Do we need a new Duty of Care?

1. Legislating unkindness

Based on the House of Lords' Report ("HOLR") at paragraph [182], the issues of concern are lawful but harmful or anti-social behaviour. That report cites the following: bullying, abuse, distressing/ disturbing content,[1] political misinformation. See Table 3 of the same. The Executive Summary of the White Paper on Online Harms refers to illegal and "unacceptable content."[2] And "[o]*ther online behaviours or content, even if they may not be illegal in all circumstances, can also cause serious harm. The internet can be used to harass, bully or intimidate, especially people in vulnerable groups or in public life. Young adults or children may be exposed to harmful content that relates, for example, to self-harm or suicide…*"

It is simply not possible or advisable to prohibit by law, rudeness, unkindness or stupidity in writing. If it was, our forefathers would surely have done so, in times when community standards were far higher. In so far as it is possible this is within the scope of the Protection from Harassment Act and other legislation, see below. [3]

In relation to falsity or misinformation, we note also that we know from defamation law that what is fact and what is opinion is such a difficult distinction that even appellate judges regularly err. It is not acceptable to require proof of opinions and we note in *Salov v Ukraine*, the ECtHR found a violation of the right to freedom of expression although the case involved the spread of false information. The Court affirmed that 'article 10 of the [ECHR] as such

[1] This includes information that is not appropriate for children as promoting violence, self-harm or bullying – and is anti-social, as indecent, misleading or profane.

[2] https://www.gov.uk/government/consultations/online-harms-white-paper/online-harms-white-paper-executive-summary--2

[3] If we try to apply the proposed approach offline—we may for example hold the bus-driver or his employer liable for the behavior of the school children on the bus home. If they behave badly and there are hurt feelings or tears—would we hold the driver responsible? Similarly, is the school liable for the taunts in the playground? Should the telephone company listen to all the calls to make sure no one says something anti-social and bleep it out? This is foolishness. This conduct cannot be legislated for and should not be. The reaction of the victim depends on subjective feelings and this is not a basis on which behavior can be rendered unlawful. See below re Freedom of Expression.

does not prohibit discussion or dissemination of information received even if it is strongly suspected that this information might not be truthful' and found that the freedom of expression had been violated due to the fact that the penalty – a prison sentence – was excessive.

2. Newsagents of the internet

There is a desire to see platforms co-regulated in a similar way to other professional media, the print press and the broadcasters by the introduction of codes of conduct and the imposition of a duty to adhere to the code and provide a remedy. Which model is appropriate? The Ofcom Code has harm and offence and taste and decency rules plus a requirement of due impartiality in news and independence (balanced reporting of both sides) in matters of current affairs and political and industrial controversy. That is all very well for an organization that creates its own content professionally or distributes other professionally created content but is it appropriate where all is third party content? There is a precedent for this in relation to print media. Traditionally, the print press was held liable for letters to the Editor that they published. This was due to the fact that they edited and published them and so were a primary publisher (author, editor, publisher) and so the defence of secondary responsibility (an intermediary defence for distributors, printers and deliverymen etc) in section 1 of the Defamation Act 1996 was not made available for "commercial publishers." This same approach was taken by the court in *Delfi v Estonia* (App. No. 64569). The question is how appropriate it is to make platforms –akin to newsagents or distributors liable for the content of the newspaper including the readers' letters section? That is what is proposed. This is not practicable or feasible.

3. Not the right model

We note that much of the same terrain is covered by the European Commission's proposed new regulation on preventing the dissemination of terrorist content online. This proposal notably includes the following provisions:

- The one-hour rule: the Commission is proposing a legally binding one-hour deadline for content to be removed *following a removal order from national competent authorities;*
- A definition of terrorist content as material that incites or advocates committing terrorist offences, promotes the activities of a terrorist group or provides instruction in techniques for committing terrorist offences;
- *A duty of care obligation for all platforms to ensure they are not misused for the dissemination of terrorist content online. Depending on the risk of terrorist content being disseminated via their platforms, service providers will also be required to*

take proactive measures – such as the use of new tools – to better protect their platforms and their users from terrorist abuse;

- A framework for strengthened co-operation between hosting service providers, Member States and Europol. Service providers and Member States will be required to designate points of contact reachable 24/7 to facilitate the follow up to removal orders and referrals;
- Content providers will be able to rely on effective complaint mechanisms that all service providers will have to put in place. *Where content has been removed unjustifiably, the service provider will be required to reinstate it as soon as possible*. Effective judicial remedies will also be provided by national authorities and platforms and content providers will have the right to challenge a removal order. For platforms making use of automated detection tools, human oversight and verification should be in place to prevent erroneous removals;
- Transparency and accountability will be guaranteed with annual transparency reports and Member States will have to put in place financial penalties for non compliance with removal orders, which may go up to 4% of the global turnover of the last business year of the target companies.

Is it an appropriate model by which to regulate all content of all kinds online? Absolutely not. The proposed regulation would be a vast improvement on the IWF process currently in place as it anticipates due process and put back options however.[4] We understand that the current state of the proposed regulation is that on 6 February 2019, the European Parliament requested an opinion from the EU Fundamental Rights Agency (FRA) on the proposal.[5] We also note the following from the Proposal at page 3:

[4] On a very practical level –we know that the UK government has no problems whatsoever having material removed from the web and that they have ample assistance from Nominet and the IWF and there is no obstacle in the nature of due (or any) process for such removals (which are actioned merely on a phone call or email from Law Enforcement by Nominet). Note the UNHRC report (below) says "58. *Notice and appeal. Users and civil society experts commonly express concern about the limited information available to those subject to content removal or account suspension or deactivation, or those reporting abuse such as misogynistic harassment and doxing. The lack of information creates an environment of secretive norms, inconsistent with the standards of clarity, specificity and predictability. This interferes with the individual's ability to challenge content actions or follow up on content-related complaints;"*

[5] FRA considered that the definition of terrorist content has to be modified, as broadens the terms on the directive on combating terrorism (2017/541). The definition of the content was considered too wide by FRA and as such would interfere with the freedom of expression and information. FRA called for the proposal to better protect journalistic, academic and artistic expression and to guarantee some type of judicial involvement and in cases of cross-border removal orders, the involvement of the host Member State's judiciary. Online providers must receive sufficient notice/information. Etc. Amendments to the draft report were due on 15 February 2019.

"*The present proposal is consistent with the acquis…and in particular the E-Commerce Directive. Notably, any measures taken by the hosting service provider in compliance with this Regulation, including any proactive measures, should not in themselves lead to that service provider losing the benefit of the liability exemption provided for, under certain conditions, in Article 14 of the E-Commerce Directive. A decision by national authorities to impose proportionate and specific proactive measures should not, in principle, lead to the imposition of a general obligation to monitor, as defined in Article 15(1) of Directive 2000/31/EC towards Member States. However, given the particularly grave risks associated with the dissemination of terrorist content, the decisions under this Regulation may* ***exceptionally derogate from this principle*** *under an EU framework. Before adopting such decisions, the competent authority should strike a fair balance between public security needs and the affected interests and fundamental rights including in particular the freedom of expression and information, freedom to conduct a business, protection of personal data and privacy. Hosting service providers' duties of care should reflect and respect this balance which is expressed in the E-Commerce Directive."*

So again, is it an appropriate model by which to regulate all content of all kinds online? Absolutely not. It is *exceptional* and should remain so. Further note at page 5: *"The proposal takes into account the burden on hosting service providers and safeguards, including the protection of freedom of expression and information as well as other fundamental rights.* ***The one-hour timeframe for removal only applies to removal orders, for which competent authorities have determined illegality in a decision which is subject to judicial review.*** *For referrals, there is an obligation to put in place measures to facilitate the expeditious assessment of terrorist content, without however imposing obligations to remove it, nor within absolute deadlines. The final decision remains a voluntary decision by the hosting service provider."* Indeed, the scheme of the regulation is for implementation of removal decisions made by LEA.

4. The Devil is in the detail

Let us see the draft duty and its scope. No doubt it will be unacceptably broad and totally unworkable (see above) and all the time spent discussing it in theory will have been wasted.
The EU Proposal for the regulation (above) gives us this example:

"(*12) Hosting service providers should apply certain duties of care, in order to prevent the dissemination of terrorist content on their services.* ***These duties of care should not amount to a general monitoring obligation.*** *Duties of care should include that, when applying this Regulation, hosting services providers act in a diligent, proportionate and non-discriminatory manner in respect of content that they store, in particular when implementing*

their own terms and conditions, with a view to avoiding removal of content which is not terrorist. ***The removal or disabling of access has to be undertaken in the observance of freedom of expression and information.***

Article 3 Duties of care
1. Hosting service providers shall take appropriate, reasonable and proportionate actions in accordance with this Regulation, against the dissemination of terrorist content and to protect users from terrorist content. In doing so, they shall act in a diligent, proportionate and non-discriminatory manner, and with due regard to the fundamental rights of the users and take into account the fundamental importance of the freedom of expression and information in an open and democratic society.
2. Hosting service providers shall include in their terms and conditions, and apply, provisions to prevent the dissemination of terrorist content."

Note the definition:

(5) 'terrorist content' means one or more of the following information:
(a) inciting or advocating, including by glorifying, the commission of terrorist offences, thereby causing a danger that such acts be committed;
(b) encouraging the contribution to terrorist offences;
(c) promoting the activities of a terrorist group, in particular by encouraging the participation in or support to a terrorist group within the meaning of Article 2(3) of Directive (EU) 2017/541;
EN 24 EN
(d) instructing on methods or techniques for the purpose of committing terrorist offences."

At least there is a definition and it is of conduct that is grounded in what is *unlawful.* Further, again, the whole scheme of the regulation is for a duty to comply promptly with removal orders and referrals—not to be responsible for monitoring (and editing all content on their services). The duty of care language itself exhorts undertaking a complex legal balancing act between taking appropriate and proportionate actions and observing freedom of expression etc. This will be near impossible for these private actors and we can expect over blocking as a result. Removal on the side of caution is incentivised, see below.

5. New liability and regulation can be unbundled

It is perfectly possible to regulate and so "Gold Plate" against existing laws without introducing new laws or statutory duties. See below.

6. Solving problems that are not the problem

We have plenty of laws. We don't need new law or duties. The real problems online are: jurisdiction, criminal enforcement and civil access to justice/remedy. The new duty fails to address them. While there may be a desire to regulate, no new law is needed for that.

The UK has laws dealing with defamation, intellectual property, privacy, confidence and data protection, harassment (criminal and civil), revenge porn, obscene publications, malicious publications and communications. We also have the Public Order Act (Incitement to Racial Hatred), the Race Relations Act and the Racial and Religious Hatred Act, the Terrorism Act, the Official Secrets Act, the law of Contempt, the Sexual Offences Act and the Protection of Children Act and the Equality Act and laws governing advertising and commercial communications and speech during elections. The law is currently adequate and there is no pressing or obvious need for additional legislation. We also have ASBOs and CPNs and CBOs. These are not being used. Arguably the real problem is LEA and cuts as a result of the long austerity. The Police are not able to deal with online criminality in anything other than the most serious cases.

Crime is increasingly online but is not adequately investigated or enforced despite recent revised and new guidelines for prosecuting offence online. See https://www.cps.gov.uk/legal-guidance/social-media-guidelines-prosecuting-cases-involving-communications-sent-social-media .[6]

[6] *"Part A: Offences – Substantive. 6.Where social media is used to facilitate a substantive offence, prosecutors should proceed under the substantive offence in question, having regard as appropriate to the Hate Crime and VAWG sections below. 7. The following are potential offences against the person, against public justice or sexual offences, with links to the relevant guidance, which prosecutors may consider: Making a threat to kill, contrary to section 16 Offences Against the Person Act 1861. Making a threat to commit criminal damage, contrary to section 2 Criminal Damage Act 1971. Harassment or stalking, contrary to sections 2, 2A, 4 or 4A Protection from Harassment Act 1997. Controlling or coercive behaviour, contrary to section 76 Serious Crime Act 2015. Blackmail, contrary to section 21 Theft Act 1968. Juror misconduct, contrary to sections 20A-G Juries Act 1974.* Contempt of court, contrary to the Contempt of Court Act 1981.* Publishing material which may lead to the identification of a complainant of a sexual offence, contrary to section 5 Sexual Offences (Amendment) Act 1992.* Intimidating a witness or juror, contrary to section 51 Criminal Justice and Public Order Act 1994. Breach of automatic or discretionary reporting restrictions, contrary to section 49 Children and Young Persons Act 1933 and section 45 Youth Justice and Criminal Evidence Act 1999. Breach of a restraining order, contrary to section 5 Protection from Harassment Act 1997. * Note that these offences require Attorney General's consent to prosecute and should be referred to the DLA Team prior to any such submission. Disclosing private sexual images without consent ("revenge pornography"), contrary to section 33 Criminal Justice and Courts Act 2015. Causing sexual activity without consent, or causing or inciting a child to engage in sexual activity, or sexual communication with a child contrary to sections 4, 8, 13, 15A Sexual Offences Act 2003. Taking, distribution, possessing or publishing indecent photographs of a child, contrary to section 1 Protection of Children Act 1978 Allegations contrary to Part III Public Order Act 1986 should be referred to Special Crime and Counter Terrorism Division. 8. The act of setting up a false social networking*

It is possible that some of the demand comes from the desire to replicate the new German law. However, that jurisdiction may have had a lacuna –we do not. If there is a UK lacuna, then we need to know very precisely what new conduct now falls outside the ambit of the law? See 1 above etc.

7. Ignoring years of jurisprudence that painstakingly and carefully tailored tort duties for speech

The law of defamation, malicious falsehood and confidence for example has been hundreds of years in the making and is carefully adapted in a very precise ways to protect truthful speech and the public interest in it –and protecting and safeguarding opinion –as required for the marketplace of ideas and to the benefit of society. There are thresholds to prevent trivial actions and *mere vulgar abuse* is excluded as it does not cause lasting harm. These can be adjusted but the beauty of the common law on defamation and the reason it has stood the test of time is that discretion and flexibility enable it to function as living law which adapts to the times—so that the test of what is defamatory shifts with societal ideas about what is harmful to reputation. For example, it is no longer defamatory to allege homosexuality (unless the claimant is married, when the allegation become hypocrisy). Why would we throw this over and start from scratch? Can or should mere rudeness or abuse be rendered unlawful civilly or criminalised? That would bring our creaking justice system to a halt. The thresholds for serious harm and real and proper torts having recently been raised by the government in defamation –a speech related torts. That plus the increase in litigants in person, means handing out new causes of action should be very carefully considered. Such laws would also be extremely difficult to draft and would catch much ordinary human conduct.

8. New law will take a long time to be litigated into shape

---

account or website, or the creation of a false or offensive profile or alias could amount to a criminal offence, depending on the circumstance For example: The former estranged partner of a victim creates a profile of the victim on a Facebook page, to attack the character of the victim, and the profile includes material that is grossly offensive, false, menacing or obscene. A "photoshopped" (digitally edited) image of a person is created and posted on a social media platform. Although many photoshopped images are humorous and inoffensive, others are disturbing or sinister, such as the merging of a person's face with the nude body of another to create obscene images, which may be accompanied by offensive comment. 9. Depending on the circumstances, this may be a way in which one of the offences above are committed, or it may be a way in which a communications offence (as these include "false" communications or messages) is committed. Part A: Offences – Communications Offences. 10. Where social media is not being used to commit another substantive offence, prosecutors may turn to consider the communications offences available."

Quite aside from the challenges from NGOs, a new duty of care will take decades to refine and define. Why not adjust whatever law is the problem? One answer is that what is wanted is not at all appropriate for speech and would not be article 10 compliant. This duty of care is a fudge against that.

9. Discriminatory

Good laws are behaviour focused and not actor focused. Justice is blind. It is entirely discriminatory to target some and not others. This fails to be an appropriate model for speech and some speakers are not better than others. Note the rise for example in citizen journalists and the cases extending the same laws to them as enjoyed by traditional media.

It is also objectionable for the state to pander to the incumbents in our print press and entrench its status, raising barriers to entry (the new art 11 press publishers ancillary right), while discriminating against new entrants with onerous legal obligations and potentially enormous liabilities.

10. Race to the bottom

In tort you must take your victim as you find him, this is the *egg shell skull* principle. This means that the providers of services caught are going to have to tailor their behaviour to the most vulnerable with the resulting unlawful and disproportionate interference to the adult users' Freedom of Expression protected rights to receive and impart information. This is not proportionate and will be challenged.

11. Tyranny by majority

There is an obvious risk that a blunt instrument like this will be used to supress or restrict speech that is unpopular or dissenting or a minority view. Freedom of Expression protects the right to offend --not just views that are widely held or accepted. See *Handyside v UK* (5493/72), *Sunday Times v UK* (No.2) [1992] 14 EHRR 123 (Spycatcher) (at [50] "*Freedom of expression constitutes one of the essential foundations of a democratic society ..it is applicable to not only to "information" or "ideas" that are favourably received or regarded as inoffensive or as a matter of indifference but also to those that offend, shock or disturb..*").

12. AVMSD II?

The duty of care is unnecessary and need not underpin any regulation or co or self-regulatory code. We note that AVMSD II imposes an enhanced duty to protect minors and against hate and incitement **and extends to platforms with audio-visual content.** Note that AVMSD II encourages co-regulation and sets out what protection would look like –flagging, reporting, age verification, rating systems, terms, shielding minors from certain conduct, encryption and parental controls for the most harmful content. New procedures are proposed for public heath, security and terrorism. This Directive covers similar ground but in a precise and

focused way as a regulatory obligation applicable to the providers of regulated audio-visual media services. Further this dovetails with the proposal to have Ofcom co-regulate (at least the Audio-Visual content providers –whoever they are). The balance should comply with the Editor's Code (used by both IPSO or IMPRESS). This seems to be what the AVMSD contemplates—the regulation follows the nature of the content.

13. Contrary to Human Rights Convention obligations

The proposal incentivizes the suppression of users' rights to Freedom of Expression, to impart and receive information, by private actors. See the UN Special Rapporter on Promotion and Protection of the Right to Freedom of Opinion and Expression (2018) Report to UNHRC on Human Rights Approach to Platform Content Regulation, ARC/38/35 at paragraph [5] noting that the activities of companies in the ICT sector implicate rights to privacy, religious freedom and belief, opinion and expression, assembly and association, and public participation, among others. See [13] "*Broadly worded restrictive laws on "extremism", blasphemy, defamation, "offensive" speech, "false news" and "propaganda" often serve as pretexts for demanding that companies suppress legitimate discourse.."*

Recall that art. 19 of the International Covenant on Civil and Political Rights provides globally established norms and rules, ratified by 170 States and echoing the Universal Declaration of Human Rights, guaranteeing "the right to hold opinions without interference" and "the right to seek, receive and impart information and ideas of all kinds, regardless of frontiers" and through any medium. Under art. 19 ICCPR and art. 10 ECHR, Freedom of Expression can only be restricted or interfered with in accordance with the three-step test: Legality (provided by law), proportionate and necessary and legitimate (for an enumerated purpose). Note also that the relationship between the right and the restriction must not be reversed.[7]

The proposal would be subject to challenge on all three grounds. As to the first, on the basis that (lacking any contours) it fails to alert the citizen in advance, he cannot foresee liability and adapt his conduct accordingly. See *Editorial Board of Pravoye Delo and Shtekel v Ukraine*, where the ECtHR observed that '*a norm cannot be regarded as a "law" unless it is formulated with sufficient precision to enable the citizen to regulate his conduct: he must be able – if need be with appropriate advice – to foresee, to a degree that is reasonable in the circumstances, the consequences which a given action may entail'*. This principle was also enunciated by the ECtHR in The Sunday Times v the United Kingdom.

[7] Human Rights Committee, *General Comment No. 34, Article 19: Freedoms of Opinion and Expression* (12 September 2011), para.21.

Restrictions must be "provided by law" and be adopted by regular legal processes and limit government discretion in a manner that distinguishes between lawful and unlawful expression with "sufficient precision". Secretly adopted restrictions fail this fundamental requirement. The assurance of legality should generally involve the oversight of independent judicial authorities.

Further, a new general duty of care cannot be necessary (in light of the existing legislation) nor proportionate (same) if it has no contours at all. Much like blanket surveillance was found to be unlawful, so too is blanket liability. The requirements of necessity and proportionality mean states must demonstrate that the restriction imposes the least burden on the exercise of the right and actually protects, or is likely to protect, the legitimate State interest at issue. States may not merely assert necessity but must demonstrate it, in the adoption of restrictive legislation and the restriction of specific expression. Restrictions pursuant to article 20 (2) of the ICCCP — which requires States to prohibit "advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence" — must still satisfy the cumulative conditions of legality, necessity and legitimacy. The UN Rabat Plan of Action offers a rigorous six-part test for expressions considered as criminal offences by analysing the context, speaker, intent, content and form, extent of the speech act, and likelihood including imminence. Further, states are generally obliged to adopt the least intrusive means of meeting one of the purposes listed in article 19(3) of the ICCPR. This principle is enunciated in the European Commission against Racism and Intolerance General Policy Recommendation No. 15 on Combating Hate Speech.

States also have a duty to ensure that private entities do not interfere with the freedoms of opinion and expression. The Guiding Principles on Business and Human Rights, adopted by the Human Rights Council in 2011, emphasize in principle 3 State duties to ensure environments that enable business respect for human rights. States may not restrict the right to hold opinions without interference.

14. Vague, overbroad regulation of speech

As the UNHRC report (above) noted at [15] *"Imposition of company obligations. Some States impose obligations on companies to restrict content* ***under vague or complex legal criteria without prior judicial review and with the threat of harsh penalties.*** *For example, the Chinese Cybersecurity Law of 2016 reinforces vague prohibitions against the spread of "false" information that disrupts "social or economic order", national unity or national security; it also requires companies to monitor their networks and report violations to the authorities. Failure to comply has reportedly led to heavy fines for the country's biggest social media platforms. [16]. Obligations to monitor and rapidly remove user-generated*

content have also increased globally, establishing punitive frameworks likely to undermine freedom of expression even in democratic societies. The network enforcement law (NetzDG) in Germany requires large social media companies to remove content inconsistent with specified local laws, with substantial penalties for non-compliance within very short time frames. The European Commission has even recommended that member States establish legal obligations for active monitoring and filtering of illegal content."

The duty of care proposed enables just such a vague and punitive approach to supress speech and opinion by private actors when it would be totally illegal and inappropriate and unworkable to have legislation prohibiting the same speech and opinion. There is also a lack of any due process and the proposal provides neither for notice or a hearing either before or after removal of speech. The rule of law and its normal operation is suspended –with obvious dangers.

15. Prior Restraint and censorship by private actors

There is no allowance whatsoever for the rule against prior restraint which is long-standing and designed to protect the courts (or private tribunals) from acting as censors. If it really is intended that these private companies censor the speech they host –the standards to which they are asked to do so, should be enshrined in law and not be a matter of their own views or opinions. This is fudged by the use of contract terms instead of national local laws.