GETTING TO THE ROOT OF THE PROBLEM

Written Representations to the Select Committee on Deliberate Online Falsehoods

SUI Yi Siong
LLB (Singapore Management University);
Advocate and Solicitor, Supreme Court of Singapore

Lyndon **CHOO** Hao Ren
4th Year Undergraduate,
School of Law, Singapore Management University

CHEN Lixin and Aaron **YOONG** Joon Wei
3rd Year Undergraduates,
School of Law, Singapore Management University

7 March 2018

Contents

I. Introduction 2

II. Old wine in new wineskins – The nature of the threat posed by deliberate online falsehoods 5

A. Deliberate falsehoods are not a new threat to society 6

B. The real mischief in deliberate online falsehoods lies in the ease of access and dissemination via internet intermediaries 9

III. Preventing and combat online falsehoods by imposing intermediary liability 13

A. An analysis of overseas approaches to regulating intermediaries 15

B. A proposed model for regulating intermediaries in Singapore 18

III. Conclusion 22

***Abstract**: The phenomenon of the spread of deliberate falsehoods has exacerbated in the current day and age with the usage of technology. An examination of this issue shows that that the motivations and reasons for spreading such falsehoods have not changed. Instead, the primary mischief lies in the near instantaneous dissemination and ease of access via internet intermediaries.*

Despite this, there is a gap in the regulatory tools available to deal with this mischief. The solution hence should lie in imposing some sort of liability on internet intermediaries to remove deliberate falsehoods. Nevertheless, it is crucial that such intermediary liability must be carefully calibrated to minimize restrictions on the right to freedom of expression. These written representations will provide some solutions and suggestions as to the form intermediary liability may take.

I. Introduction

1. It is acknowledged that deliberate online falsehoods present a real and pressing danger in today's world. As set out in the Ministry of Communications and Information and Ministry of Law's Green Paper on "*Deliberate Online Falsehoods: Challenges and Implications*" ("**Green Paper**"), the impact of online falsehoods can be significant.[1]

2. However, the motivations that drive the spread of online falsehoods are neither unique nor contemporary. Foreign state actors seeking to influence domestic politics, and unscrupulous individuals attempting to make a quick buck, have all perpetrated deliberate falsehoods long before the Internet was even conceived.[2] But what has changed is the medium in which deliberate falsehoods now spread – through internet intermediaries like social media platforms and search engines.

[1] (Green Paper, *Misc 10 of 2018*, 5 January 2018) at paras 14–50.
[2] Green Paper at para 78.

Given the "*ubiquity, universality and utility*" of such mediums,[3] deliberate falsehoods can now reach a virtually unlimited audience instantaneously.

3. Hence, the logical solution to deliberate online falsehoods should be to regulate intermediaries. But how far and how much regulation should be imposed? And to what extent should an intermediary be held liable? These written representations to the Select Committee on Deliberate Online Falsehoods – Causes, Consequences and Countermeasures (the "**Select Committee**") will deal with these issues and proffer some suggestions for getting to the root of the problem. In so doing, it will pay particular attention to the following two areas under the Select Committee's Terms of Reference:

    a. The motivations and reasons for the spreading of such falsehoods, and the types of individuals and entities, both local and foreign, which engage in such activity; and

    b. How Singapore can prevent and combat online falsehoods, including the principles that should guide Singapore's response, and any specific measures, including legislation, that should be taken.

4. In summary, our views are as follow:

    a. First, the motivations and reasons for spreading deliberate falsehoods have not changed. Political and financial motivations remain the chief reasons for deliberate falsehoods. However, while existing legislation may be sufficient to deal with the authors of such falsehoods, there is a gap in the regulatory tools available to deal with the chief mischief of deliberate online falsehoods, which is their near instantaneous dissemination and ease of access.

[3] *Barrick Gold Corp. v Lopehandia* [2004] CanLII 12938 [28]–[31]; See also *Dow Jones & Company Inc v Gutnick* [2002] HCA 56 [80]; Potts, *Cyberlibel: Information Warfare in the 21st Century?* (Irwin Law Inc, 2011) 29.

b. Second, bearing in mind the gap in tools to combat deliberate online falsehoods, some sort of intermediary liability should be imposed to combat and prevent the spread of deliberate online falsehoods. However, any such intermediary liability must be carefully calibrated to minimize restrictions on the right to freedom of expression.

5. We hope that these Written Representations will be of assistance to the Select Committee.

II. Old wine in new wineskins – The nature of the threat posed by deliberate online falsehoods

6. As a preliminary point, there have been concerns that any approach taken towards the phenomenon known as "*fake news*" may be overly restrictive.[4] While lies and falsehoods ought to be restricted, care must be taken not to stifle genuine discourse. Indeed, this was a point raise by several Members of Parliament who questioned whether there was a need to legislate to tackle "*fake news*".[5]

7. However, any concerns in this regard are addressed by the Select Committee's specific mandate to investigate "*deliberate online falsehoods*", rather than the more generic label of "*fake news*". The Green Paper has also characterized the issue as such.[6] Hence, it is clear that it is only falsehoods which are spread with a deliberate intent, rather than those spread negligently or recklessly, that is targeted.

8. The scope of "*deliberate online falsehoods*" also excludes what could be charitably called "*legitimate but biased commentary*",[7] and a whole range of phenomena that do not necessarily call for restriction, like parodies and bad journalism.[8] At the end of the day, contrary opinions should not be stifled by dint of simply being disagreeable.[9] Furthermore, it cannot be denied that the Internet has had a democratizing effect in relation to political speech and public discourse, equalizing opportunities for communication and facilitating the freedom to disseminate and receive information.[10]

---

[4] Ivan Teo, "Take care not to stifle genuine discourse", (*Straits Times Forum*, 15 January 2018) <www.straitstimes.com/forum/letters-in-print/take-care-not-to-stifle-genuine-discourse> (accessed 26 February 2018).
[5] *Singapore Parliamentary Debates*, Official Report (10 January 2018) vol 94 at p 45 (Ms Sun Xueling, Pasir Ris-Punggol); p 50 (Mr Kok Heng Leun, Nominated Member); and p 68 (Asst Prof Mahdev Mohan, Nominated Member).
[6] Green Paper at para 50.
[7] *Singapore Parliamentary Debates*, Official Report (10 January 2018) vol 94 at p 50 (Mr Kok Heng Leun, Nominated Member).
[8] Damian Tambini, Fake *News, Public Policy Responses* (Media Policy Brief 20, London School of Economics and Political Science, Media Policy Project, 2017).
[9] *Singapore Parliamentary Debates*, Official Report (10 January 2018) vol 94 at p 50 (Mr Kok Heng Leun, Nominated Member).
[10] Thio Li-ann, *A Treatise on Singapore Constitutional Law* (Academy Publishing, 2012) ("**Thio's Treatise**") at para 14.037.

9. As for spreading deliberate falsehoods, few would argue that any utility can be gained from it, whether online or offline. Indeed, as the US Supreme Court in the case of *Schenck v United States* noted in its oft-cited observation, "*the most stringent protection of free speech would not protect a man falsely shouting fire in a theater and causing panic*".[11] More recently, the UN *Joint Declaration on Freedom of Expression and "Fake News", Disinformation and Propaganda*[12] has recognized that deliberate falsehoods interfere with the public's right to know, and the right of individuals to seek, receive and impart information and ideas of all kind.[13] Furthermore, specific kinds of falsehoods are intrinsically harmful, such as those that harm individual reputations and privacy, or incite to violence or hostility against identifiable groups in society.[14]

10. With these parameters in mind, we turn now to the motivations and reasons for spreading deliberate online falsehoods. In so doing, we will show that these motivations and reasons have not changed over the years. Rather, what has changed is the medium in which they are now spread, namely internet intermediaries.

A. Deliberate falsehoods are not a new threat to society

11. According to the Green Paper, there is an increasing use of digital technologies to spread falsehoods. There appear to be two main types of actors: private individuals and entities, and foreign state actors. The latter aim to engineer specific outcomes in elections and referenda, while private actors seem to be motivated by financial considerations. The Green Paper also observes that many online falsehoods are aimed at interfering with elections and referenda.[15]

12. However, deliberate falsehoods are nothing new in the arena of politics. The highly bruising 2016 US Presidential elections has often been compared with that of

[11] 249 US 47 (1919) at 52 (Holmes J).
[12] UNHRC, 'Joint Declaration on Freedom of Expression and "Fake News", Disinformation and Propaganda' (3 March 2017) FOM GAL/3/17 ("**2017 Joint Declaration"**).
[13] 2017 Joint Declaration at preamble.
[14] 2017 Joint Declaration at preamble.
[15] Green Paper at para 15.

1800, where Democratic-Republican Thomas Jefferson and Federalist incumbent John Adams also used media (traditional newspapers) to target each other.[16] In any case, the "weaponisation" of information,[17] and specifically, deliberate falsehoods finds its roots even further back in history. One such example comes from World War I, where both Germany and the UK made use of fabricated stories about wartime atrocities to undermine each other's military efforts.[18]

13. As for private actors, supermarket tabloids have long trafficked in a mix of partially true and outright false stories.[19] The publication of deliberate falsehoods for financial profit has a long and storied past, going all the way back to the 19th century and what was known as "*yellow journalism*".[20] Large corporations have also been known to finance suspect research in order to ensure their dominance – tobacco companies that fund research that challenges the correlation of smoking with lung cancer, and fossil fuel manufacturers that fund research that attributes climate change to natural causes rather than human activities or carbon emissions.[21]

14. Turning to Singapore, Parliament during its debate on the Select Committee also heard examples of deliberate falsehoods. Some were recent; others were propagated by old fashion means of rumor-mongering. In the 1960s-70s, false rumours were used to turn races against each other, such as during the race riots

---

[16] Hunley at 499; A. Roger Ekirch, "Echoes from the Presidential Election of 1800" (*HUFFINGTON POST*, June 7, 2017) <www.huffingtonpost.com/a-roger-ekirch/echoes-from-the-president_b_10326026.html> (accessed 26 February 2018); Katie Reilly, "Lin-Manuel Miranda: 2016 Race 'No More Bizarre' Than Election of 1800" (*TIME*, 1 June 2016) <http://time.com/4353996/hamilton-lin-manuelmiranda- 2016-election/> (accessed 26 February 2018).

[17] A term used by The Rt Hon Theresa May MP, *PM Speech to the Lord Mayors Banquet 2017*, GOV.UK (Nov 13 2017), www.gov.uk/government/speeches/pm-speech-to-the-lord-mayors-banquet-2017.

[18] Barbaras Starmans, Ten Examples of Fake News from History, The Social Historian http://www.thesocialhistorian.com/fake-news/

[19] Hunt Allcot & Matthew Gentzkow, "Social Media and Fake News in the 2016 Election" (2017) 31:2 Journal of Economic Perspectives 211 ("**Allcot & Gentzkow**") at 214.

[20] The Public Domain Review, "Yellow Journalism: The 'Fake News' of the 19th Century" <https://publicdomainreview.org/collections/yellow-journalism-the-fake-news-of-the-19th-century/> (accessed 26 February 2018); Alexandra Samuel, "To fix fake news, look to yellow journalism" (*JSTOR Daily*, 29 November 2016) <https://daily.jstor.org/to-fix-fake-news-look-to-yellow-journalism/> (accessed 26 February 2018).

[21] Bruce McComiskey, *Post-Truth Rhetoric and Composition* (Utah State University Press, 2017) ("**McComiskey**") at p 13.

of 1969 which lasted for seven days and resulted in four dead and 80 injured – this was sparked off partly because of rumours of Malays attacking Chinese in Malaysia.[22] A more recent example given was NTUC Fairprice's constant battle with false online rumours about its products, like "*halal pork*", "*plastic rice*" and "*man-made eggs from China*".[23]

15. It can thus be seen that deliberate falsehoods are not new to either Singapore or the world. This recognition is crucial as it acknowledges that there are existing laws that can be used to combat the most common types of deliberate online falsehoods, such as falsehoods with a view to undermining national security or national elections, or stirring up hatred towards minority racial or religious groups.

16. For example, the Sedition Act prohibits any act or publication which has a seditious tendency, which includes promoting feelings of ill-will and hostility between different races.[24] Similarly, the Maintenance of Religious Harmony Act allows for a restraining order to be made against a person who attempts to cause feelings of hatred or hostility between different religious groups.[25]

17. A recent example of the Sedition Act being used against deliberate online falsehoods is the recent prosecution of the editors of the now-defunct The Real Singapore blog. Ai Takagi and Yang Kaiheng were charged under the Sedition Act for publishing articles which promoted feelings of ill-will and hostility between different classes of Singapore society. Just like the Macedonian teenagers who earnt staggering profits from creating false articles about the 2016 US Presidential election,[26] Takagi and Yang fabricated articles which sensationalized social fault lines for the purpose of boosting their advertising revenue from high internet traffic.

---

[22] *Singapore Parliamentary Debates*, Official Report (10 January 2018) vol 94 at p 73 (Er Dr Lee Bee Wah, Nee Soon).
[23] *Singapore Parliamentary Debates*, Official Report (10 January 2018) vol 94 at p 61 (Mr Seah Kian Ping, Marine Parade).
[24] Cap 290, 2013 Rev Ed, ss 3 – 4.
[25] Cap 167A, 2001 Rev Ed, ss 8 – 9.
[26] Green Paper at paras 17, 56.

18. However, while holding the author of a deliberate falsehood is always a viable remedy, it is not necessarily practicable. First, many authors of such "*fake news*" are anonymous, and second, even if one could discover their identity, these falsehoods are sometimes developed outside the jurisdiction.[27] More importantly, the length of time and amount of resources necessary to bring a criminal prosecution often renders it a blunt tool at best for dealing with the real threat of deliberate online falsehoods. And it is to this consideration that we now turn.

B. The real mischief in deliberate online falsehoods lies in the ease of access and dissemination via internet intermediaries

19. Given that deliberate falsehoods have always been present, what is so dangerous and different about deliberate online falsehoods? We submit that the real mischief lies in the ease of access and dissemination of information online.

20. Support for this proposition can be found in the jurisprudence of the European Court of Human Rights ("**ECtHR**"), the supranational court that ensures the compliance of contracting states parties to the European Convention of Human Rights. In the ECtHR Grand Chamber decision in *Delfi AS v Estonia* ("**Delfi**"),[28] the court found that holding an online news outlet liable for offensive comments posted by its readers did not violate the right to freedom of expression under Article 10 of the European Convention on Human Rights. In doing so, the ECtHR observed that the "*ease, scope and speed of the dissemination of information on the Internet, and the persistence of the information once disclosed … may considerably aggravate the effects of unlawful speech on the Internet compared to traditional media*".[29] The case of *Delfi* follows a line of ECtHR cases which have all recognized the risk of harm posed by content and communications on the internet.[30]

---

[27] Lee Royster, "Fake News: Political Solutions to the Online Epidemic" (2017) 96 NCL Rev 270 at 285.
[28] *Delfi AS v Estonia* App no 40287/98 (ECtHR, 16 June 2015) ("***Delfi***").
[29] *Delfi* at [147].
[30] See also *Editorial Board of Pravoye Delo and Shtekel v Ukraine* App no 33014/05 (ECtHR, 5 August 2011) at [63] and *Mosley v UK* App no 48009/08 (ECtHR, 15 September 2011) at [130].

21. In other words, and to use an oft-misattributed quote, a lie can travel halfway around the world before the truth can get its shoes on. The Internet simply gives the lie more of a head start. For example, during the lead-up to the 2016 US Presidential election, a rumour was spread that Democratic nominee Hilary Clinton and her chief of staff were running a paedophile ring out of a pizza restaurant. This "*Pizzagate*" Incident not only resulted in threats and demonstrations against the restaurant and its owners,[31] but also led a 28 year-old man to arrive at the restaurant with an AR-15 rifle to rescue the imaginary children.[32] What would have once been dismissed as an obvious hoax was given the cloak of legitimacy simply by virtue of "*going viral*" on social media.[33]

22. Closer to home, the longevity of the "NTUC Fairprice sells *halal* pork" rumour is testament to the longevity of a deliberate online falsehood. As explained by Member of Parliament Mr Seah Kian Peng, this false rumour first surfaced in 2007, but despite police reports and subsequent public announcements rebutting this claim, "*halal pork*" reappeared in 2011 and in 2014.[34]

23. Indeed, the danger of instantaneous dissemination to a virtually unlimited audience was recognized in Parliament. Member of Parliament Ms Rahayu Mahzam summed it up it best in the following extract from her speech:

> *Mr Deputy Speaker, Sir, the deliberate spreading of falsehoods is not a novel phenomenon. For as long as human societies existed, there would always be incidence of rumours, untruths being spread in various media. Such falsehoods affect people, relationships, the way people see things, and ultimately, impacts on the society. Human societies have managed to cope and deal with such incidents and its consequences. However, with the*

---

[31] See Green Paper at para 20, describing the event more commonly known as "Pizzagate".
[32] Cathleen Decker and Michael Memoli, "Online conspiracy theory turns into a real danger after armed man shows up at a Washington pizza parlour" (*LA Times*, 5 December 2016) < http://www.latimes.com/politics/la-na-pol-comet-conspiracy-theory-20161205-story.html> (accessed 3 March 2018).
[33] See McComiskey at pp 15 – 16.
[34] *Singapore Parliamentary Debates*, Official Report (10 January 2018) vol 94 at p 61 (Mr Seah Kian Ping, Marine Parade).

advent of digital technology, the spread of falsehoods and its impact on communities have reached unprecedented levels. The spread of deliberate online falsehood is, therefore, of particular concern.

Firstly, the manner in which we receive and share information has changed significantly. There are now many more information systems and platforms which translates to more means and ways by which falsehoods can be spread. With the proliferation via social media, falsehoods spread much faster. We now have the concept of "going viral", a problem which we never had to deal with before. There is now instant communication – Twitter, Facebook Messenger, Instagram, Telegram – and with a click of a button, an individual could share information to many others almost immediately. On Facebook, an individual could share a post to 5,000 people. On a single WhatsApp group, you can have up to 256 members. Imagine the extent of the reach in numerous WhatsApp groups.[35]

24. We argue that this is the real danger posed by deliberate online falsehoods. The ease of dissemination afforded through the internet and social media platforms leads to a "crowding-out" of truth and fact. The propagation of deliberate online falsehoods devalue and delegitimize voices of expertise, authoritative institutions, and the concept of objective data, all of which undermines society's ability to engage in rational discourse based upon shared facts.[36] Furthermore, attempts to correct errors may be pointless, given the ease of reproducing these errors on another online platform.[37] Going back to the Pizzagate Incident, mainstream media

[35] *Singapore Parliamentary Debates*, Official Report (10 January 2018) vol 94 at p 47 (Ms Rahayu Mahzam, Jurong).
[36] Sandra Bacon & Rebecca Crootof, "Fighting Fake News: Workshop Report" <https://law.yale.edu/system/files/area/center/isp/documents/fighting_fake_news_-_workshop_report.pdf> (accessed 26 February 2018>.
[37] Thio's Treatise at para 14.042.

articles that exposed the falsity of this rumour failed to have any discernible impact.[38]

25. In fact, this is why it has been argued that the "*marketplace of ideas*" is no defence to deliberate online falsehoods. The marketplace of ideas is the theory that in free and public discourse, all ideas should be available to the community – including false ones – because a restriction on speech of any kind might incidentally restrict the truth. It holds that without government interference, truth and falsity will compete in a marketplace, and truth will emerge victorious.[39]

26. However, the marketplace of ideas requires two implicit assumptions to function properly: first, that truth is in fact discoverable and able to be proven, and second, that participants in the marketplace are actually seeking to uncover the truth.[40] Both these assumptions do not appear to be applicable to deliberate online falsehoods. The sheer volume of information available online through a variety of sources, not all of which are credible or trustworthy, inhibits the discernment of truth.[41]

27. Similarly, people may filter and receive information from certain preferred sources only, precluding an engagement with competing views that may nonetheless provide the truth.[42] This is known as the "*echo chamber*" effect, where participants in online communities constantly have their own opinions echoed back to them, which reinforces their original (potentially false) beliefs.[43] Also, given the way social media algorithms work, it may be that it is simply impossible or unlikely for consumers of false information to be provided with alternative (true) information – for example, Facebook's algorithms are designed to populate users' news feeds

[38] Cecilia Kang and Adam Goldman, "In Washington Pizzeria Attack, Fake News Brought Real Guns" (*The New York Times*, 5 December 2016) <https://www.nytimes.com/2016/12/05/business/media/comet-ping-pong-pizza-shooting-fake-news-consequences.html> (accessed 3 March 2018).
[39] Annie Hundley, "Fake News and the First Amendment: How False Political Speech Kills the Marketplace of Ideas" (2017) 92 Tul L Rev 497 ("**Hundley"**), 502.
[40] Hundley, at 502–503.
[41] Thio's Treatise at paras 14.039 and 14.042.
[42] Thio's Treatise at para 14.044.
[43] Thio's Treatise at para 14.044.

with content similar to material previously "*liked*".[44] As such, people who get news primarily from social media may be less likely to receive information that would counter or expose a false story.[45]

28. Overall, it can be seen that the usual weapons for combatting false news – rebuttals via objective reporting which is publishing in mainstream media – are not effective in the context of deliberate online falsehoods. Setting up officially sanctioned fact-checking websites, while useful, may not be completely effective in tackling the main mischief posed by deliberate online falsehoods.

29. In addition, it is clear that there is a gap in the regulatory or legislative tools that can be used to combat the speed of dissemination of online falsehoods. The process of instituting a criminal prosecution takes time, in that a disclosed offence has to be investigated for sufficient evidence before the decision to bring charges is made.[46] By the time the alleged purveyor of deliberate falsehoods is charged in court, the falsehood could have incited acts of violence or other illegal behaviour. Indeed, in the ECtHR case of *Delfi*, the Grand Chamber agreed that the intermediary ought to have taken down hate speech "*without delay*" because of the threat that hate speech poses to society.[47]

III. Preventing and combat online falsehoods by imposing intermediary liability

30. How then, to combat and prevent the spread of deliberate online falsehoods? We submit that the solution is to restrict access to them – in other words, restrict them at the source, which in this day and age, is through internet intermediaries.[48] Such intermediaries include Google's search engine and Twitter's social network, so called because they mediate online communication and enable various forms of

[44] See McComiskey at 19.
[45] Allcott & Gentzkow at 221.
[46] See Lucien Wong SC, "Prosecution in the Public Interest", Singapore Law Review Lecture 2017 at paras 7 – 9.
[47] *Delfi* at [159].
[48] Electronic Frontier Foundation, "The Manila Principles on Intermediary Liability Background Paper" (*EFF*, 22 March 2015) <www.eff.org/sites/default/files/manila-principles-background-paper-0.99.pdf> ("**Manila Principles Background Paper**") (accessed 26 February 2018) p 2.

online expression.[49] Social media platforms are also regarded as intermediaries, as they connect individual content users by allowing them to exchange text, photos, videos, as well as by allowing them to post content to their network of contacts, or to the public at large.[50]

31. Internet-related misconduct is arguably more effectively regulated through the imposition of responsibility on intermediaries.[51] After all, intermediaries are best placed to ensure that the information disseminated through them meet certain minimum standards. Depending on the type of intermediary and the services offered, intermediaries control how their users communicate and with whom, and have access to information created by users as well as a range of information directly related to users.[52]

32. Hence, we submit that the most appropriate and effective way of combating deliberate falsehoods is to place liability on internet intermediaries. As the modern gatekeepers of information, intermediaries ought to assume a degree of responsibility for content that they host or publish.

33. However, even if it is accepted intermediary liability is the best way of combating deliberate online falsehoods, there are several ways in which it can be implemented, and not all may be suitable for Singapore's context. The next section will give a brief overview of the types of intermediary liability that may be imposed, before we conclude by evaluating and giving suggestions as to which model could be adopted by Singapore.

---

[49] Rebecca MacKinnon, Elonnai Hickok, Allon Bar, and Hae-in Lim, "Fostering Freedom Online: The Role of Internet Intermediaries" (UNESCO, 2014) <http://unesdoc.unesco.org/images/0023/002311/231162e.pdf> ("**Fostering Freedom Online**") (accessed 26 February 2018) p 15.

[50] Fostering Freedom Online, p 22.

[51] Ronald Mann and Seth Belzley, 'The Promise of Internet Liability' (2005) 47 William & Mary Law Review 239, 265; Aleksandra Kuczerawy, 'The Power of Positive Thinking: Intermediary Liability and the Effective Enjoyment of the Right to Freedom of Expression' (2017) 8 Journal of Intellectual Property, Information Technology and Electronic Commerce Law 266, para 7; Lilian Edwards, 'Role and Responsibility of Internet Intermediaries in the Field of Copyright and Related Rights' (*WIPO*, 2010) <http://www.wipo.int/export/sites/www/copyright/en/doc/role_and_responsibility_of_the_internet_intermediaries_final.pdf> accessed 21 January 2018, 26.

[52] Fostering Freedom Online, p 23.

A. An analysis of overseas approaches to regulating intermediaries

34. It is commonly accepted that there are 3 models of regulating intermediaries: (i) broad immunity from liability; (ii) strict liability for third party content; and (iii) conditional immunity from liability, or "*safe harbour*".[53]

35. Under the first model, broad immunity from liability, intermediaries are generally exempt from liability for most third-party content because they are not considered as publishers. Instead, these intermediaries are treated as mere "messengers" who are not responsible for the content they disseminate, and are not required to monitor their platforms.

36. One of such example is section 230 of the Communications Decency Act ("**CDA**") in USA, which states that "[*n]o provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider*." The CDA shields online intermediaries from most claims such as defamation and negligence, with the exception of matters of intellectual property infringement.[54] Additionally, intermediaries are not required to remove harmful content - they are only encouraged to do so in good faith under section 230(c). Under such a broad immunity framework, courts have found the intermediaries immune from liability, even where false information was transmitted. This was the case in unsurprising *Gentry v eBay, Inc,* and *Goddard v Google*. In fact, under this model it is rare for intermediaries to be found liable for unlawful content posted by third parties, even if the post contains obviously illegal materials such as explicit sexual content of minors.[55]

---

[53] Center for Democracy and Technology, "Shielding the Messengers: Protecting Platforms for Expression and Innovation" (2012) <https://cdt.org/files/pdfs/CDT-Intermediary-Liability-2012.pdf> ("**Shielding the Messengers**") (accessed 26 February 2018) pp 4–12; Article 19, "Internet Intermediaries: Dilemma of Liability" (2013) <www.article19.org/data/files/Intermediaries_ENGLISH.pdf> ("**Dilemma of Liability**") (accessed 26 February 2018) p 7; Manila Principles Background Paper, pp 12–15.
[54] Fostering Freedom Online, p 42.
[55] *Doe v. America Online*, 783 So. 2d 1010, 1013-1017 (Fl. 2001); *Doe v. MySpace*, 528 F.3d 413 (5th Cir. 2008).

37. The next model is known as strict liability for third party content. On the other extreme from the broad immunity model, the strict liability approach holds intermediaries liable for third-party content as if they are the content creator, even though they might not be aware of the existence of the unlawful content and of its illegality. Such an approach does not differentiate between the publisher of original content and an intermediary.

38. An example of such strict liability imposed for third party content can be found in China. Under Article 36 of the Tort Liability Law of the People's Republic of China,[56] an internet intermediary can bear either primary or joint liability for infringing acts. Internet intermediaries' liability depends on whether they directly infringed the interest of other individuals, or whether they chose inaction when they had knowledge of, or was notified of the presence of unlawful user content on their websites.[57] This section of the legislation is applicable for general claims against internet intermediaries beyond claims of intellectual property infringements.[58]

39. It appears that the Chinese courts have taken a liberal approach in considering whether internet intermediaries possess knowledge of unlawful user content both before and after the enactment of the Tort Liability Law.[59] This essentially makes internet intermediaries strictly liable for users' content.[60] This is seen in the *Tianya* decision predating the legislation,[61] where the internet intermediary was liable for a user's defamatory post, even if no notification was given to the internet intermediary. More recently, Sina.com faced sanctions for the mere presence of user generated pornographic material which were detected on its network.[62]

---

[56] Tort Liability Law of the People's Republic of China was adopted by the Standing Committee of the Eleventh National People's Congress on 26 December 2009, and came into force on 1 July 2010, art 36.
[57] *Ibid*.
[58] Qian Tao, "Legal Framework of Online Intermediaries' Liability in China" (2012) 14 Info: The Journal of Policy, Regulation and Strategy for Telecommunications, Information and Media 59, 61, 64.
[59] *Ibid.*
[60] Fostering Freedom Online, p 40.
[61] *Zhang Fu vs Tianya Online Network Technology Co, Ltd*, People's Court of Taojiang County, Hunan Province.
[62] Glyn Moody, "China's Internet Giant Sina.com Loses Publication License for Publishing Pornography – 20 Articles and Four Videos" (*Techdirt*, 30 April 2014)

40. Thailand also broadly imposes liability on intermediaries for third party content under the 2007 Computer Crimes Act ("**CCA**").[63] While an individual may be punished for posting content that is likely to injure national security or a "*false*" publication online,[64] internet intermediaries that "*intentionally*" transmit or host such content are also made "*liable to the same penalty*" that the content creators face.[65] Without a safe harbour regime and precise definitions of unlawful speech, the CCA has been criticised for being inappropriately used to punish political dissent.[66] One such example is the case of Chiranuch Premchaiporn, where the executive director of Prachatai (an independent news website) was criminally convicted for being too slow in removing user comments that were deemed insulting to Thailand's monarchy.[67]

41. The third and final model is known as the conditional immunity from liability model, or "*safe harbor*" model. Under the "*safe harbour*" model, intermediaries receive protection against liability for user content only if they meet certain criteria, such as removing third-party content after certain conditions are met, such as upon receiving notice (also known as "*notice and takedown*") or a court order (also known as "*notice and judicial takedown*"). Such safe harbour regime has typically been found in legislation that pertains to copyright infringement claims, in countries such as the United States, Chile, Colombia, the Republic of Korea, and Singapore.[68] For example, the Digital Millennium Copyright Act in the United States exempt liability only for intermediaries that "*reasonably implement" a policy of terminating the accounts of repeated offenders*.[69]

<www.techdirt.com/articles/20140425/09451027029/chinas-internet-giant-sinacom-loses-publication-license- publishing-pornography-20-articles-four-videos.shtml> (accessed 26 February 2018).

[63] Computer Crimes Act (No. 2) 2017.

[64] Section 14(1) of the 2007 Computer Crimes Act.

[65] Section 15 of the 2007 Computer Crimes Act; Shielding the Messengers, p 14.

[66] Shielding the Messengers, p 14.

[67] Sean W. Crispin, "Internet Freedom on Trial in Thailand" (CPJ Blog, 4 February 2011) <http://cpj.org/blog/2011/02/internet-freedom-on-trial-in-thailand.php> (accessed 26 February 2016) p 7; Centre for Democracy and Technology, "CDT Objects to Conviction of Thai Webmaster as Threat to Free Expression" (May 30, 2012) <https://cdt.org/press/cdt-objects-to-conviction-of-thai-webmaster-as-threat-to-free-expression> (accessed 26 February 2018).

[68] Shielding the Messengers, p 7.

[69] (Public Law 105–304, 112 Stat. 2860) Section 512(i).

42. Specifically with regard to removing unlawful content, the EU E-Commerce Directive has established the criteria under which different categories of intermediaries can seek exemption from liability.[70] In order to benefit from limitation of liability, the intermediary has to "*act expeditiously to remove or disable access" to information upon obtaining actual knowledge of illegal activities*". Going back to the ECtHR case of *Delfi*, the intermediary did have a "*notice and takedown*" system in place – any reader could mark a comment as inciting hatred, and the comment would be removed.[71] However, the intermediary in *Delfi* was held liability because it did not act expeditiously to remove hate speech, which was only removed some six weeks after the it was notified.[72]

B. A proposed model for regulating intermediaries in Singapore

43. Having considered the three models of intermediary liability above, we now evaluate each of them and consider their suitability for Singapore.

44. The broad immunity model is consistent with existing concerns regarding freedom of expression at international law. International human rights bodies have been hesitant in approving the imposition of intermediary liability, instead emphasizing that intermediaries should not be liable "*for content produced by others when providing technical services*"[73] and that "*censorship measures should never be delegated to a private entity*".[74] As the argument goes, without legal pressure, intermediaries will be able to preserve their ability to function as a platform for individuals to express freely. However, considering the increasing influence that these online intermediaries wield, it is only practical for them to shoulder some of the responsibility in ensuring that the online platforms remain conducive to facilitate expression. Furthermore, as the Green Paper more than amply illustrates, the

[70] Council Directive (EC) 2000/31/EC on Certain Legal Aspects of Information Society Services, in Particular Electronic Commerce, in the Internal Market [2000] OJ L178/1.
[71] *Delfi* at [13].
[72] *Delfi* at [17] – [19].
[73] Organisation for Security and Co-operation in Europe, "2011 Joint Declaration on Freedom of Expression and the Internet" (1 June 2011) <www.osce.org/fom/78309?download=true> (accessed 26 February 2018).
[74] United Nations Human Rights Council, "UN Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression" (16 May 2011) Doc A/HRC/17/27 para 43.

consequences of allowing deliberate online falsehoods to spread unchecked are dire. The broad immunity model is therefore wholly inappropriate for Singapore.

45. The strict liability approach is also not appropriate, considering its practical implications. Such a model requires intermediaries to closely monitor user content and remove any publication that may potentially be unlawful, lest they face penalties for allowing its dissemination.[75] However, intermediaries themselves are typically ill-equipped to make judgments as to what amounts to objectionable material that requires to be taken down, due to a variety of reasons such as resource constraints, lack of oversight and accountability, or potential conflicts of interest.[76] By imposing an obligation to actively regulate their platforms, intermediaries will tend to err on the side of caution, resulting in excessive self-censorship that is detrimental to the freedom of expression. This precise concern was shared by the ECtHR in *Magyar Tartalomszolgáltatók Egyesülete v Hungary,*[77] where it held that demanding an intermediary to filter comments instead of relying on a reporting mechanism for the removal of such objectionable material would "*[require] excessive and impracticable forethought capable of undermining freedom of the right to impart information on the Internet*".[78]

46. This leaves us with the "*safe harbour*" model. We submit that this model strikes the appropriate balance between protecting intermediaries from liability and defining their responsibilities with regard to hosting unlawful content on their platforms. More importantly, it addresses the current concerns pertaining to the dissemination of unlawful content on the Internet, while instituting safeguards to avoid excessive private censorship. As noted by Frank LaRue, UN Special Rapporteur on Freedom of Opinion and Expression:[79]

---

[75] Shielding the Messengers, p 15; Manila Principles Background Paper, p 18.
[76] United Nations General Assembly, "2016 Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression" (11 May 2016) Doc A/HRC/32/38 para 44.
[77] *Magyar Tartalomszolgáltatók Egyesülete v Hungary* App no 22947/13 (ECtHR, 2 February 2016) ("***MTE***").
[78] *MTE* at [82].
[79] Organisation for Security and Co-operation in Europe, "2011 Joint Declaration on Freedom of Expression and the Internet" (1 June 2011) <www.osce.org/fom/78309?download=true> (accessed 26 February 2018).

"When assessing the proportionality of a restriction on freedom of expression on the Internet, the impact of that restriction on the ability of the Internet to deliver positive freedom of expression outcomes must be weighed against its benefits in terms of protecting other interests."

47. By suitably allocating the burden of regulating online content on all parties involved - the intermediary, its users, and the state, the "*safe harbour*" model is best able to accommodate this tension between protecting the intermediaries' right to freedom of expression (as well as their users') and pursuing other citizen's rights or interests of the State.[80]

48. The main issue underlying this model, however, pertains to choosing between the "*notice and takedown*" regime and the "*notice and judicial takedown*" regime. To reiterate, under the former, intermediaries may not avail themselves of the safe harbour exception if they fail to remove unlawful content upon receiving user notice on their platform within a reasonable period of time. On the other hand, intermediaries will only be liable under the "*notice and judicial takedown*" regime for not removing a post if the notice is in the form of a judicial order, confirming that the content is unlawful.

49. The "*notice and takedown*" regime adopted by states such as Germany has been criticised for placing the intermediary in a *quasi*-judicial position.[81] This is because the intermediaries are forced to make a decision on whether to remove the reported post upon user notification, effectively having to evaluate the legality of reported post. However, intermediaries generally lack the capacity to determine legality, which is contextual and subjective.[82] Even if they can exercise such a task,

---

[80] An example of such competing rights to freedom of expression includes the right to private life as enshrined under Article 12 of the Universal Declaration of Human Rights, which is usually relevant in defamation cases. The State may also prosecute speech in the interest of protecting national security or public order. Such public order concerns are also recognised as legitimate aims that a state can pursue under Article 19(3)(b) of the International Covenant on Civil and Political Rights.

[81] Article 19, "Germany: The Act to Improve Enforcement of the Law in Social Networks" (*Article 19*, August 2017) <https://www.article19.org/wp-content/uploads/2017/09/170901-Legal-Analysis-German-NetzDG-Act.pdf> (accessed 26 February 2018).

[82] Dilemma of Liability, p 14; *Scarlet Extended SA v SABAM* C–70/10 (CJEU, 24 November 2011) paras 49–53; *SABAM v Netlog NV* C–360/10 (CJEU, 16 February 2012) paras 47–51.

such an obligation should not be imposed because evaluating each post would be expensive and time-consuming, incentivizing the intermediary to err on the side of caution and exercise excessive self-censorship instead.[83]

50. Additionally, practical experiences with the notice and takedown procedures have not been entirely satisfactorily. There is often a lack of due process e.g. the content creator is usually not given an opportunity to appeal or consider the complaint provided by the third party before the post is removed.[84] Such an obligation eventually results in a chilling effect on the freedom of expression.

51. It thus follows that the "*notice and judicial takedown*" regime is the more appropriate solution as the judicial body is best placed to determine legality of each post. However, there also exist practical concerns - such a takedown process may be too slow and cumbersome.[85] This is especially so given the large volumes of such reported content, which will require substantial amount of time for the courts to process such complaints. Further, considering the ability of the Internet to rapidly disseminate such content, the harm arising from the post would already have materialised before the issuance of a judicial order.

52. Considering the pros and cons surrounding the two different regimes, it is suggested that both approaches should be combined. In instances where a post is clearly unlawful, the intermediary should expeditiously remove the post upon gaining knowledge of its existence through user complaints, without having to rely on a judicial order. On the other hand, when a post fails to fall neatly within established categories of unlawful content, the intermediary should still be entitled to wait for a judicial order to inform itself of the appropriate measure to take.

53. It also bears mentioning that such categories of unlawful content should be limited to specific circumstances. After all, expanding the scope of restrictions comes with

---

[83] Fostering Freedom Online, p 52; Manila Principles Background Paper, p 17.
[84] Article 19, "Response to EU consultation on the E-Commerce Directive" (*Article 19*, November 2010) <https://www.article19.org/data/files/pdfs/submissions/response-to-eu-consultation.pdf> (accessed 26 February 2018); Manila Principles Background Paper, pp 3, 11.
[85] Shielding the Messengers, p 10.

the risk of undermining the right to freedom of expression. While Singapore ultimately should develop its own unique restrictions that fits within its autochthonous legal system, guidance on establishing the prohibited categories can be gleaned from the United Nations. In particular, the report by the UN Special Rapporteur has set out the four categories of content that are prohibited under international law: (1) child pornography; (2) direct and public incitement to commit genocide; (3) advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence; and (4) incitement to terrorism.[86]

54. On a final note, in the event that requiring notice in the form of a court order is not an appropriate solution due to resource considerations, the burden of regulating unlawful posts can also be shifted to existing regulatory bodies. Such a solution has also been proposed by Professor Warren Chik, such that the Info-Communications Media Development Authority ("**IMDA**") should share the responsibility of dealing with deliberate online falsehoods.[87] This would not present substantial additional burdens, given that the IMDA presently supervises the removal of websites that host content that is in breach of the guidelines under the Broadcasting Act. Furthermore, as it appears that the existing categories set out by the guidelines[88] can be extended to capture deliberate falsehoods, this task of regulating deliberate falsehoods can be undertaken by the same authority.

III. Conclusion

55. As can be seen from the above discussion, the scourge of deliberate online falsehoods is not new. Rather, the true mischief posed by deliberate online falsehoods is the ease of access and dissemination to a virtually unlimited audience. Falsehoods which would have ordinarily died out on their own for want

---

[86] UNHRC, "Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression" (16 May 2011) UN Doc A/HRC/17/27 para 25–32.
[87] Warren B. Chik, "Fact or Fake News The "Role of Law" for Data Accuracy" (*Singapore Law Gazette,* June 2017) <http://v1.lawgazette.com.sg/2017-06/1879.htm> accessed 26 February 2018.
[88] Infocomm Media Development Authority, "Internet Code of Practice" <https://www.imda.gov.sg/-/media/imda/files/regulation-licensing-and-consultations/codes-of-practice-and-guidelines/acts-codes/19-policiesandcontentguidelinesinternetinternecodeofpractice.pdf?la=en> accessed 26 February 2018.

of an audience can now be sustained almost indefinitely so long as someone, somewhere, spreads it via an internet intermediary.

56. Because of this, the solution to deliberate online falsehoods should be targeted and calibrated to address the true mischief, and no more than that. There is no need to enact some sort of "deliberate online falsehood act", which would risk duplicating existing legislation. Rather, the key to combatting and preventing deliberate online falsehoods is to require intermediaries accept the responsibility of policing the information disseminated through them. Still, intermediaries should only be required to do what is reasonable in the circumstances, and not be saddled with overly onerous obligations. Adopting the "*safe harbour*" conditional liability model with the takedown framework as suggested above would strike the appropriate balancing between the various competing rights.

57. In closing, we hope that our written representations have contributed to the ongoing discussion regarding the vexed question of "*fake news*". There is consensus that this is a problem, but less agreement about how best to solve it. It is therefore highly laudable that Singapore is attempting to canvass all possible views before deciding on the appropriate course of action.

58. For the Select Committee's consideration please.