Technology and Regulation

~~Content Not Available~~

Why The United Kingdom's Proposal For A "Package Of Platform Safety Measures" Will Harm Free Speech

Mark Leiser* & Edina Harbinja**

#online harms, duty of care, platform regulation, online safety

m.r.leiser@law.leidenuniv.nl

e.harbinja@aston.ac.uk

This article critiques key proposals of the United Kingdom's "Online Harms" White Paper; in particular, the proposal for new digital regulator and the imposition of a "duty of care" on platforms. While acknowledging that a duty of care, backed up by sanctions works well in some environments, we argue is not appropriate for policing the White Paper's identified harms as it could result in the blocking of legal, subjectively harmful content. Furthermore, the proposed regulator lacks the necessary independence and could be subjected to political interference. We conclude that the imposition of a duty of care will result in an unacceptable chilling effect on free expression, resulting in a draconian regulatory environment for platforms, with users' digital rights adversely affected.

1. Introduction

In April 2019, the UK Government's Department of Digital, Culture, Media and Sport ("DCMS") released its White Paper for "Online Harms" which, if accepted, would impose a new duty of care standard for online platform users to be overseen by an independent regulator.[1] If the White Paper proposals are brought into force, a regulator will be empowered to decide what activities and content are deemed harmful to Internet users.[2] After making this determination, the regulator can mandate intervention by internet providers to protect users from these harms.[3] If the recommendations in the DCMS White Paper are enacted into law, the UK could soon have a new Internet regulator (provisionally referred to as "OfWeb"[4]) that will have the statutory obligation of imposing a duty of care on online services to prevent a series of yet-to-be defined "online harms."[5] It moves enforcement of laws regulating content and free speech from courts and passes the obligation to private actors like Facebook, Google, and Twitter under the threat of penalties for non-compliance (possibly a loss of "license to operate"[6] or attaching personal liability to directors[7]). Rather than using the courts or other legitimate democratic institutions, platforms are obliged to determine and assess the harmfulness of user behavior before-and-after content is generated by users. The "duty of care" and the imposition of liability will change platforms and social media from a safe space for exercising fundamental speech rights to one where the state forces platforms to regulate content and decide what actions could be harmful.

However, the White Paper's "world-leading package of safety measures"[8] leaves important terms undefined, empowering politicians of the day to force platforms to respond to harms where there is little evidence to support its dangers. Based on a statutory duty of care to prevent users from harm, platforms will be forced to monitor, moderate, and remove user-generated content under the threat of "substantial fines".[9] As tight compliance deadlines strongly incentivize online service providers to comply with complaints swiftly, there is ample evidence from takedown regimes that platforms err on the side of caution, regardless of the actual merits of the claims.[10] The proposal also "empowers users" to hold platforms to account for failing to live up to their duty of care.[11] Fortunately, for people who care about the value of public discourse and are willing to resist the moral panic about the dangers of unregulated platforms, the White Paper should unite a disparate crew of civil society groups, desperate for a cause to rally behind since the Digital Economy Act 2010, the

1 Online Harms White Paper, Gov.UK, https://www.gov.uk/government/consultations/online-harms-white-paper (last updated Feb. 12, 2020) [hereinafter Online Harms White Paper].
2 Online Harms White Paper (n 1) ¶ 2.2 and ¶ 5.15.
3 Idem at ¶ 6.5.
4 Akin to Ofcom (The Office of Communications), broadcast and telecoms regulator, https://www.ofcom.org.uk/home, It is imagined that the office of the web' would be a newly created regulator named Ofweb.
5 Online Harms White Paper (n 1) 7 §-
6 Online Harms White Paper (n 1) 60 ¶ 6.5.
7 Idem.
8 Online Harms White Paper (n 1).
9 Online Harms White Paper (n 1) 59 § 6.
10 Hosting Intermediary Services and Illegal Content Online, Inst. for Info. L. (2018), https://www.ivir.nl/publicaties/download/hosting_intermediary_services.pdf.
11 Online Harms White Paper (n 1) 10 §16-18.

* Dr. Mark Leiser is Assistant Professor in Law and Digital Technologies, Leiden Law School, The Netherlands.
** Dr Edina Harbinja, Senior Lecturer in Media/Privacy Law, Aston University, United Kingdom.

Received 30 March 2020, Accepted 8 Aug 2020, Published 8 Sept 2020

SOPA[12]/ACTA[13] protests of 2012[14], and Articles 10 and 17 of the new Copyright Directive.[15] The DCMS intervention might just also help the typical citizen understand why Article 10 of the European Convention of Human Rights[16] is so important to the functioning of our modern society and end the present cycle of everything on the Internet is bad. With free expression at the heart of Western concepts like the marketplace of ideas, democratic deliberation, and the media's role in holding power to account, the challenge of any regulatory intervention online is targeting the effort at the right people, the legitimacy of the intervention, the proportionality of the measure and the effectiveness of steps taken, while ensuring media pluralism and the protection of low-level speech.[17]

This Article provides a brief overview of the events and the numerous hearings and interventions by the UK Parliament and the Government that led to the production of the White Paper. The Article then critiques the "duty of care" proposed in the White Paper, concluding that the imposition of a duty will chill free speech. The next section focuses on the role and independence of the proposed regulator, "OfWeb." This is followed by a critique of the harms identified by the Department of Culture, Media and Sport as justification for the duty of care. The Article concludes with recommendations and next steps.

2 The Bureaucratic Responses

The White Paper mirrors large parts of the output from the House of Lords' Communication Committee Report[18] and a previous report from the House of Commons DCMS Committee titled "Disinformation and 'fake news': Final Report"[19]: an amalgamation, in part, of what special interest groups believe is in the best interests of their members, rather than the wider digital community at large. The White Paper relies heavily on evidence from NGOs and charities like Doteveryone, the Children's Commissioner, Internet Matters, and Girl Guiding.[20] Upon reading the evidence submitted, one could easily conclude that the digital environment remains extremely hazardous and generally unsafe. On the contrary, the public generally believes the Internet is a good thing. In Ofcom's examination of online users, 59% of adults said the benefits outweigh the risks. Only a small percentage said the opposite. Furthermore, 61% of children said the Internet makes their lives better.[21] So where does this nuanced vision come from?

After the Cambridge Analytica scandal,[22] a series of hearings across a wide range of UK government entities were launched into the role of platforms in everything from disrupting democracy to causing long-term harm to children, facilitating abusive content.[23] The general consensus was that "something must be done."[24] Ironically, those same Members of Parliament (MPs) took to platforms to publicize how their plan was going to make the Internet safe again. The White Paper follows the Government's proposals set out in the Internet Safety Strategy Green Paper from October 2017[25] and the Digital Charter from January 2018.[26] The key principles for future regulation are parity ("what is unacceptable offline should be unacceptable online"), *openness*, *transparency*, the *protection of human rights*, and the *protection of children*.[27] In February 2019, the House of Commons DCMS Committee published their report, "Disinformation and 'Fake News'," [28] calling for government action to curtail the dissemination of deceptive content.[29] The White Paper goes even further, including disinformation within the list of harms that platforms will be under a duty of care to prevent.[30]

The Government and both Houses of Parliament agreed that there needs to be extensive regulation of the Internet and, in particular, social media.[31] To justify the need for intervention, they cite everything from issues with political advertising[32] (in particular, the UK's referendum on the Continued Membership of the European Union (Brexit)), to fake news and online manipulation,[33] data breaches by the tech giants,[34] the lack of competition in the Internet's mainstream (social

12 https://www.congress.gov/bill/112th-congress/house-bill/3261/text
13 https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX-:52010IP0432&from=GA.
14 The US Congress debated two bills (Stop Online Piracy Act (SOPA) was designed to protect the copyright creative industries. The bills were ultimately rejected after unprecedented protests and complaints to American representatives and Senators. The Anti-Counterfeiting Trade Agreement (ACTA) protests flared in Europe out of the belief that the openness of the Internet would be compromised.
15 Council Directive (EU) 2019/790, 2019 O.J. (L 130) 92 (EC).
16 The Convention for the Protection of Human Rights and Fundamental Freedoms was opened for signature in Rome on 4 November 1950 and came into force on 3 September 1953. As of 16 May 2018, it counts 47 States parties.
17 Jacob Rowbottom, To Rant, Vent and Converse: Protecting Low Level Digital Speech, 71 *Cambridge L.J.* 355, (2 Apr. 2012).
18 Regulating in a digital world: Final report published, House of Lords Communication Committee (9 Mar. 2019), https://www.parliament.uk/business/committees/committees-a-z/lords-select/communications-committee/inquiries/parliament-2017/the-internet-to-regulate-or-not-to-regulate/.
19 Disinformation and 'fake news': Final Report published, House of Commons Digital, Culture, Media and Sport Committee (14 Feb. 2019), https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf.
20 Online Harms White Paper (n 1).
21 Online Nation, Ofcom 3 (30 May 2019) https://www.ofcom.org.uk/__data/assets/pdf_file/0025/149146/online-nation-report.pdf 141.
22 The Cambridge Analytica files, (*The Guardian*) https://www.theguardian.com/news/series/cambridge-analytica-files (last accessed 6 May 2019).
23 See ICO issues maximum £500,000 fine to Facebook for failing to protect users' personal information, Info. Commissioner's Off. (25 Oct. 2018), https://ico.org.uk/facebook-fine-20181025; House of Commons Digital, Culture, Media and Sport Committee, (n 19); Addressing harmful online content, Ofcom (18 Sept. 2018), https://www.ofcom.org.uk/__data/assets/pdf_file/0022/120991/Addressing-harmful-online-content.pdf; Government response to the Internet Safety Strategy Green Paper, HM Government (May 2018), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/708873/Government_Response_to_the_Internet_Safety_Strategy_Green_Paper_-_Final.pdf.
24 Sonia Livingstone, Rethinking the rights of children for the Internet Age, Available at https://blogs.lse.ac.uk/parenting4digitalfuture/2019/04/03/rethinking-the-rights-of-children-for-the-internet-age/ (last accessed 25 April 2019).
25 HM Government, Internet Safety Strategy – Green paper, October 2017, Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/650949/Internet_Safety_Strategy_green_paper.pdf (last accessed 6 May 2019).
26 Digital Charter, Gov.UK, https://www.gov.uk/government/publications/digital-charter (last accessed 8 Apr. 2019).
27 Ibid.
28 House of Commons Digital, Culture, Media and Sport Committee (n 19).
29 (n 19) 64.
30 Online Harms White Paper (n 1) 31
31 Online Harms White Paper (n 1); House of Lords Communication Committee (n 18) and House of Commons Digital, Culture, Media and Sport Committee (n 19).
32 Online Harms White Paper (n 1) 28-29. See also Vote Leave's targeted Brexit ads released by Facebook, (*BBC News*, 26 July 26 2018), https://www.bbc.co.uk/news/uk-politics-44966969.
33 Online Harms White Paper (n 1) 22-23. See Articles on Fake news, The Conversation, https://theconversation.com/uk/topics/fake-news-33438 (last accessed 30 May 2019).
34 Idem at 31-32; see generally Sam Schechner & Mark Secada, You Give Apps Sensitive Personal Information. Then They Tell Facebook, (*Wall St. J.*, 22 Feb. 2019), https://www.wsj.com/articles/you-give-apps-sensitive-personal-information-then-they-tell-facebook-11550851636.

networking, search engines, advertising),[35] child abuse[36] and harms to children (including self-harm threats),[37] terrorist and extremist content,[38] and even knife crime.[39] The Committee expressed confidence that their proposal will address all of these issues, regardless of how different their causes and consequences.[40]

The White Paper recommends establishing a new independent regulator for the Internet[41] and the adoption of a co-regulatory model similar to broadcast regulation,[42] despite the fact that wireless internet is transmitted in a similar manner to the broadcasting signal, the Internet has almost zero resemblance to broadcasting. The new "OfWeb" will draft codes of conduct that set out principles of online safety and the "duty of care,"[43] backed up by reporting requirements and effective enforcement powers.[44] The regulator will also have responsibilities to promote education and awareness-raising about online safety, take a risk-based approach, 'prioritising action to tackle activity or content where there is the greatest evidence or threat of harm, or where children or other vulnerable users are at risk.'[45] Additionally, the regulator will be tasked to safeguard innovation, and to protect digital rights, 'being particularly mindful to not infringe privacy and freedom of expression.'[46] This regulatory effort actually goes hand-in-hand with what Facebook CEO Mark Zuckerberg recently proposed in his regulatory vision for Facebook.[47] Understandably, the proposal has attracted opposition and warnings from civil society groups, human rights advocates, lawyers, and academics.[48] Given the DCMS's promises and passion for new regulation, it is unlikely that the submissions substantially alter the proposal, no matter how well-evidenced and reasonable.

The White Paper is the latest in a long line of government reports and policy documents emanating from, among others, the controversy surrounding the vote on the United Kingdom's continued membership in the European Union (Brexit) and concerns about Russian interference in democratic discourse. The DCMS White Paper is the latest of these reports and focuses on the identification of "online harms" that are then used to justify the creation of a new regulator for Internet platforms. The harms are linked to and supported by evidence and reports filed by a large number of civil society groups, NGOs, charities, and child protection advocates.[49] The DCMS White Paper argues that these online harms are severe enough to warrant new and Internet-specific regulation. It also claims to reflect the changing tide in the way the government and society think about the Internet.

Despite numerous laws already in place to tackle some of the identified harms and numerous laws regulating content, actions, and behavior,[50] the White Paper attempts to pass the government's own policing responsibilities onto platforms; in other words, "it is *your* platform, *you* have to deal with it". The ethos of the White Paper is simple: platforms are seen as public spaces and are no different from theme parks, offices, and restaurants. Risk-based legal regimes like the UK's Health and Safety Act and the EU's General Data Protection Regulation have successfully deployed a duty of care before; therefore, as Facebook et al. are places where people gather, the imposition of a duty of care will work between platforms and users too. As we discuss in the next section, there are fundamental flaws with this argument.

3 How a Duty of Care Will Chill Free Speech

A duty of care normally carries with it a "three-stage test of foreseeability, proximity, and policy."[51] Foreseeability and proximity involve an examination of the factual circumstances of the parties.[52] Policy considerations usually require the court to deploy foresight into the consequences for other parties, not part of the dispute. The test requires the court to determine whether there is a legal relationship between the parties of the dispute;[53] for example, does an employer have a duty of care to its employees? Does a business have a duty of care to its customers? Does a building site operator have a duty of care to its visitors? A legal requirement to keep a place safe not only makes sense, but also puts prevention at the heart of the legal regime. Secondly, one of the central purposes of a duty of care is to apply similar duties to similar facts.[54] Once a court establishes that a duty of care is owed between x and y in circumstances z, then that decision applies to all future cases of the same kind. Of course, this duty will then be foreseeable and more certain, and not vague as suggested in the White Paper, both in terms of harms and individuals owed to.

35 Online Harms White Paper (n 1) 27-28; Matt Binder, Google hit with $1.7 billion fine for anticompetitive ad practices, (*Mashable, 20* Mar. 2019), https://mashable.com/article/google-eu-antitrust-fine-ads/.

36 Online Harms White Paper (n 1) 50; Jamie Grierson, Met police 'overwhelmed' by surge in online child sexual abuse, (*The Guardian*, 28 Mar. 2019), https://www.theguardian.com/uk-news/2019/mar/28/london- metropolitan-police-overwhelmed-by-surge-in-online-child-sexual-abuse-watch-dog-finds.

37 Online Harms White Paper (n 1) 19; Sarah Marsh & Jim Waterson, Instagram bans 'graphic' self-harm images after Molly Russell's death, (*The Guardian*, 7 Feb. 2019), https://www.theguardian.com/technology/2019/feb/07/instagram-bans-graphic-self-harm-images- after-molly-russells-death.

38 Online Harms White Paper (n 1) 14; Preventing the dissemination of terrorist content online, European Parliament (Sept. 2018), http://www.europarl.europa.eu/legislative- train/theme-area-of-justice-and-fundamental-rights/file-preventing-the-dissemination-of-terrorist-content-online.

39 Online Harms White Paper (n 1) 13. The report implies that online content allegedly glorifies gangs, and has led to an increase in knife crimes (the Report cites ONS (2019). Crime in England and Wales, Year Ending September 2018. Available at: https://www.ons.gov.uk/ peoplepopulationandcommunity/crimeandjustice/bulletins/crimeinenglandandwales/yearendingseptember2018). This is unsupported and the causation/correlation is vague. See generally Knife Crime, (*The Guardian)*, https://www.theguardian.com/uk/knifecrime (last accessed 31 May 2019).

40 Online Harms White Paper (n 1) 11–41.

41 Idem at 57 (OfWeb is the suggested name. However, the paper suggests Ofcom may initially be given the task.).

42 Idem.

43 William Perrin & Lorna Woods, Reducing harm in social media through a duty of care, CarnegieUK Trust (8 May 2018), https://www.carnegieuktrust.org.uk/blog/reducing-harm-social-media-duty-care/.

44 Online Harms White Paper (n 1) Section 7.42 and 6.

45 Ibid, p. 53.

46 Ibid.

47 Regulating in a Digital World, House of Lords Select Committee on Comm. (9 Mar. 2019), https://publications.parliament.uk/pa/ld201719/ldselect/ldcomuni/299/29902.htm.

48 See, e.g., PI's take on the UK government's new proposal to tackle "online harms", Privacy Int'l (8 Apr. 2019), https://privacyinternational.org/news/2779/pis-take-uk-governments-new-proposal-tackle-online-harms; Jim Killock & Amy Shepherd, The DCMS Online Harms Strategy must "design in" fundamental rights, Online Rts. Group (8 Apr. 2019), https://www.openrightsgroup.org/blog/2019/the-dcms-online-harms-strategy-must-design-in-fundamental-rights; UK: More regulation of content online risks right to privacy and free expression, Article 19 (19 June 2018), https://www.article19.org/resources/uk-more-regulation-of-content-online-risks-rights- to-privacy-and-free-expression/.

49 Examples include reports cited in the Online Harms White Paper (n 1) 13 – 14.

50 Abusive and Offensive Online Communications: A Scoping Report, L. Commission at 66-96 (1 Nov. 2018), https://s3-eu-west-2.amazonaws.com/lawcom-prod-storage- 11jsxou24uy7q/uploads/2018/10/6_5039_LC_Online_Comms_Report_FINAL_291018_WEB.pdf.

51 *Caparo Industries v. Dickman* [1990] 2 AC 605 (HL).

52 Ibid per Lord Roskill at 629 - 627

53 Ibid per Lord Oliver of Aylmerton at 632 - 634

54 Ibid at 618 – 619 per Lord Bridge of Harwich; Brennan J. in the High Court of Australia in *Sutherland Shire Council v. Heyman* (1985) 60 A.L.R. 1, 43–44.

Applying a duty of care between platforms and users to speech, however, will require platforms to block entire categories of speech, based on a legal obligation to block specific kinds of harm in the future. Cyber-bullying might cause individualized harms, but another user might not view others' comments as bullying. Trolling and swearing, for example, might be completely unacceptable to one person, but acceptable to another. Trolling is purely subjective speech that may, on occasion, rise to the threshold of criminal speech. An example would be grossly offensive, obscene, indecent or menacing communications regulated by s. 127 of the Communications Act 2003. Trolling that does not pass this threshold, would not be considered criminal now. For this reason, there are no laws regulating this kind of speech or content wherein the legal test of harm or offence is subjective and a recipient of speech/content gets to be the sole determinant of whether something causes harm.[55] In a recent high-profile event, a columnist for the New York Times accused an academic of abuse for referring to him as a "metaphorical bedbug".[56] The DCMS report offers no guidance about how platforms should police metaphors.

In the UK, the test required before criminal charges will be brought against content posted on social media is one of such gross offensiveness that criminal charges should be brought.[57] After several controversial and high-profile prosecutions,[58] the Public Prosecutor issued guidelines for the prosecution of grossly offensive speech.[59] These limit prosecutions under Section 127 of the Communications Act 2003 to cases which go beyond those which are "[o]ffensive, shocking or disturbing; or [s]atirical, iconoclastic or rude; or [t]he expression of unpopular; or unfashionable opinion about serious or trivial matters, or banter or humor, even if distasteful to some or painful to those subjected to it."[60] The threshold for bringing criminal charges is high, yet the DCMS bases their report on broad categories of speech that does not come close to the threshold of criminality.

One of the challenges of regulating content is differentiating between the risk of harm and the *uncertainty* that the harm may or may not occur.[61] Imposing a duty of care on platforms to tackle harms associated with content *ex post* is fundamentally different from the imposition of a duty of care on platforms for the *uncertainty* that users may generate content that causes harm. Risk can be accounted and modelled for and quantified through pricing strategies, while *uncertainty* is a risk than cannot be measured. Its reliance on the evidence of numerous NGOs, charities, consumer protection groups, and other advocates informs us of numerous *types* of harms[62], but none can predict *when* or *how* these harms will take place.

The imposition of a duty of care backed by financial sanctions onto platforms, spins both the *uncertainty* about the frequency and the likelihood and validity of the harms themselves into the *risk* of harms associated with user-generated content. Once this occurs, risk can be modelled and priced. This is a dangerous path leading to a chilling effect on permitted content. First, are the unknowns. As Douglas Adams writes in the Hitchhiker's Guide to the Galaxy, "the major difference between a thing that might go wrong and that thing that cannot possibly go wrong is that when a thing that cannot possibly go wrong goes wrong, it usually turns out to be impossible to get at or repair."[63] Second, once something is priced, the imposition of a duty of care establishes a transaction cost for content; speech deemed too costly to the platform will be filtered, blocked and/or removed *ex ante* rather than *ex post*, especially when the uncertainty surrounding content is determined to have too high a transaction cost, regardless of its *actual risk*. In other domains, where one sees the imposition of a duty of care (i.e. environmental law[64] or health and safety[65]), the law serves to mitigate the distribution costs of uncertainty through legal conventions like the precautionary principle[66] or the preventative measures rule.[67] Prentice-Dunn & Rogers argue that preventative regimes operate best when it is possible to predict outcomes that are contrary to totally rational decision making.[68] But regulating speech through precaution or prevention is a disproportionate response to the various forms of uncertainty. An online harm may come about from one or more causes, and if it occurs, one or more effects. In isolation, an innocuous comment may cause little harm, but the cumulative effect might cause substantial harm.

The duty of care would require platforms to avoid content that would place them at fault for a list of harms. In practice, this means that platforms would be under a statutory obligation to take reasonable care. If there was no duty of care to prevent a certain harm; in these circumstances, the law would permit platforms to act unreasonably. Yet, cases in English law about duty of care are limited to whether

55 Section 127 of the Communications Act 2003 should be interpreted as an objective test. Would a reasonable person view the communication as 'grossly offensive or of an indecent, obscene or menacing character'.

56 @davekarpf, Twitter (26 Aug. 2019, 2:07 PM), https://twitter.com/davekarpf/status/1166094950024515584. See Allan Smith, A professor labeled Bret Stephens a 'bedbug.' Here's what the NYT columnist did next, (*NBC News*, 27 Aug. 2019), https://www.nbcnews.com/politics/politics-news/professor-labeled-bret-stephens-bedbug-here-s-what-nyt-columnist-n1046736.

57 Communications Act 2003, c. 21, § 127 (UK).

58 *Chambers v Director of Public Prosecutions* [2012] EWHC 2157 (PC). See also Azhar Ahmed Sentencing Remarks, Available at https://www.judiciary.uk/wp-content/uploads/JCO/Documents/Judgments/azhar-ahmed-sentencing-remarks-09102012.pdf (last accessed 12 Sept 2019); "Man who racially abused Stan Collymore on Twitter spared prison", (*The Guardian*, 21 Mar 2012), Available at https://www.theguardian.com/technology/2012/mar/21/man-racially-abused-collymore-twitter-spared-prison (last accessed 11 Sept 2019).

59 Director of Public Prosecutions, "Social Media - Guidelines on prosecuting cases involving communications sent via social media", Revised: 21 August 2018, Available at: https://www.cps.gov.uk/legal-guidance/social-media-guidelines-prosecuting-cases-involving-communications-sent-social-media (last accessed 23 March 2020).

60 Director of Public Prosecutions, "Interim guidelines on prosecuting cases involving communications sent via social media", Available at https://adam1cor.files.wordpress.com/2012/12/117342720-social-media-dpp.pdf, (last accessed 12 September 2019).

61 For a detailed explanation on the difference between risk and uncertainty, see Volz, K. G., & Gigerenzer, G. (2012). Cognitive processes in decisions under risk are not the same as in decisions under uncertainty. *Frontiers in Neuroscience*, 6, 105.

62 Department for Digital, Culture, Media & Sport and UK Council for Internet Safety, Online harms research publications, Available at https://www.gov.uk/government/collections/online-harms-research-publications (last accessed 3 Sept 2019).

63 Douglas Adams (2000), *Mostly Harmless* (New York: Del Rey).

64 At the European level, the precautionary principle was enshrined in the Maastricht Treaty in 1992. It is now included in Article 191 of the Treaty on the Functioning of the European Union among the principles underpinning EU environmental policy.

65 The Health and Safety at Work etc Act 1974; See also Management of Health and Safety at Work Regulations (MHSWR) 1999.

66 M.D. Rogers. Scientific and technological uncertainty, the precautionary principle, scenarios and risk management. *Journal of Risk Research*, 4(1):1–15, 2001; See also doteveryone, "A digital duty of care", Available at https://doteveryone.org.uk/wp-content/uploads/2019/02/Doteveryone-briefing-a-digital-duty-of-care.pdf (last accessed 3 September 2019).

67 Niskanen, T., Naumanen, P., & Hirvonen, M. L. (2012). An evaluation of EU legislation concerning risk assessment and preventive measures in occupational safety and health. *Applied ergonomics*, 43(5), 829-842.

68 Prentice-Dunn, S., & Rogers, R. W. (1986). Protection motivation theory and preventive health: Beyond the health belief model. *Health Education Research*. 1(3), 153-161.

the duty of care was applied properly and whether the party with the obligation acted reasonably.[69] Most duties of care cases involve an examination of the application of vague concepts like "foreseeability," "proximity," and "fair, just and reasonable."[70] The scope of the duty is unrelated to legal causation: should we limit the extent of the defendant's responsibility for the harm despite the fact that the defendant's fault was a but-for cause of the harm? It is difficult to contemplate that platforms, rather than user-generated content, is the but-for cause of harm, that can only be identified by a new regulator.

4 Online Harm Offensive

The government defines "online harm" as "online content or activity that harms individual users, particularly children, or threatens our way of life in the UK, either by undermining national security or by reducing trust and undermining our shared rights, responsibilities and opportunities to foster integration."[71] The vague definition ironically refers to the UK's "way of life" and "rights," but the list of harms contradicts this proposition. Having analyzed the White Paper,[72] the EDPS report on Online Manipulation,[73] the High Level Working Group's Report on Disinformation,[74] the House of Lords' Communication Committee Report on Regulating the Internet,[75] and any of the hundred other reports into deceptive media online, one could be forgiven for thinking that "harms" are an invention of the World Wide Web. Although it is clear that the Internet's architecture and scale make some harms easier to facilitate,[76] it is also true that the diffusion of harms, especially within the context of communication, has always been a danger for society.

The White Paper's framework aims to regulate harms "based on an assessment of their prevalence and impact on individuals and society."[77] Rather than relying on compelling evidence, this is based on a handful of surveys and media reports, largely provided by a variety of outside groups with their own agendas.[78] Even more troubling is the fact the list of harms is "by design, neither exhaustive nor fixed."[79] This is justified by claiming that a "static list could prevent swift regulatory action to address new forms of online harm, new technologies, content and new online activities."[80] Consequently, in the right political climate, OfWeb could theoretically proclaim *anything* harmful, including academic articles that criticize OfWeb's approach for its chilling effect on free expression.

The White Paper identifies twenty-three harms in total.[81] Some of them are already criminal offenses, others are so vague it would be a regulatory achievement for OfWeb to come up with a definition that sounds good in theory, but also works in practice. The DCMS report categorizes these harms in three groups: harms with a clear definition; harms with a less clear definition and underage exposure to legal content.[82] In the table below, we provide an overview of all the harms included in these three groups, referring to their current legal status, i.e. whether some of them are already a criminal offence, of their status is less clear from the perspective of the current laws.

Table 1 Harms with a Clear Definition[83]

Group 1: Harms with a Clear Definition	Status	Criminal Law Provision
Child sexual exploitation and abuse	Criminal Offense	Sections 47-51 of the Sexual Offences Act 2003.
Terrorist content and activity	Criminal Offense	Section 58 of the Terrorism Act 2000; Section 3 of the Counter-Terrorism and Border Security Act 2019.
Organized immigration crime	Criminal Offense	Modern Slavery Act 2015; Section 1 of the Asylum and Immigration (Treatment of Claimants, etc.) Act 2004; Sections 57 to 59 of the Sexual Offences Act 2003.
Modern slavery	Criminal Offense	Modern Slavery Act 2015.
Extreme pornography	Criminal Offense	Section 63 of the Criminal Justice and Immigration Act 2008.
Revenge pornography	Criminal Offense	Section 33 of the Criminal Justice and Courts Act 2015.
Harassment and cyber-stalking	Criminal Offense	Section 2, 2A, 4, 4A, Protection from Harassment Act 1997; Section 1 Malicious Communications Act 1988.
Hate crime	Criminal Offense	Public Order Act 1986; Racial and Religious Hatred Act 2006; Criminal Justice and Immigration Act 2008; Criminal Justice and Public Order Act 1994; For England, Wales, and Scotland, the Crime and Disorder Act 1998 makes hateful behavior towards a victim based on the victim's membership (or presumed membership) in a racial group an "aggravating factor" for the purpose of sentencing in respect of specified crimes. Sections 2, 2A, 4, 4A of the Protection from Harassment Act 1997 also apply for racially and religiously aggravated offences of harassment and stalking and putting people in fear of violence, and stalking involving fear of violence. Finally, there are communication offence under section 127(1) of the Communications Act 2003, or section 1 of the Malicious Communications Act 1988, with enhanced sentencing due to hostility towards one of the five protected characteristics.
Encouraging or assisting suicide	Criminal Offense	Section 2 and 2A of the Suicide Act 1961.
Incitement of violence	Criminal Offense	Section 44, 45 of the Serious Crime Act 2007.
Sale of illegal goods/ services such as drugs and weapons on the open internet	Criminal Offense	Section 1(1) of the Criminal Law Act 1977; Section 46 of the Serious Crime Act 2007; Fraud Act (2006), Misuse of Drugs Act (1971), or Firearms Act (1968).
Content illegally uploaded from prisons	Criminal Offense	Section 40D(3A) Prison Act 1952.
Sexting of indecent images by under 18s	Criminal Offense	Section 45 of the Sexual Offences Act 2003.

As seen in the table above, these harms are *already* illegal and there is no need to introduce entirely new laws for them. Some of them could benefit from further clarifications (e.g. terrorist content or harass-

69 Key cases: *Donoghue v Stevenson* [1932] AC 562; *Topp v London Country Bus* [1993] 1 WLR 976; *Home Office v Dorset Yacht Co Ltd* [1970] AC 1004.
70 Howarth, D. (2006). Many Duties of Care—Or a Duty of Care? Notes from the Underground. *Oxford Journal of Legal Studies*, 26(3), 449-472.
71 Online Harms White Paper (n 1) 30.
72 Online Harms White Paper (n 1).
73 EDPS Opinion on online manipulation and personal data, Opinion 3/2018, available at: https://edps.europa.eu/sites/edp/files/publication/18-03-19_online_manipulation_en.pdf.
74 The final report "A multi-dimensional approach to disinformation" is available at https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation.
75 House of Lords Communication Committee (n 18).
76 Council of Europe, Parliamentary Assembly, Resolution 1970 (2014), Internet and politics: the impact of new information and communication technology on democracy, Available at https://assembly.coe.int/nw/xml/XRef/Xref-XML2HTML-en.asp?fileid=20447&lang=en, Accessed 30 May 2019.
77 Online Harms White Paper (n 1) 30.
78 Ofcom (2018). Adults' Media Use and Attitudes Report. Available at: https://www.ofcom.org.uk/__data/assets/pdf_file/0011/113222/Adults-Media-Use-and-Attitudes-Report-2018.pdf Ofcom and ICO (2018). Internet users' experience of harm online 2018. Available at: https://www.ofcom.org.uk/__data/assets/pdf_file/0018/120852/Internet-harm-research-2018-report.pdf Accessed 6 May 2019.
79 Online Harms White Paper (n 1) 30.
80 Ibid.
81 Ibid, p. 31.
82 Ibid.
83 Ibid.

ment and cyberstalking).[84] We argue that this should not be done within the overarching bundle of dissimilar harms as suggest in the White Paper.

Table 2 Harms with 'less clear' definition[85]

Group 2: Harms with 'less clear' definition	**Status**	**Provision/Comment**
Cyberbullying and trolling	Potentially criminal in some instances, but vaguely defined and with serious implications for free speech.	Potentially subset of communication offences under section 2, 2A, 4, 4A of the Protection from Harassment Act 1997; Section 1 Malicious Communications Act 1988, but vague and depends on the definitions, which are vague and overlapping.[86]
Extremist content and activity	Criminal in many instances, but vaguely defined and difficult to apply uniformly.	Potentially Section 58 of the Terrorism Act 2000; Section 3 of the Counter-Terrorism and Border Security Act 2019, but this is already covered by terrorist content, so it is unclear why the extremist content is necessary as a "new harm."
Coercive behavior	Vaguely formulated.	Potentially Section 2, 2A, 4, 4A, Protection from Harassment Act 1997; Section 1 Malicious Communications Act 1988, but vague and depends on the definition. This harm can also be potentially confused with existing offences, such as harassment. Further offence is found in section 76 of the Serious Crime Act 2015, but it only relates to domestic abuse cases.
Intimidation	Potentially criminal, but also vaguely defined.	Section 4 and 4A, Protection from Harassment Act 1997, already illegal, and serious fear of violence offences in Section 4 and 4a of the Public Order Act 1986, so unnecessary as a vaguely defined and subjective harm here. Its vagueness could mean that the harm may include legitimate free speech.
Disinformation	Vague, regulation of "fake news" is in progress.	Potentially covered by Section 127(2)(a) or (b) of the Communications Act 2003 and Section 1 Malicious Communications Act 1988, ongoing law reform in the area. Section 51 of the Criminal Law Act 1977 covers a bomb hoax; Hoaxes involving noxious substances or things are covered under section 114(2) of the Anti-Terrorism, Crime and Security Act 2001; giving a false alarm of fire exists under section 49(1) of the Fire and Rescue Services Act 2004; impersonating a police officer - section 90 of the Police Act 1996; section 106 of the Representation of the People Act 1983 offence to make or publish any false statement of fact in relation to the personal character of a candidate prior to or during an election.[87]
Violent content	Vaguely defined and problematic - any violent content online, including artistic speech could be harmful.	It is unclear how this is different from harassment, fear of violence, threat, stalking and extreme pornography, and other already existing criminal offences, as noted above. Does it include artistic speech, video games, films and what implication can this vaguely defined harm have on free speech?
Advocacy of self-harm	Dangerous precedent, blurs the lines between free speech and 'advocacy', as well as support self-harm support groups on social media.	Not illegal, but the UK government has threatened to introduce legislation if platforms do not remove content promoting self-harm. The Law Commission notes "[publicizing] or glorifying self-harm is not ostensibly criminal either."[88] However, offence of causing grievous bodily harm with intent, contrary to section 18 of the Offences Against the Person Act 1861, could be used here, provided that the victim caused herself serious harm with intent, so assisting or encouraging such behavior could be guilty of an offence under sections 44 to 46 of the Serious Crime Act 2007.[89]
Promotion of female genital mutilation	Criminal	Female Genital Mutilation Act 2003 makes the Act illegal, but there is no offence relating to its promotion. but see ss. 44 - 46 of the Serious Crime Act, intentionally encouraging or assisting an offence; encouraging or assisting an offence believing it will be committed; and encouraging or assisting offences believing one or more will be committed.[90]

84 The Law Commission, 'Abusive and Offensive Online Communications: A Scoping Report', Law Com No 381, at: https://s3-eu-west-2.amazonaws.com/lawcom-prod-storage-11jsxou24uy7q/uploads/2018/10/6_5039_LC_Online_Comms_Report_FINAL_291018_WEB.pdf.

85 Online Harms White Paper (n 1)31.

86 The Law Commission (n 84) ¶ 8.8. The Commission also rightly points out the issues with studies that analyze these phenomena, such as: 'it is unclear whether the offending behavior being discussed would constitute a criminal offence under the criminal law, or whether the study is focused on generally unkind and unacceptable behavior, which falls short of an offence but is capable nevertheless of causing harm.' (emphasis by the authors).

87 But there are no offences of creating or spreading fake news per se in the UK, see Disinformation and 'fake news': Interim Report, Report of the Digital, Culture, Media and Sport Committee (2017-19) HC 363, available at https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/363/363.pdf. There are however civil remedies in defamation and media regulation. The ECHR in *Salov v Ukraine* (2007) 45 EHRR 51 held that Article 10 does not prohibit discussion or dissemination of information received, even if it is strongly suspected that this information might not be truthful. The Court suggested that otherwise it would deprive persons of the right to express their views and opinions about statements made in the mass media, and would therefore place an unreasonable restriction on the freedom of expression. For a broader discussion see B McNair, Fake News Falsehood, Fabrication and Fantasy in Journalism, Routledge, 2018; T McGonagle, "'Fake news': False fears or real concerns?" (2017) 35(4) *Netherlands Quarterly of Human Rights* 203

88 The Law Commission (n 84) ¶ 12.93, see also A Gillespie (2016), *Cybercrime: Key Issues and Debates*, Routledge, 200. As the Commission also finds, online communications promoting self-harm would need to pass the threshold of "obscene, indecent or grossly offensive" to be prosecuted under section 127 of the CA 2003. If considered obscene, its publication may be prosecuted under section 2 of the Obscene Publications Act 1959. This would need to ensure compatibility with the Human Rights Act 1998, as both the Commission and Gilespie warn. See The Law Commission (n 84) ¶ 12.95, or Gilespie ibid p. 201.

89 See Law Commission (n 84) ¶ 12.94. The Commission also questions the appropriateness of using criminal law in this context, ¶ 12.99.

90 The Law Commission (n 84) ¶ 12.64 notes: 'However, a general glorification of certain conduct, without an intention for a specific crime to be committed, would be difficult to fit within the terms of sections 44 to 46 of the SCA 2007'..

Table 3 Underage exposure to Legal Content[91]

Group 3: Underage exposure to Legal Content	Status	Comment
Children accessing pornography	Service providers' liability for making pornographic content available to persons under 18.	Digital Economy Act 2017, s 14 requires providers to prevent children from accessing pornography, under a threat of financial penalties (implementation has been delayed).[92]
Children accessing inappropriate material	Vague, undefined, and problematic - who decides what is 'inappropriate' and who decides whether a child can access? What is the role of parents and education in helping kids understand what is appropriate for them to engage with? Do we really want parents determining what content a child accesses about sexual health is appropriate.	There are existing provisions preventing children from accessing pornographic, obscene and other prohibited materials, as noted above. 'Inappropriate' as a category is extremely vague and open to interpretation, it is not certain whether it includes harmful online advertising, for example. It could also affect free speech of adults, children as well as other rights such as privacy.[93]
Under 13s using social media and under 18s using dating apps	Already the rule; however, rarely enforced; moral panic.	This is a question of adequate enforcement and age verification, as noted above.[94]
Excessive screen time	Moral panic.	Evidence that the risk outweighs benefits have not been conclusive and the real harm is often overestimated by the media and advocacy groups.[95]

Even for "clear" harms, there is some dispute whether the definition is clear enough and to what extent these should even be criminalized (e.g. the definition of what defines "extreme pornography" is limited to the anus, genitals, breasts; necrophilia and bestiality, while excluding other injuries to the body).[96] Some other harms are, of course, indisputably illegal, e.g. content related to child abuse. In the "less clearly" defined harms group,[97] the scope of harm goes far beyond what is permitted by UK law in the offline world.

Some harms are not illegal in the offline world and some are difficult to define without adversely affecting protected free speech (e.g. trolling, violent content, intimidation or disinformation). The legislators have been avoiding to criminalize all trolling for instance, as a legal definition would potentially include legitimate free speech. Furthermore regulating "to achieve equivalent outcomes online and offline" requires a platform to determine the comparable offline offence. For example, Section 127 of the Communications Act 2003 is far more restrictive to online speech than any offline equivalent.[98] Thus, DCMS are suggesting that in order to achieve equivalent outcomes one needs tighter restrictions on online speech.

5 Free Speech

Article 10 of the European Convention on Human Rights reflects a middle ground between unfettered speech under the First Amendment of the US Constitution and the authoritarian approach to direct control of the media within a territory proposed by the Soviets.[99] Beyond the Convention, the European "project" is rooted in concepts of dignity and social justice. The first pillar of the Treaty of Maastricht on European Union refers to and of social protection and equality between men and women.[100] The European Social Model is based on the concept of social cohesion. That individuals should not have to put up with promulgation of views deeply hurtful to themselves or their communities is a basic tenet of this approach to the needs of a pluralistic society.

Article 10, Para 1 ECHR provides that the freedom of expression "shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers." The same wording appears almost verbatim in Article 19, Para 2, of the International Covenant on Civil and Political Rights (ICCPR).[101] Recent judgments from the ECtHR have highlighted State obligations to not only respect freedom of expression by refraining from interference and censorship, but also to ensure a favorable environment for inclusive and pluralistic public debate.[102] Thus, Article 10 requires the adoption of "positive measures" to protect the "passive" element (right to be correctly informed) of free expression.[103] There is a strong link between the two fundamental rights: an election process is "free" when, not only the electorate's choice is determined by access to the widest possible range of proposals and ideas, but also if the election results do not risk being distorted or altered by the dissemination of false information.

Article 10(2) of the European Convention contains a list of the exceptions to the right of free expression contained in 10(1).[104] Any limitations to this must be not only proportional but achieve a legitimate aim for which the restriction is in place, in accordance with the law, and is necessary in a democratic society. European Union activities must respect the proportionality principle and, in areas that did not fall within its exclusive competence, the principle of subsidiarity[105] which encourages regulation at the local level, "as close to the citizen

91 Online Harms White Paper (n 1) 31.
92 The UK Government has struggled to find the appropriate way to implement the age verification system, so the implementation and enforcement of this provision has been delayed a few times already, see J Waterson and A Hern, 'UK age-verification system for porn delayed by six months', (*The Guardian*, 20 Jun 2019), at: https://www.theguardian.com/technology/2019/jun/20/uks-porn-age-verification-system-to-be-delayed-indefinitely
93 Regulation is not the silver bullet for all the risks associated with children using the Internet. The matrix of opportunities and risk associated with this is very complex, and researchers have identified various model to assess and address this risk, see e.g. Livingstone, Sonia, Mascheroni, Giovanna and Staksrud, Elisabeth, 'European research on children's internet use: assessing the past and anticipating the future', 2018 *New Media and Society*, 20 (3). pp. 1103-1122; E Staksrud (2013), *Children in the Online World: Risk, Regulation, Rights*, Aldershot: Ashgate.
94 There are, however, many myths associated with this issue and the picture may not always be as presented in the media. See e.g. Livingstone, Mascheroni and Staksrud, ibid.
95 Livingstone, Sonia and Franklin, Keely (2018) Families with young children and 'screen time' advice. *Journal of Health Visiting*, 6 (9). pp. 434-439.
96 Section 63. Criminal Justice and Immigration Act 2008.
97 Online Harms White Paper (n 1) 31.
98 Section 127 Communications Act 2003.
99 For a good overview of the debates of the foundations of free expression in international law, see Leiser M.R. (2019), 'Regulating Computational Propaganda: Lessons from International Law', *Cambridge International Law Journal* 8(2): 218-240.
100 Article 1 and 2 of the Treat of Maastricht on European Union; See also https://www.europarl.europa.eu/RegData/etudes/fiches_techniques/2013/010103/04A_FT(2013)010103_EN.pdf
101 UN General Assembly, International Covenant on Civil and Political Rights, 16 December 1966
102 See *Lingens v Austria* (App No 9815/82) 8 July 1986; *Janowski v Poland*, Judgment (Merits), (App No 25716/94) 21 January 1999; *Tammer v Estonia*, (App. 41205/98), 6 February 2001; *Janowski v Poland*, Judgment (Merits), (App No 25716/94) 21 January 1999
103 J.F. Akandji-Kombe, Positive obligations under the European Convention on Human Rights. *A guide to the implementation of the European Convention on Human Rights—Human rights handbooks*, No. 7. 2007, Strasbourg: Council of Europe. Available at: https://rm.coe.int/168007ff4d.
104 Council of Europe, European Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by Protocols Nos. 11 and 14, 4 November 1950
105 Article 5 of the EC Treaty; now Article 5(3) of the Treaty on European Union (TEU) and Protocol (No 2) on the application of the principles of subsidiarity and proportionality.

as possible."[106]

While the European Court of Human Rights will generally not interfere with a Member State's margin of appreciation to determine whether a particular measure is "necessary," the ECtHR does not take kindly to measures that are not properly prescribed or satisfy the quality of law test, especially when it comes to expression.[107] However, Strasbourg is only engaged in "...applying the principle of subsidiarity when national authorities have demonstrated in cases before the court that they have taken their obligations to secure Convention rights seriously."[108]

What is it about free speech that irritates UK regulators? The Internet is a communications system without any front-end filter for user-generated content. It exists inside a legal system that has historically regulated different forms of non-digital speech. The framework for free expression as a *fundamental right* had to find a way to slot on top of an existing body of law that restricts speech in certain instances; for example, the longstanding law of copyright restricts the use of intellectual creation[109] and defamation law generally restrains speech that lowers the standing of one's reputation in the eyes of right-minded members of society.[110] Attributing responsibility (and in many cases, liability) was relatively straightforward when systems of information dissemination were limited to one-on-one or mediated communication.

Mass communication systems, like broadcasting and print journalism, have fought to be subject to narrow controls.[111] Media freedom in Europe has come about from hard battles that have established a set of legal principles and rules, alongside general duties and obligations.[112] As a result, the media and the political class have a symbiotic relationship wherein the press might sit as a vital check on political power in one newspaper column and play the role of public relations conduit in another. The roles are enshrined in Convention Law and the European Union's Charter of Fundamental Rights and judgements of the European Court of Justice.[113]

The question that requires a clearer answer is what is an "offense" and when/why people do take offense. This may include some of the following considerations: unwarranted critique of/interference with sense of personal/collective identity; political opportunism; taking offence seriously: questions of principle/practice. Further, this begs three questions: first, where are appropriate limits to be drawn? Second, who should decide where these limits are? Third, is it legitimate ever to restrict freedom of expression to avoid the causing of offense to recipients of message?

Starting point to answering these key questions should be the assertion of the importance of free speech: "Freedom of expression constitutes one of the essential foundations of such a society, one of the basic conditions for its progress and for the development of every man. Subject to Article 10 (2), it is applicable not only to 'information' or 'ideas' that are favorably received or regarded as inoffensive or as a matter of indifference, but also to those that offend, shock or disturb the State or any sector of the population. Such are the demands of that pluralism, tolerance and broadmindedness without which there is no 'democratic society'."[114]

Will users be banned or have their speech curtailed for spreading fake news, just because they are misinformed or not well-educated? Is sharing a review of a violent Korean film or a death metal song with a bloody video 'violent harmful content'? Does a platform need to take down Quentin Tarantino film previews too? The right to offend, shock and disturb is part of free expression. Speech should never be judged on its subjective effects on a user, but carefully weighed against clearly defined public interest and other human rights. The vague nature of harms as a group that is not illegal could be challenged under principles of the rule of law, proportionality and legal certainty.[115] This also contravenes the longstanding principle from Handyside v UK:[116]

> Freedom of expression ... is applicable not only to "information" or "ideas" that are [favorably] received or regarded as inoffensive or as a matter of indifference, but also to those that offend, shock or disturb the State or any sector of the population.[117]

However, the Internet has brought a wholesale change in not only how information was disseminated, but also in respect of the actors doing the broadcasting. Users now generate most of the Internet's content, with Article 10 engaging everything from low-level speech to news commentary to search engines display results.[118] With a variety of technology available to hide user identities and the web's architecture empowering the user to not only speak without fear of social censure, but also automate and propagate their voice, the search for the *right* actor to regulate has frustrated regulators who have spent the last decade searching for a way to characterize platforms as publishers to attach a regulatory code of content for the "harms" associated with social media platforms.

As platforms have no general obligation to monitor content, the White Paper also fails to address how it intends to comply with the e-Commerce Directive and corresponding case law on platform liability.[119] The Online Harms White Paper claims, without explanation, that a 'duty of care' will somehow increase compliance:

> 'The new regulatory framework will increase the responsibility of

106 http://www.europarl.europa.eu/factsheets/en/sheet/7/the-principle-of-subsidiarity (last accessed 18 September 2019).
107 *Sunday Times v United Kingdom* (Application no. 6538/74), 26 April 1979
108 *Handyside v United Kingdom* (5493/72) at ¶ 49.
109 UK, Copyright, Designs and Patents Act 1988 c. 48.
110 UK, Defamation Act 2013, c. 26.
111 See e.g. F S Siebert (1952), *Freedom of the Press in England, 1476–1776: The Rise and Decline of Government Controls* (Urbana: University of Illinois Press) or J Rowbottom (2015), 'Entick and Carrington, the Propaganda Wars and Liberty of the Press' in A Tomkins and P Scott (eds), *Entick v Carrington: 250 Years of the Rule of Law* (Oxford, Hart).
112 J Rowbottom (2018), *Media Law* (Oxford, UK: Hart Publishing) 1-7.
113 Council of Europe, European Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by Protocols Nos. 11 and 14, 4 November 1950; *Sunday Times v United Kingdom* (Application no. 6538/74), 26 April 1979
114 *Handyside v United Kingdom* (5493/72) at ¶ 49.
115 UN, Declaration of the High-level Meeting of the General Assembly on the Rule of Law at the National and International Levels (2012), ¶ 8 http://www.unrol.org/article.aspx?article_id=192; EU, Charter of Fundamental Rights of the EU (2009), Article 49 (concerning the principles of legality and proportionality of criminal offences and penalties); European Convention on Human Rights, in particular 6(1), 7, 8(2), 9(2), 10(2) and 11(2).
116 ECtHR (1976) *Handyside v UK* (5493/72).
117 Ibid ¶ 49.
118 See e.g. *Google Spain, SL, Google Inc v Agencia Espanola de Proteccion de Datos* Case C-131/12 for search engine results; *Sunday Times v United Kingdom* (Application no. 6538/74), 26 April 1979 for news or *Mosley v United Kingdom* (Application no. 48009/08) 10 May 2011 for celebrity gossip as 'low level' speech.
119 EU, 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 ('Directive on electronic commerce'), article 15; *Scarlet Extended SA v Société belge des auteurs, compositeurs et éditeurs SCRL* (SABAM) Case C-70/10; or the ECtHR case *Delfi AS v. Estonia* (Application no. 64569/09).

online services in a way that is compatible with the EU's e-Commerce Directive, which limits their liability for illegal content until they have knowledge of its existence, and have failed to remove it from their services in good time'.[120]

However, 'intermediary liability' is not merely a creation of the e-commerce Directive. It may have been established by the e-Commerce Directive,[121] but it was implemented into UK law through the e-Commerce Regulations[122] and developed by subsequent case law in both the CJEU and UK Courts.[123]

This will have implications on the status of the law post brexit as well. Under Section 2 of the European Union (Withdrawal) Act 2018, the directives are not 'retained EU law', yet the domestic legislation that gives a directive effect in national law remains:

> 'Whereas other provisions of the e-Commerce Directive were implemented by the Electronic Commerce (EC Directive) Regulations 2002, article 15 was not specifically implemented through UK domestic legislation. Under section 2 of the European Union (Withdrawal) Act 2018 directives are not in themselves "retained EU law", only the domestic legislation made to implement them. However, under section 4 of the Act any prior obligations or restrictions of EU law which are "recognised and available in domestic law" will continue after Brexit. As article 15 has been recognised by domestic courts, including the Supreme Court in Cartier International AG and others v British Telecommunications Plc, it is likely to be considered retained law, **but uncertainty may remain until the matter is tested by the courts**'[124] [Emphasis Added]

The Directive provides a safe harbor for internet "hosts" (most of the companies the Government aims to regulate would fit into this category, including social media) and the protection from liability for illegal content stored on their platforms, provided that they do not have the *actual knowledge* about this content, and that they act promptly upon obtaining this knowledge.[125] The Directive prohibits the general monitoring of Internet users for the purpose of detecting such content.[126] There is extensive CJEU case law on the matter[127] as well as the related ECtHR jurisprudence on Articles 8 and 10 of the ECHR and the liability of Internet platforms.[128] The Government claims the new regime will be compatible with the Directive, but given the scope and requirements of duty of care, this is uncertain.[129] The prohibition of general monitoring will almost certainly have to be violated; it would be practically impossible to identify and remove all the potentially "harmful" content without monitoring activities of all the users on a certain platform.

6 Meet Ofweb, the UK's New Internet Overlord

The proposed model, placed on statutory footing, will be co-regulation with initial responsibility handed to Ofcom, the United Kingdom's broadcasting regulator, while a new regulator (provisionally called 'Ofweb') is established. [130]

The obvious questions that arises in a legal analysis is whether decisions made by Ofcom or Ofweb are subject to judicial review and what effects judicial review might have. In the UK, 'parliamentary sovereignty' ensures legislation cannot be reviewed by inferior courts; accordingly, this new regulatory system rooted in primary legislation cannot be subjected to judicial review. As the DCMS envisages that Ofcom's responsibilities will eventually be handed to Ofweb, judicial review can only have a role in most limited of circumstances; for example, codes of conduct and enforcement decisions. Furthermore, one can only raise proceedings in the UK on one of four grounds – (a) illegality, (b) procedural unfairness, (c) irrationality, or (d) incompatibility with human rights that are given effect by the Human Rights Act 1998. Therefore, the threshold for raising a judicial review against *any* specialist regulator in the UK is incredibly high. As a result of a failed petition for judicial review in *RT v Ofcom*[131], one commentator was prompted to note: "the judgment was an all-out win for Ofcom. It demonstrated yet again how difficult it is to succeed in a judicial challenge against the decision of a specialist regulator unless it has failed to comply with its own procedures or due process".[132]

Furthermore, judicial review should not be seen as a viable appeals process for decisions. In fact, one cannot apply to the courts if there is an open and valid appeals process that could have been followed, nor can judicial review overturn an earlier decision; it can only determine whether that decisions were illegal, improper, or irrational. Judicial review can only nullify the act (i.e. it never actually happened) rather than overturning it. A regulator like Ofweb would then be free to make the same exact decision having corrected for the earlier, for example, procedural error.

Admittingly, putting platform regulation on a co-regulatory framework has its benefits (e.g. stronger legitimacy than self-regulation, based on powers given by the Parliament, expertise, principle-based regulation, flexibility, cooperation with the industry),[133] yet there is a danger in Ofweb uncritically replicating the existing model of broadcast regulation, which has a very different historical rationale and justification, onto the Internet. Broadcast regulation affects entities who have access to scarce resources, such as spectrum,[134] who produce and distribute content at a large scale, and exercise editorial control with little user-created and/or generated content.[135] The Americans have also rejected this approach. In *ACLU v Reno*,[136] the US Supreme Court famously rejected the argument that regulating the Internet was

120 Ibid ¶ 41.
121 Directive 2000/31/EC.
122 Regulations 17, 18, and 19 of the E-Commerce (EC Directive) Regulations 2002 SI 2002/2013.
123 In the UK, for example, *Godfrey v Demon Internet Service* [1999] EWHC 244 (QB).
124 House of Lords Communications Committee Report (n 18) 185.
125 EU, 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 ('Directive on electronic commerce'), article 14.
126 Ibid, article 15.
127 For example, *Scarlet Extended SA v Societe Belge des Auteurs, Compositeurs et Editeurs SCRL* (SABAM) (C-70/10) [2011] E.C.R. I-11959 (24 November 2011).
128 *Delfi AS v. Estonia* (Application no. 64569/09); *Tamiz v the United Kingdom* (Application no. 3877/14) ECHR (12 October 2017); *Magyar Jeti Zrt v. Hungary* (Application no. 11257/16), 4 December 2018.
129 For example, *Tamiz v the United Kingdom* (Application no. 3877/14) [2017] ECHR (12 October 2017).
130 Online Harms White Paper (n 1) ¶ 5.15.
131 https://www.judiciary.uk/wp-content/uploads/2020/03/RT-v-Ofcom-approved-judgment-27.3.20.pdf.
132 https://smab.co.uk/first-court-decision-ofcom-impartiality.
133 See generally Christopher T. Marsden, 'Internet Co-Regulation and Constitutionalism: Towards a More Nuanced View' (August 29, 2011). Available at SSRN: https://ssrn.com/abstract=1973328 or Marsden (2011), Internet Co-Regulation: European Law, Regulatory Governance and Legitimacy in Cyberspace (Cambridge, UK: Cambridge University Press).
134 The radio spectrum is the part of the electromagnetic spectrum, widely used in modern technology, particularly in telecommunications and broadcasting. Examples of its use include TV, radio, mobile internet etc. see Wikipedia, Radio spectrum, https://en.wikipedia.org/wiki/Radio_spectrum
135 Rowbottom (n 112) 280-288.
136 *Reno v. American Civil Liberties Union*, 521 U.S. 844 (1997).

justified on grounds that broadcasting platforms are heavily regulated for content.[137] The Court stated that regulating content was permitted in broadcasting contexts because viewers had little control over what they were exposed to; however, users have to take a series of affirmative action to access the online content they want to see.[138] In addition to this, users also produce different types of content in ways unimaginable for broadcast.

In a case before the European Court of Human Rights, the Court recognized the difficulty applying broadcasting codes to Internet platforms:

> It is true that the Internet is an information and communication tool particularly distinct from the printed media, especially as regards the capacity to store and transmit information. The electronic network, serving billions of users worldwide, is not and potentially will never be subject to the same regulations and control.[139]

The Court also recognizes that the risk of harm online is different to that of broadcast and press media.

> The *risk of harm* posed by content and communications on the Internet to the exercise and enjoyment of human rights and freedoms, particularly the right to respect for private life, is certainly higher than that posed by the press. Therefore, the policies governing reproduction of material from the printed media and the Internet may differ. The latter undeniably have to be adjusted according to the technology's specific features in order to secure the protection and promotion of the rights and freedoms concerned.[140]

Note the Court's two concerns about the risk of harm by content and communications to first, the interference with other rights; and, second, that technology-specific features require adjustments to "secure the protection and promotion of the rights and freedoms concerned."[141]

Regulation over broadcasting has a much smaller impact than on individual speech. There are no scarce resources; however, there is peer-to-peer sharing, user-generated content, and individually created, but non-filtered speech. Furthermore, the press' self-regulatory model is a result of a long and exhausting struggle against historically regulated sectors like the press.[142] On the other hand, the Internet was founded on - and still largely embraces - the libertarian principle of openness.[143] Chapter Six of the House of Lords Communications Committee report suggests a new Parliamentary Joint Committee to ensure the regulator does not act on their own.[144] If implemented, there would be a tripartite regulatory relationship between 'Ofweb' (the regulator), the Government via the Cabinet Office and Parliament via a new Joint Committee. This is not independence; on the contrary, it is government using platforms as proxies to control the Internet.

Beyond the inappropriateness of using broadcasting's model to regulate the Internet, the intricacy of meeting "economically effective and socially just"[145] targets means that there is no uniformly accepted regulatory technique for digital technologies.[146] Rather, numerous possibilities exist within the categories of self-regulation, state regulation and "multi stakeholder co-regulation."[147] Each option carries advantages and disadvantages,[148] satisfying and undermining different notions of legitimacy,[149] such that implementation is fraught and "cynicism is at least partly justified."[150]

Asking a public authority to make specific rules can result in regulatory capture or a climate of resistance between the regulated and regulators and an impediment to higher performance.[151] Moreover, the strong intervention of public authority who represent the overall interests of the state may cause undue influence on the assessment through external factors. Both political and economic considerations can damage the advantages associated with the top-down model of platform regulation. There is also the risk of the most restrictive content laws becoming the norm across multiple platforms, regardless of the audience and user demographics. Schultz calls this the "slowest ship in the convoy problem"– the universal availability of information on the Internet might produce universal effects.[152] All platforms would have to comply with the most restrictive (i.e. the "slowest ship") standard.

7 Recommendations

Placing a duty of care on platforms for user-generated content that may cause harm will chill free expression and conflates the well-established common and statutory duty of care with clear duties and actual injuries. The least we could do is refer to the "duty" as "the duty to comply with existing regulation," or just maintain general terms of legal and regulatory obligations and duties. The government needs to reacquaint itself with the historical rationales for regulating broadcast (initially unregulated with increasing regulation amid scarce resources) or the press (initially heavily regulated with gradual deregulation and strengthening of media freedom).[153]

Whilst it is clear that there are problems with online platforms, their power and different harms that arise as a consequence,[154] most of the identified harms in the White Paper could be remedied with proper co-regulation of actors operating online, and enhanced obligations to cooperate with law enforcement over a variety of existing forms of

137 Ibid at 845, 870.
138 Idem at 854.
139 ECtHR, Judgment of 5 May 2011, *Case of Editorial Board of Pravoye Delo and Shtekel v Ukraine,* (Application No. 33014/05) at ¶ ;See also Judgment of 16 June 2015, Case of *Delfi AS v Estonia* (Application no. 64569/09).
140 Editorial Board at ¶ 36 (emphasis added).
141 Editorial Board at ¶ 63.
142 Rowbottom (n 111).
143 John Perry Barlow, A Declaration of the Independence of Cyberspace, available at: https://www.eff.org/cyberspace-independence.
144 House of Lords Communications Committee (n 18) chapter 6.
145 Brown, I., & Marsden, C. T. (2013). *Regulating code: Good governance and better regulation in the information age.* (Boston: MIT Press) ix; See also Orla Lynskey (2015), *The Foundations of EU Data Protection Law* (Oxford: OUP) at Page 47.
146 Brown & Marsden (n 145) 1; See also Terry Flew (2018), 'Technology and Trust: The Challenge of Regulating Digital Platforms' (Korean Association for Broadcasting and Telecommunications Studies) 9-11.
147 Ibid at Page 2.
148 Ibid at Page 2-3.
149 Black, J. (2008). Constructing and contesting legitimacy and accountability in polycentric regulatory regimes. *Regulation & governance*, 2(2), 137-164 at Page 145.
150 Brown & Marsden (n 145) 3.
151 Baldwin R, Cave M, Lodge M. (2012) *Understanding regulation: theory, strategy, and practice*. (Oxford University Press on Demand)108-110.
152 Schultz, T. (2008). Carving up the Internet: jurisdiction, legal orders, and the private/public international law interface. *European Journal of International Law*, 19(4), 799-839 at 813 citing Zittrain, ' Be Careful What You Ask For: Reconciling a Global Internet and Local Law', in A. *Thierer and C.W. Crews (eds) (2013), Who Rules the Net? Internet Governance and Jurisdiction* (Cato Institute) 17.
153 Rowbottom (n 153) 2 -5, 256 – 288.
154 As the EU recognises and addresses in the ongoing attempt to reform platform liability, inter alia. See e.g. European Commission, 'Shaping Europe's digital future' (COM (2020)0067), 19 February 2020.

criminal speech and behavior. Because of the extent of the impact of regulation on the digital rights of users, judicial oversight is crucial, and any regulator should be independent with pathways for judicial remedies and reviews. Furthermore, it is insufficient to base platform regulation on a handful of user submissions and surveys. Although Ofcom's annual survey is widely cited throughout, albeit quite selectively, any additional harms subject to further regulation need to be based on clear and unambiguous evidence.[155] As our analysis shows, the concept of "online harms" is vague and it should be dropped entirely. Any additional harm criminalized in the future needs to be clearly defined, well evidenced and regulated in the public interest.

Additionally, it is suggested that companies do not rely on technology solely, but human oversight should also be a requirement wherever there are takedown procedures in place. Automated systems and AI are not reliable enough to be used alone, as we have seen in the case of the YouTube Content ID system[156] and the likelihood of errors.[157] Speech assessment includes qualitative questions on whether content should be treated differently to information offline for every individual user (*the parity principle*, for example).[158] For this to happen, the platform will need to understand the context of exchanges between every user on a platform and how people communicate offline with one another.[159] Different platforms have different social norms and communication practices and this should be respected (e.g. it is not realistic to expect the same language on 4Chan, Reddit, and Mumsnet).

Using technology to search for fake news is potentially problematic with false positives potential affecting media pluralism.[160] In *Jersild v Denmark*,[161] the court stated that "the methods of objective and balanced reporting may vary considerably, depending among other things on the media in question."[162] Second, Article 10 ECHR "protects not only the substance of the ideas and information expressed, but also the *form* in which they are conveyed."[163] The observation that "the methods of objective and balanced reporting may vary considerably"[164] takes on increased importance in contemporary times. In the current "post-truth" era, fake news, misinformation and disinformation are widely generated and disseminated by a range of actors (and algorithmic techniques) and they compete fiercely with one another for the public's attention and acceptance.

More generally, the White Paper also lacks the clarity necessary in law. The regulatory framework must be accessible: users "must be given an indication that is adequate in the circumstances of the legal rules applicable to a given case."[165] Secondly, users must be able to moderate their behavior in line with what is reasonably foreseeable.[166] Users "must be able - if need be with appropriate advice - to foresee, to a degree that is reasonable in the circumstances, the consequences which a given action may entail."[167] The White Paper's proposed framework is vague and insufficient and lacks the clarity to "give individuals an adequate indication as to the circumstances in which and the conditions on which the authorities are empowered to resort" to any such measures.[168]

The White Paper falls short by not properly considering alternatives to its proposed measures. We point out some alternatives below with the purpose of demonstrating that alternative recommendations would be conceivable, rather than attempting to develop these fully.

1. **Reform Intermediary liability**. An alternative way to respond to the "online harms" identified in the White Paper is reformation of the liability provisions of the e-Commerce Directive, in line with the Regulations adopted in the last mandate of the European Commission. The general principle of a harmonized, graduated, and conditional exemption continues to be needed as a foundational principle of the Internet. The principle, however, needs to be updated and reinforced to reflect the nature of the services in use today. This could mean that the notions of mere conduit, caching and hosting service could be expanded to explicitly include other services. In some instances, this can amount to codifying existing case law (e.g. for search engines or Wi-Fi hotspots), while in other cases a clarification of its application to collaborative economy services, cloud services, content delivery networks, domain name services, etc. is necessary. Building on concepts like editorial responsibility[169], actual knowledge[170] and degree of control[171], the concept of active/passive hosts should be replaced by more appropriate concepts that reflect the technical reality of today's services.[172]

2. **General monitoring and automated filtering**. While the prohibition of general monitoring obligations should be maintained as another foundational cornerstone of Internet regulation, specific

155 For a good example of how evidence submitted to the Committee has been bastardized to make a political point, see Goldman, Eric, The U.K. Online Harms White Paper and the Internet's Cable-ized Future (2019). *Ohio State Tech. L.J.*, Forthcoming. Available at SSRN: https://ssrn.com/abstract=3438530 at Page 2.
156 YouTube, 'How Content ID Works' https://support.google.com/youtube/answer/2797370?hl=en; J Bailey, 'YouTube Beta Testing Content ID For Everyone' (*Plagiarism Today*, 2 May 2018) https://www.plagiarismtoday.com/2018/05/02/youtube-beta-testing-content-id-for-everyone/.
157 J M Urban, J Karaganis, and B Schofield,'Notice and Takedown in Everyday Practice', *UC Berkeley Public Law Research Paper* No. 2755628. http://dx.doi.org/10.2139/ssrn.2755628.
158 Online Harms White Paper (n 1)
159 This is virtually impossible, see Banerjee, S., Chua, A. Y., & Kim, J. J. (2017). Don't be deceived: Using linguistic analysis to learn how to discern online review authenticity. *Journal of the Association for Information Science and Technology*, 68(6), 1525-1538.
160 Heins, M., & Beckles, T. (2005). *Will fair use survive? Free expression in the age of copyright control.* Marjorie Heins.
161 *Jersild v Denmark*, ECtHR 23 September 1994 (GC), ECLI:CE:ECHR:1994:0923JUD001589089, Series A no. 298.
162 Ibid ¶ 31.
163 ECtHR, *Autronic AG v Switzerland* (1990) 12 EHRR 485, [47]
164 *Bladet Tromso and Stensaas v Norway*, 20 May 1999, [59].
165 ECtHR The Sunday Times v. the United Kingdom (No. 1), 6538/74, 26 April 1979 at ¶ 49.
166 *Rekvényi v Hungary*, 25390/94, 20 May 1999, At ¶ 34f.
167 ECtHR *The Sunday Times v. the United Kingdom* (No. 1), 6538/74, 26 April 1979 at ¶ 49.
168 Ibid, at at ¶ 49; See also *Malone v United Kingdom* Application no. 8691/79, 2 August 1984 at [67].
169 For example, See The New Definitions Of "Audiovisual Media Service" (Article (1)(a)(i)) and "Video-Sharing Platform Service" (Article (1)(b)(aa) and Article 1(D)(Bb) of Directive (EU) 2018/1808 of the European Parliament and of the Council of 14 November 2018 amending Directive 2010/13/EU on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive) in view of changing market realities.
170 Note 123, supra.
171 *L'Oreal v eBay* Case C324/09, 12 July 2011 at ¶ 116, 123, 145; See also Article 14(2), e-Commerce Directive
172 For some other approaches see, for instance: C. Angelopoulos C. and S. Smet, 'Notice-and-fair-balance: how to reach a compromise between fundamental rights in European intermediary liability', (2016) *Journal of Media Law*, 8(2); S. Stalla-Bourdillon (2017), 'Internet Intermediaries as Responsible Actors? Why It Is Time to Rethink the Ecommerce Directive as Well', in Taddeo M., Floridi L. (eds), *The Responsibilities of Online Service Providers.* Law, Governance and Technology Series, vol 31. (Heidelberg etc: Springer).

provisions governing algorithms for automated filtering technologies - where these are used - should be considered, to provide the necessary transparency and accountability of automated content moderation systems.

3. **Regulating content moderation**. Uniform rules for the removal of illegal content like illegal hate speech should be made binding across the EU. Replacing notice-and-takedown with notice-and-action rules could be tailored to the types of services, e.g. whether the service is a social network, a mere conduit, or a collaborative economy service, and where necessary to the types of content in question, while maintaining the maximum simplicity of rules. The feasibility of introducing thresholds could be examined in this context, taking due account of the size and nature of the service provider and of the nature of the potential obligations to be imposed on them. Building on the Recommendation on Illegal Content,[173] binding transparency obligations would also be at the heart of a more effective accountability framework for content moderation at scale and would complement recently adopted rules under the Audiovisual Media Services Directive[174] or the modernization of the EU copyright rules.[175] Increasing transparency for algorithmic recommendation systems of public relevance like social media news feeds should be examined. At the same time, these rules should prohibit allowing Member States to impose parallel transparency obligations at national level, providing for a simple set of rules that comply with the Manila principles[176] on content moderation and intermediary liability in the European Union.

8 Conclusions: Broad and Flawed

The Internet is not a "safe space," nor was it intended to be. Without a doubt, the Internet is a complicated space; however, it also makes us look at humankind's most unsavory characteristics in a way never imagined before. Accordingly, we should look at platforms as a blessing, not a burden. How else could we know that so many people think like us at the same time as hold such divergent, even abhorrent views? Yet the White Paper goes beyond turning the Internet into a virtual soft play area where everyone has to watch what they say, what they do, and how they act. It burdens the platform with a duty of care to police the speech of its patrons, under the threat of sanctions for what might be offensive or intimidating and might cause harm. This is a prime example of the "chilling effects" of content moderation. Furthermore, platforms have undertaken significant self-regulatory responses to mitigate the threat of co-regulation. For example, Facebook launched an Independent Oversight Board and charter for content moderation on its site.[177]

In what feels like ancient history, the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, David Kaye, warned in 2011 against the adverse effects that disproportionate regulation of content might have on free speech.[178] To address this, he recommends:

> States should only seek to restrict content pursuant to an order by an independent and impartial judicial authority, and in accordance with due process and standards of legality, necessity and legitimacy. States should refrain from imposing disproportionate sanctions, whether heavy fines or imprisonment, on Internet intermediaries, given their significant chilling effect on freedom of expression.[179]

Furthermore, there is little evidence that the data published in the Online Harms White Paper or the House of Lords Communications Committee will change anyone's opinion or behavior. Opinion on platform regulation will always remain divided along deeply held beliefs about the constitutional merits of criminal prohibitions in areas like hate speech and the role of private actors in content regulation. Rather than assuming every view different from our own results in harm and place a burden on online services to remove them, we need to develop techniques and strategies for defeating ideology through competition in the marketplace of ideas. With so many special interests competing with each other for the attention of lawmakers, each with their own agenda in protecting identifiable stakeholders, regulation should be forward-thinking and dynamic, and protect the principles of free expression and media pluralism, rather than take action to inhibit and control. In hindsight, it is quite surprising that there was not more emphasis on enhanced cooperation between platforms and law enforcement.

The Internet is for expression – it is for argument, emotion, anger, purchasing, love, sex, and sharing. Expression is its bread and butter. All of the above, of course, come with negative consequences. Arguments can turn into violence, emotions can run high and lead to regret, anger can cause permanent damage, love can turn to heartbreak, sex can lead to objectification and pain, and sharing can be a violation of someone else's rights. We are already well-equipped to deal with this through different forms of online offences such as harassment, revenge porn and other communication offences. Some of these, as suggested by the Law Commission, should be reviewed and consolidated,[180] but this will be dealt with through criminal law reforms, and not the vaguely imposed duty of care that threatens fundamental rights online. Behind all of this is the harm associated with regulatory capture, the protectionist mindset of 'something must be done,' and the problem of regulating the wrong actors.

Imposing a duty of care on platforms inadvertently creates a framework for crushing dissent, plurality, diversity, "British values,"[181] and

173 Commission Recommendation on measures to effectively tackle illegal content online, Commission Recommendation of 1.3.2018 on measures to effectively tackle illegal content online (C(2018) 1177 final), Available at https://ec.europa.eu/digital-single-market/en/news/commission-recommendation-measures-effectively-tackle-illegal-content-online, (last accessed 19 September 2019).

174 Directive 2010/13/EU of the European Parliament and of the Council of 10 March 2010 on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services.

175 Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC

176 Manila Principles on Intermediary Liability Best Practices Guidelines for Limiting Intermediary Liability for Content to Promote Freedom of Expression and Innovation, Available at https://ec.europa.eu/digital-single-market/en/news/commission-recommendation-measures-effectively-tackle-illegal-content-online (last accessed 19 September 2019).

177 Establishing Structure and Governance for an Independent Oversight Board, Available at https://newsroom.fb.com/news/2019/09/oversight-board-structure (last accessed 19 September 2019).

178 HRC, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 06 April 2019, available at: https://documents-dds- ny.un.org/doc/UNDOC/GEN/G18/096/72/PDF/G1809672.pdf

179 Ibid 19.

180 Note 76, Supra 328- 334.

181 UK, Department for Education, Promoting fundamental British values as part of SMSC in schools, Departmental advice for maintained schools, November 2014, at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/380595/SMSC_Guidance_

ultimately, free speech. Populism, is, by its very definition,[182] wedded to the preservation of the status quo. That said, it is surely a wonderful thing that, for all its faults, there is at least one remaining space in our culture where words still matter and where promises made in the form of written undertakings ("laws") have consequences. However, for the Internet, the trick is getting the law right. A society that stops being governed by the authority and rule of law and reverts to that of the "populist," the priest, or "the people" is not a place where freedom, openness and democracy will long survive. It seems a long way from Frank La Rue, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 2011 Report to the UN[183] when he stated:

> The Special Rapporteur believes that censorship measures should never be delegated to a private entity, and that no one should be held liable for content on the Internet of which they are not the author. Indeed, no State should use or force intermediaries to undertake censorship on its behalf.[184] Of concern, Subject to abuse by state and private entities; Risk of liability causes intermediary to err on the side of taking content down; Lack of transparency on decision making practices obscures discriminatory practices or political pressure affecting their decisions; and companies shouldn't be making the assessment of legality of content.[185]

More recently, Catalina Botero Marino strongly endorsed transparency in her 2013 report, stating:

> [w]ith respect to the duty of transparency, intermediaries should have sufficient protection to disclose the requests received from government agencies or other legally authorized actors who infringe upon users' rights to freedom of expression or privacy. It is good practice, in this respect, for companies to regularly publish transparency reports in which they disclose at least the number and type of the request that could lead to the restrictions to users' rights to freedom of expression or privacy.[186]

A flat-earther that has been called an idiot or an imbecile could have a claim of "abuse" and/or intimidation. Empowering users might be a noble objective, but that requires empowering the *right* users and educating everyone.

Maintained_Schools.pdf.

[182] David Molloy, What is populism, and what does the term actually mean?, (*BBC News*, 6 March 2018), https://www.bbc.co.uk/news/world-43301423.

[183] Frank La Rue, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression Human Rights Council, Seventeenth session, Agenda item 3, Promotion and protection of all human rights, civil, political, economic, social and cultural rights, including the right to development, Available at https://www2.ohchr.org/english/bodies/hrcouncil/docs/17session/A.HRC.17.27_en.pdf (last accessed 11 September 2019).

[184] Ibid ¶ 43.

[185] Ibid ¶ 42.

[186] IACHR Office of the Special Rapporteur for Freedom of Expression (OSRFE), Freedom of Expression and the Internet, (Dec. 31, 2013), Available at https://www.oas.org/en/iachr/expression/docs/reports/2014_04_08_internet_eng%20_web.pdf ¶ 113.

Copyright (c) 2020 Mark Leiser, Edina Harbinja

Creative Commons License

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Technology and Regulation (TechReg) is an open access journal which means that all content is freely available without charge to the user or his or her institution. Users are permitted to read, download, copy, distribute, print, search, or link to the full texts of the articles, or to use them for any other lawful purpose, without asking prior permission from the publisher or the author. Submissions are published under a Creative Commons BY-NC-ND license.