/pdfs/Trusted-Notifier-Summary.pdf>.

[48] Annamarie Bridy, 'Notice and Takedown in the Domain Name System: ICANN's Ambivalent Drift into Online Content Regulation' (2017) Washington and Lee L Rev 1345, 1386.

regulation. In the Communication *Tackling Illegal Content Online*, this point is reinforced by endorsing the view that:

> In order to ensure a high quality of notices and faster removal of illegal content, *criteria* based notably on respect for fundamental rights and of democratic values *could be agreed by the industry at EU level*. This can be done *through self-regulatory mechanisms or within the EU standardisation framework*, under which a particular entity can be considered a trusted flagger, allowing for sufficient flexibility to take account of content-specific characteristics and the role of the trusted flagger. Other such criteria could include internal training standards, process standards and quality assurance, as well as legal safeguards as regards independence, conflicts of interest, protection of privacy and personal data, as a non-exhaustive list.[49]

In addition, especially in the domain of terrorist propaganda, extremism, and hate speech, as mentioned, the European Commission 'encourages that the notices from trusted flaggers should be able to be fast-tracked by the platform', and user-friendly anonymous notification systems.[50] The goal of these mechanisms is to standardize procedures, relationship with the notifying parties and technologies used to implement them in order to further increase efficiency of law enforcement within the existing legal framework.

1.6. Codes of Conduct

In the aftermath of the refugee crisis, the fight against online hate speech became one of the important political issues. In a wave of regulatory euphoria, the German political leaders were perhaps the most inclined to regulate the removal of hate speech. Because self-regulations expected by governments appears to lag behind, Germany and later some EU member have threatened to bring in a law to impose heavy fines on a platform failing to take down hate-based criminal content[51]

In response to this, the European Commission acted swiftly by coordinating an EU-wide self-regulatory efforts by which online platforms should be directed to fight hate speech, incitement to terrorism and prevent cyber-bullying.[52] As an immediate result of this new policy trend, the Commission recently agreed with all major online hosting providers—including Facebook, Twitter, YouTube, Microsoft, Instagram, Snapchat and Dailymotion—on a code of conduct that endorses a series of commitments to combat the

---

[49] See Communication (n 1) s 3.2.1

[50] ibid s 3.2.1. and 3.2.3.

[51] See eg 2017 The Network Enforcement Act (*Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken*) (NetzDG) (DE).

[52] See Communication (n 7) 10. Several other documents coming out of the EU on anti-radicalization and countering extremism, including the UK Counter Extremism Strategy and the EU Parliament's Civil Liberties committee draft report on anti-radicalization, emphasize a stronger role for intermediaries in policing online content. See European Commission, 'Proposal for a Directive on Combating Terrorism and Replacing Council Framework Decision 2002/475/JHA on Combating Terrorism' (2 December 2015) COM(2015) 0625 final; European Parliament, 'Draft Report on Prevention of Radicalization and Recruitment of European Citizens by Terrorist Organizations' (1 June 2015) 2015/2063(INI); Home Department (UK), *Counter-Extremism Strategy* (Cmd 9148, 2015).

spread of illegal hate speech online in Europe.[53] The code spells out commitments such as faster notice and takedown for illegal hate speech that will be removed within 24 hours or special channels for government and NGOs notice to remove illegal content.[54] In partial response to this increased pressure from the EU regarding the role of intermediaries in the fight against online terrorism, major tech companies—Facebook, Microsoft, Twitter and YouTube—announced that they will begin sharing hashes of apparent terrorist propaganda.[55]

The Code of Conduct for hate speech is not the only EU-brokered self-regulatory mechanism increasing responsibility. Historically the first was the Memorandum of Understanding in the area of trademark infringements.[56] Recently, the European Commission adopted a new Code of Practice against disinformation.[57]

1.7. Filtering

Filtering and proactive monitoring have been increasingly sought—and deployed—as enforcement strategy online.[58] Proactive monitoring comes first—and largely—as a private ordering approach following rightholders and government pressures to purge the Internet from allegedly infringing content or illegal speech. In the mist of major lawsuits launched against them,[59] YouTube and Vimeo felt compelled to implement filtering mechanisms on their platforms on a voluntary basis. Google lunched Contend ID in 2008.[60] Vimeo adopted Copyright Match in 2014.[61] Both technologies rely on digital fingerprinting to match an uploaded file against a database of protected works provided by rightholders.[62]

Technologies from these initiatives inspired part of the solutions debated within the 2019 EU Copyright Reform. According to its Article 17, selected providers are subject to a preventive obligation in case they fail to conclude licensing agreements and are given

---

[53] See European Commission, European Commission and IT Companies Announce Code of Conduct on Illegal Online Hate Speech, Press Release (31 May 2016) <http://europa.eu/rapid/press-release_IP-16-1937_en.htm>.

[54] ibid

[55] See 'Google in Europe, Partnering to Help Curb the Spread of Terrorist Content Online' (*Google Blog*, 5 December 2016) <https://blog.google/topics/google-europe/partnering-help-curb-spread-terrorist-content-online>.

[56] See European Commission, 'Memorandum of Understanding on online advertising and IPR' (May 2011) <https://ec.europa.eu/growth/industry/intellectual-property/enforcement/memorandum-of-understanding-online-advertising-ipr_en>.

[57] See European Commission, 'Code of Practice on Disinformation' (28 September 2018) <https://ec.europa.eu/digital-single-market/en/news/code-practice-disinformation>.

[58] See Chapter 28 and 29

[59] See *Viacom Int'l v YouTube Inc*, 676 F3d 19 (2nd Cir 2012); *Capitol Records LLC v Vimeo*, 826 F3d 78 (2nd Cir 2015).

[60] See YouTube, How Content ID Works <https://support.google.com/youtube/answer/2797370?hl=en>.

[61] See Chris Welch, 'Vimeo Rolls Out Copyright Match to Find and Remove Illegal Videos' (*The Verge*, 21 May 2014) <https://www.theverge.com/2014/5/21/5738584/vimeo-copyright-match-finds-and-removes-illegal-videos>.

[62] YouTube (n 60).

necessary information to trigger such technologies. According to some, this effectively means imposition of filtering content recognition technologies to prevent the availability of infringing content.[63]

Enforcing online behaviour through automated or algorithmic filtering and fair use is heavily debate in the literature. Julie Cohen and Dan Burk argued that fair use cannot be programmed into an algorithm, so that institutional infrastructures will always be required instead.[64] In general, it was noted that 'the design of copyright enforcement robots encodes a series of policy choices made by platforms and rightsholders and, as a result, subjects online speech and cultural participation to a new layer of private ordering and private control.'[65] According to Matthew Sag, automatic copyright filtering systems 'not only return platforms to their gatekeeping role, but encode that role in algorithms and software' and fair use only nominally applies online.[66] On the other hand, Niva Elkin-Koren[67] and Husovec[68] argued that technologies might be the only way how we address the concerns of over-blocking on scale and with necessary speed.

You Tube and Facebook have been using other matching tools to filter 'extremist content.'[69] In this context, tech companies plan to create a shared database of unique digital fingerprints—known as hashes—that can identify images and videos promoting terrorism.[70] When one company identifies and removes such a piece of content, the others will be able to use the hash to identify and remove the same piece of content from their own network.[71] Similar initiatives equally relying on hashing technologies were also implemented in the area of child abuse material. PhotoDNA is Microsoft's technology that has been widely used to find the pictures and stop their distribution.[72] The technology is being used by the Internet Watch Foundation, which operates its dedicated

[63] See Chapter 28, s 4.2.

[64] See Dan Burk and Julie Cohen, 'Fair Use Infrastructure for Copyright Management Systems' (2000) Georgetown Public Law Research Paper 239731/2000 <https://ssrn.com/abstract=239731>.

[65] See Matthew Sag, 'Internet Safe Harbors and the Transformation of Copyright Law' (2017) 93 Notre Dame L Rev 499, 538.

[66] ibid

[67] See Niva Elkin-Koren, 'Fair Use by Design' (2017) 64 UCLA L Rev 22 (2017).

[68] See Martin Husovec, 'The Promises of Algorithmic Copyright Enforcement: Takedown or Staydown? Which is Superior? And Why?' (2019) Columbia J of Law & the Arts (forthcoming).

[69] See Joseph Menn and Dustin Volz, 'Excusive: Google, Facebook Quietly Move Toward Automatic Blocking of Extremist Videos' (*Reuters*, 25 June 2016) <http://www.reuters.com/article/us-internet-extremism-video-exclusive-idUSKCN0ZB00M>.

[70] Olivia Solon, 'Facebook, Twitter, Google and Microsoft Team up to Tackle Extremist Content (*The Guardian,* 6 December 2016) <https://www.theguardian.com/technology/2016/dec/05/facebook-twitter-google-microsoft-terrorist-extremist-content>.

[71] See 'Partnering to Help Curb Spread of Online Terrorist Content' (*Facebook Newsroom*, 5 December 2016) <https://newsroom.fb.com/news/2016/12/partnering-to-help-curb-spread-of-online-terrorist-content>.

[72] See Microsoft, PhotoDNA <https://www.microsoft.com/en-us/photodna>.

internet crawler,[73] and private firms, such as by Microsoft, Twitter, Google and Facebook for some of their own products.[74]

The European Commission would like to provide a regulatory framework for these initiatives with special emphasis on tackling the dissemination of terrorist content online. In a recent Recommendation, the Commission has singled out automated filtering means as the optimal policy solution:

> Hosting service providers should take proportionate and specific proactive measures, including by using automated means, in order (1) to detect, identify and expeditiously remove or disable access to terrorist content (36) (2) in order to immediately prevent content providers from re-submitting content which has already been removed or to which access has already been disabled because it is considered to be terrorist content[75]

A proposal for a Regulation for preventing dissemination of terrorist content online endorses similar principles and is under consideration before the EU Parliament.[76]

1.8. Website-blocking

Another enforcement tool popularized in the area of IP law has been the website blocking measures.[77] These injunctions led to considerable case-law in some of the Member States. Over the years, the courts tried to fleshing out conditions, legal and technical modalities under which such orders should be available on the national level in the European Union.[78] These measures were then sometimes adopted in the national law by means of administrative regulations which entrusted authorities with special powers to block websites under specific conditions. [79]

---

[73]See 'Using Crawling and Hashing Technologies to Find Child Sexual Abuse Material – the Internet Watch Foundation' (*NetClean*, 11 February 2019) <https://www.netclean.com/2019/02/11/using-crawling-and-hashing-technologies-to-find-child-sexual-abuse-material-the-internet-watch-foundation>.

[74] See Wikipedia, PhotoDNA <https://en.wikipedia.org/wiki/PhotoDNA>.

[75] European Commission, 'Recommendation on measures to effectively tackle illegal content online' C(2018) 1177 final.

[76] See European Parliament, 'Legislative resolution of 17 April 2019 on the proposal for a regulation on preventing the dissemination of terrorist content online' [2019] P8_TA-PROV(2019)0421.

[77] See, *inter alia*, Chapter 4, Chapter 16, Chapter 20, Chapter 29.

[78] See Martin Husovec and Lisa Van Dongen, 'Website Blocking, Injunctions and Beyond: View on the Harmonization from the Netherlands' (2017) 7 GRUR Int; Pekka Savola,' Proportionality of Website Blocking: Internet Connectivity Providers as Copyright Enforcers' (2014) 5(2) JIPITEC 116, 116-138.

[79] See eg AGCOM Regulations regarding Online Copyright Enforcement, 680/13/CONS, 12 December 2013 (IT) (providing AGCOM with administrative power to enforce online copyright infringement); Royal Legislative Decree no 1/1996, enacting the consolidated text of the Copyright Act, 12 April 1996 (as amended by the Law No. 21/2014, 4 November 2014) (ES) (creating an administrative body—the Second Section of the Copyright Commission (CPI)—which orders injunctions against information society services who infringe on copyright); Omnibus Bill no 524 of 26 June 2013, amending provisions in various laws and decrees including Law no 5651 'Regulation of publications on the internet and suppression of crimes committed by means of such publications', Law No 5809 'Electronic Communications Law' and others (TR) (empowering the Presidency of Telecommunications and Communications with broad administrative enforcement prerogatives online); Federal Law no 139-FZ, on the protection of children from information

Even before these court-imposed injunctions entered the European landscape, a number of providers were engaging in voluntary website blocking schemes. Perhaps the most prominent of these was the anti-child abuse program operated by the Internet Watch Foundation (IWF).[80] In 2002, IWF started distributing its URL list for the purposes of implementing blocking or filtering solutions.[81] Next to internet access providers, a number of other technological companies voluntarily subscribe to the list.[82]

2. Mechanisms and Legal Challenges

After reviewing the most significant ways how the landscape of responsibilities is shifting beyond mere legal liability-imposed baseline, it is time to highlight how these are results of different mechanisms.

2.1. Market and Private Ordering

In particular in the area of IP, a lot of increased responsibilities is a result of private ordering achieved through markets. Broadly speaking, this happens for two reasons. Either it is in intermediary's self-interest to implement such enforcement tools, or appears rational given the business dealings with the right holders.

As for the first category, a number of factors contribute to self-interest in increasing one's own responsibility.[83] First of all, it is user experience. Often, illegal content misleads users or attempts to defraud them. For instance, it makes commercial sense for a newspaper to remove abusive or spam comments because they can hurt users' feelings or expose them to fraud. If an environment is dominated by offensive comments, many readers are discouraged from contributing themselves[84] and this is bad for the business of intermediaries. Second, it is credibility and reputation, which services often strive for. More accurate user content is more competitive, and has a better potential to attract advertising or other investments. For instance, Yelp, despite having no legal obligation to

harmful to their health and development and other legislative acts of the Russian Federation (aka 'Blacklist law'), 28 July 2012 (putting the *Roskomnadzor* in charge of the Registry and site blocking enforcement); Act on the establishment and operation of Korea Communications Commission (KCC) last amended by Act no 11711 of 23 March 2013 (establishing the KCC implementing deletion or blocking orders according to the request and standards of the Korea Communications Standards Commission, also instituted by the same law). For further in-depth discussion of administrative enforcement of IP rights online, see *Chapter 30*.

[80] See Internet Watch Foundation (IWF), URL List Policy <https://www.iwf.org.uk/become-a-member/services-for-members/url-list/url-list-policy>.

[81] ibid

[82] IWF, IWF URL List recipients <https://www.iwf.org.uk/become-a-member/services-for-members/url-list/iwf-url-list-recipients>

[83] See Marin Husovec, *Injunctions Against Intermediaries in the European Union: Accountable But Not Liable?* (CUP 2017) 13.

[84] For an overview of industry practices and their corresponding business reasons, see Emma Goodman, *Online Comment Moderation: Emerging Best Practices* (WAN-IFRA 2013) <https://www.wan-ifra.org/reports/2013/10/04/online-comment-moderation-emerging-best-practices>.

do so, has incorporated a right-to-reply into its review service after public pressure from the business community.[85]

Perhaps more typical situations are when increased responsibility results from market transactions. Some right holders might be in a position to cut deals with platforms, or leverage their existing business relationships. To give an example, Amazon, asking its users to review their purchasing experience with sellers, is in a business relationship with both sellers and users (buyers). Sellers will certainly voice their concerns about fraudulent reviews in negotiations about conditions for sale and failure to respond to such demands could lead them away from Amazon to its competitors. Provided that the market is competitive, Amazon must internalize harm of its customers by an action. A different type of right holders is existing business partners and thus have other leverage points. It has been observed in many countries that voluntary enforcement schemes were usually initiated when intermediaries such as Internet access providers tried to vertically integrate into markets where they had to do business with major right holders and license their content (eg video on demand). License agreements then often served as a tool for negotiating higher enforcement efforts of the same intermediaries for their other services.[86]

From the measures discussed in the earlier part, especially filtering and changes to online search could be said to result from these incentives. YouTube's ContentID was a win-win solution for YouTube which was not interested to continuous takedown of content and needed a way how to credibly monetize videos and increase collaboration with the right-holders. The downside of these privately agreed upon solutions is that they happen entirely in the dark, and thus the public has very little information about them. Terms of their operation are often confidential. Especially for human rights law, this creates a challenge because, without governmental intervention, the exposure to the legal safeguards is more challenging. The human rights, after all, were designed to protect against the state.

2.2. Corporate Social Responsibility

Corporate social responsibility theory have been ported to cyberspace to deploy human rights principles to non-public bodies, which operate largely outside the remit of traditional human rights law.[87] Arguments have been made that obligations pertaining to States—such as those endorsed by the UN Human Rights Council declaration of Internet

[85] See Claire Miller, 'The Review Site Yelp Draws Some Outcries of Its Own' *New York Times* (New York, 3 March 2009); Claire Miller, 'Yelp Will Let Businesses Respond to Web Reviews' *New York Times* (New York, 10 April 2009).

[86]For instance, in the Netherlands, US right holders, such as Disney and Warner, attempted to leverage their rights to content, when some Dutch providers decided to start providing video on demand. Right holders were reported to only license if the providers implemented some form of disconnection strategy. See Door A. Vermeer, 'Vrije internettoegang ook in Nederland onder vuur' (Bits of Freedom, 4 January 2011) accessed 19 November 2014 (Dutch provider @Home, currently Ziggo, in a press release from 2006 stated that, in the course of a VOD-deal, they also agreed to a three strikes-regime)

[87] See Laidlaw (n 4) (noting that ultimately, however, the largely voluntary nature of CSR instruments makes it a problematic candidate as a governance tool for IIGs and freedom of speech).

freedom as a human right[88]—should be extended to online platforms as well.[89] Other international instruments to that effect have been identified in the Declaration of Human Duties and Responsibilities,[90] the preamble of the UN Norms on the Responsibilities of Transnational Corporations and Other Business Enterprises,[91] and the UN Guiding Principles on Business and Human Rights.[92] Recently, the United Nations Human Rights Council adopted a resolution on the promotion, protection and enjoyment of human rights on the internet, which also addressed a legally binding instrument on corporations' responsibility to ensure human rights.[93]

In the European Union, the Directive on the disclosure of non-financial and diversity information by certain large undertakings and groups prescribes certain level of transparency and accountability to large companies at least when it comes to 'environmental, social and employee matters, *respect for human rights*, anti-corruption and bribery matters'.[94] Recital 9 directly references the United Nations (UN) Global Compact, the Guiding Principles on Business and Human Rights implementing the UN 'Protect, Respect and Remedy' Framework. This non-financial performance information should help investors, consumers, policy makers and other stakeholders to evaluate the large companies in the economy and potentially indirectly encourages them to develop a responsible approach to business.[95]

Corporate social responsibility is sometimes hard to distinguish from another reason why intermediaries increase their responsibility—the desire to avoid regulation. A good example in this space is Facebook's increased focus on tackling spread of disinformation.[96] Unlike in other areas, the risk of Facebook being held liable for disinformation is often not too severe because such information is not always illegal. At the same time, not acting could provide ground for intervention by legislation. Facebook,

---

[88] See Human Rights Council of the United Nations, Resolution on the Promotion, Protection and Enjoyment of Human Rights on the Internet (2012).

[89] See Florian Wettstein, 'Silence as Complicity: Elements of a Corporate Duty to Speak out Against the Violation of Human Rights' (2012) 22(01) Business Ethics Quarterly 37, 37–61; Stephen Chen, 'Corporate Responsibilities in Internet-Enabled Social Networks' (2009) 90(4) Journal of Business Ethics 523, 523–536.

[90] See UNESCO Declaration of Human Duties and Responsibilities (Valencia Declaration) (1998).

[91] See UN Norms on the Responsibilities of Transnational Corporations and Other Business Enterprises (August 13, 2003).

[92] See United Nations, Human Rights, Office of the High Commissioner, Guiding Principles on Business Human Rights: Implementing the United Nations 'Protect, Respect, and Remedy' Framework (2011) [hereinafter UN GPBHRs].

[93] See United Nations Human Rights Council, The Promotion, Protection and Enjoyment of Human Rights on the Internet, A/HRC/RES/26/13 (20 June 2014).

[94] Directive 2014/95/EU of the European Parliament and of the Council of 22 October 2014 amending Directive 2013/34/EU as regards disclosure of non-financial and diversity information by certain large undertakings and groups Text with EEA relevance [2014] OJ L 330/1, Art 29a (emphasis added). See also ibid Art 19a.

[95] This information should be available in the company reports starting 2018.

[96] See eg Adam Mosseri, 'Working to Stop Misinformation and False News' (*Facebook for Media*, 7 April 2017) <https://www.facebook.com/facebookmedia/blog/working-to-stop-misinformation-and-false-news>.

however, sells its efforts as part of its ambition to be a good citizen, its corporate social responsibility.

The obvious downside of the corporate social responsibility approach is that it does not prescribe any specific steps and rather tries to create environment in which companies will act in a responsible way. The expectations are very often very vague and thus hard to measure or evaluate.

2.3. Involuntary Cooperation in IP Rights Enforcement

Increasingly, governments—and interested third parties such as intellectual property rightholders—try to coerce online intermediaries into implementing voluntary measures and bear much of the risk of online enforcement. Husovec argued that the European Union law increasingly forces Internet intermediaries to work for the rightsholders by making them accountable even if they are not tortiously liable for actions of their users.[97] According to Husovec, the shift from liability to accountability has occurred by derailing injunctions from the tracks of the tort law.[98] The practical outcome of this was that right holders could potentially ask for all sorts of help in enforcement of their rights without having to argue about what intermediaries did wrong. This is because the sole reason for their involvement was that they are in a position which attracts responsibility as such. From the discussed examples, we could see a number of enforcement tools to originate in this mechanism. The website-blocking orders are the primary example.

The approach of adding responsible duties beyond those provided in the liability framework is becoming more present in the policy debates. The recent UK Government's *Online Harms White Paper* reinforces this discourse by proposing a new duty of care towards users, holding companies to account for tackling a comprehensive set of online harms, ranging from illegal activity and content to behaviours which are harmful but not necessarily illegal.[99] The goal of the proposal is to set out 'high-level expectations of companies, including some specific expectations in relation to certain harms'.[100] The violation of the duty of care would be assessed separately from liability for particular items of harmful content. This 'systemic form of liability' essentially superimposes a novel duty to cooperate on the providers and turns it into a separate form of responsibility, which is enforceable by public authorities by means of fines.[101] At the same time, it leaves the underlying responsibility for individual instances of problematic content intact.

2.4 Public Deal-making

One of very prevalent mechanisms observed in the responsibility landscape is the phenomenon of deal-making with public authorities. As a result of the liability safe

---

[97] See Husovec (n 83).

[98] ibid

[99] See Department for Digital, Culture, Media & Sport and Home Department, *Online Harm* (White Paper, Cp 59, 2019).

[100] ibid 67.

[101] ibid 59.

harbours, the providers are partially freed from responsibility for their users' content. Thus, they effectively have power to decide about the content which users post. However, this power is not supplement by a responsibility towards their users to respect their speech rights in some particular form. This has famously led Tushnet to call it 'power without responsibility'.[102] This responsibility gap,[103] then invites the government to pressure for removal of information without following a proper process. Again, public deal-making is here closely intertwined with the interest of platforms to avoid new forms of regulation.

In *Against Jawboning*, Derek Bambauer discusses government pressure on Internet intermediaries that spans a large variety of content types and subject matter.[104] Bambauer cites Representative James Sensenbrenner, pressing US Internet Service Provider Association to adopt putatively voluntary data retention scheme in the following terms: 'if you aren't a good rabbit and don't start eating the carrot, I'm afraid we're all going to be throwing the stick at you'.[105] A cost-benefit analysis would most likely suggest online intermediaries to play along and adopt pushed solutions.

Many of the enforcement tools earlier presented result, at least in part, from such governmental pressure.[106] The EU Code of Conduct for hate speech is perhaps the most prominent European example. The problem of these solutions is due process, prior restraint, and generally applicability of the human rights safeguards.[107]

2.4 Circulation of Solutions

As demonstrated by a number of examples, sometimes solutions that first originated in private ordering or injunction cases, eventually inspired changed in the law. For instance, ContentID inspired the European plaintiffs to ask for filtering, and that has inspired the European Commission to propose it in the law. However, the same cycle has also a reverse order. For instance, HADOPI law inspired private plaintiffs to demand similar

[102] Rebecca Tushnet, 'Power Without Responsibility: Intermediaries and the First Amendment' (2008) 76(4) George Washington L Rev 986, 986.

[103]See ibid; Daphne Keller, 'Who Do You Sue? State and Platform Hybrid Power over Online Speech' (2019) Aegis Series Paper no 1902 <https://www.lawfareblog.com/who-do-you-sue-state-and-platform-hybrid-power-over-online-speech>.

[104] See also Derek Bambauer, 'Against Jawboning' (2015) 100 Minnesota L Rev 51 (discussing federal and state governments increasing regulation of on-line content through informal enforcement measures, such as threats, at the edge of or outside their authority).

[105] ibid 51-52.

[106] See Danielle Keats Citron, 'Extremist Speech, Compelled Conformity, and Censorship Creep' (2018) 93 Notre Dame L Rev 1035.

[107] See Evelyn Aswad, 'The Role of US Technology Companies as Enforcers of Europe's New Internet Hate Speech Ban' (2016) 1(1) Columbia Human Rights L Rev Online 1, 6.

solutions in countries where legislation was absent, eg Ireland,[108] or similar private ordering schemes in the United States.[109]

3. Conclusions

Responsibility of intermediaries has emerged as a powerful slogan for policy makers. The European Commission has plainly admitted in recent documents that the Digital Single Market to come has been shaped according to the idea that 'responsibility of online platforms is a key […] issue', stressing that the path is set 'towards an enhanced responsibility of online platforms'. The new terminology, however, does represent a substantial shift in intermediary liability theory that apparently would move away from a well-established utilitarian approach toward a moral approach by rejecting negligence-based intermediary liability arrangements. In turn, this theoretical approach portends the enhanced involvement of private parties in online governance and a broader move towards private enforcement online. Public enforcement lacking technical knowledge and resources to address an unprecedented challenge in terms of global human semiotic behaviour would coactively outsource enforcement online to private parties. The deployment of miscellaneous self-regulation and voluntary measures—such as graduated response, monitoring and filtering, website-blocking, online search manipulation, payment blockades and follow-the-money strategies, and private DNS content regulation—reflects this change in perspective.

This development poses plenty of challenges. First, enforcement through private ordering and voluntary measures moves the adjudication of lawful and unlawful content out of public oversight. In addition, private ordering—and the retraction of the public from online enforcement—does push an amorphous notion of responsibility that incentivizes intermediaries' self-intervention to police allegedly infringing activities in the Internet. Further, enforcement would be looking once again for an 'answer to the machine in the machine'.[110] By enlisting online intermediaries as watchdogs, governments would *de facto* delegate online enforcement to algorithmic tools—with limited or no accountability.[111] Finally, tightly connected to the points above, transferring regulation and adjudication of Internet rights to private actors highlights unescapable tensions with fundamental rights—such as freedom of information, freedom of expression, freedom of business or a fundamental right to Internet access—by limiting access to information, causing chilling effects, or curbing due process.

---

[108] See *Sony Music Entertainment (Ireland) Limited v UPC Communicaitons Ireland Limited (No 1)* [2015] IEHC 317

[109] See Kerry Sheehan, 'It's the End of the Copyright Alert System (as We Know It)' (*Electronic Frontier Foundation*, 6 February 2017) <https://www.eff.org/deeplinks/2017/02/its-end-copyright-alert-system-we-know-it>.

[110] Charles Clark, 'The Answer to the Machine is in the Machine', in Bernt Hugenholtz (ed), *The Future of Copyright in a Digital Environment* (Kluwer Law Int'l 1999) 139. See also Christophe Geiger, 'The Answer to the Machine Should Not Be the Machine, Safeguarding the Private Copy Exception in the Digital Environment' (2008) 30 EIPR 121

[111] See Joshua Kroll and others, 'Accountable Algorithms' (2017) 165 U Pa L Rev 633.