

# A PANDEMIC OF MISINFORMATION: LEGAL ISSUES CONCERNING INTERMEDIARY LIABILITY IN THE COVID-19 ERA\*

*Shiela Marie L. Rabaya\*\**

## I. INTRODUCTION

The unprecedented COVID-19 pandemic led millions of people toward the confines of their homes. Physical interaction ceased to be the norm, and the world of the Internet became the primary hub for people to interact with one another. While the physical aspects of society—restaurants, fitness centers, offices, and schools, among others—have shut down, the hustle and bustle of daily life continued by transitioning to online platforms. Business and pleasure are now mixed inside the home, with the Internet as the biggest propeller of the “new normal.”

The COVID-19 pandemic brought a host of uncertainties as to how people can conduct their daily lives. People peruse the Internet for answers to questions brought by the sudden and massive change in how the world works. Physical libraries are closed, professors can only be reached virtually, and the easiest way to quench burning curiosity is to flock toward information available to the public. These sources are made available on several Internet platforms, with the reach of these pieces of information highly dependent on the popularity of the platform.

In legal terms, these Internet platforms are known as “Internet Service Providers” (“ISP”) or intermediaries, which host third-party content that can be accessed by the public. These platforms are known as intermediaries precisely because they serve as the middle man between the third party content provider and the content consumer (the public). In the Philippines, ISPs or “service providers” are defined in the Cybercrime Prevention Act (“CPA”) and its Implementing Rules and Regulations (“IRR”) as “any public or private entity that provides to users of its service the ability

---

\* *Cite as Shiela Marie Rabaya, A Pandemic of Misinformation: Legal Issues Concerning Intermediary Liability in the COVID-19 Era*, 93 (Special Online Feature) PHIL. L.J. 126, [page cited] (2020).

\*\* J.D., University of the Philippines (2020); B.A. Psychology, *magna cum laude*, University of the Philippines (2016); Member, Order of the Purple Feather; Editor, Student Editorial Board, PHILIPPINE LAW JOURNAL Vol. 92.

to communicate by means of a computer system”<sup>1</sup> and “any other entity that processes or stores computer data on behalf of such communication service or users of such service.”<sup>2</sup>

Social media platforms fall within the broad definition of an ISP.<sup>3</sup> Although platforms such as Facebook, Twitter, Instagram, and Youtube were already popular prior to 2020, the global pandemic magnified the reach and importance of these platforms exponentially.<sup>4</sup> With the ubiquity of ISPs, it is imperative to examine the role that these platforms play in the current global situation. Left unchecked, it would not only be businesses and socialization which would transition to the Internet; crimes would also begin to proliferate in the online world. Particularly, it is important to examine the role that ISPs play in disseminating information to the public and in serving the public conversation. Given the nature of the new disease, public health in the age of COVID-19 involves a global effort.<sup>5</sup> It is thus important for people to know the truth about all matters—whether scientific or societal—that involve COVID-19.

In 2020, misinformation can pose a serious threat to public health.<sup>6</sup> A lot of things about COVID-19 remain unknown to the public, meaning that people may not be able to discern fact from fiction among the hundreds of pieces of information available online. The case of President Trump’s statement regarding the ingestion of bleach is illustrative.<sup>7</sup> This shows how

---

<sup>1</sup> Rep. Act. No. 10175 (2012), § 3(n).

<sup>2</sup> Rep. Act. No. 10175 Rules & Regs. § 3(ff).

<sup>3</sup> Data Privacy Philippines, *Internet Service Providers: Can they be held liable for cybercrimes?*, DATA PRIVACY PHILIPPINES WEBSITE, at <https://privacy.com.ph/articles/internet-service-providers-can-they-be-held-liable-for-cybercrimes> (last accessed June 4, 2020).

<sup>4</sup> Lucas Matney, *The lockdown is driving people to Facebook*, TECHCRUNCH WEBSITE, Apr. 30, 2020, at <https://techcrunch.com/2020/04/29/the-quarantine-is-driving-record-usage-growth-at-facebook>. Facebook CEO Mark Zuckerberg disclosed that in April more than 3 billion internet users logged onto a Facebook service, including its central app, Instagram, Messenger, or WhatsApp.

<sup>5</sup> Charlotte West, *‘A Global Effort’: Public Health, Medicine, and International Education in the Time of COVID-19*, NAFSA: ASSOCIATION OF INTERNATIONAL EDUCATORS WEBSITE, May 6, 2020, at <https://www.nafsa.org/ie-magazine/2020/5/6/global-effort-public-health-medicine-and-international-education-time-covid-19>

<sup>6</sup> Kimberly Rodgers, *Misinformation: A Threat to the Public’s Health and the Public Health System*, 26 J. PUB. HEALTH MGMT. & PRAC. 294 (2020), available at [https://journals.lww.com/jphmp/Fulltext/2020/05000/Misinformation\\_\\_A\\_Threat\\_to\\_the\\_Public\\_s\\_Health.15.aspx](https://journals.lww.com/jphmp/Fulltext/2020/05000/Misinformation__A_Threat_to_the_Public_s_Health.15.aspx)

<sup>7</sup> Kristen Brown & Justin Sink, *Trump’s Comment on Disinfectant Prompts Experts to Warn Against Inhaling Bleach to Kill Coronavirus*, TIME, April 24, 2020, available at <https://time.com/5826882/coronavirus-trump-heat-bleach>

not only is it difficult to control sources of information, but also that the public is not equipped to filter what is true and what is false. The remaining point of control is then on the ISP or the intermediary, since the Internet is now the primary bearer of information.<sup>8</sup> This task necessarily imposes a burden upon content providers and ISPs to distribute reliable information. Further, any manner of regulation must be within the bounds of freedom of expression. This obvious responsibility, however, is left unchecked as current laws on intermediary liability do not address the new pivotal role of ISPs on public health. Intermediary liability with regard to misinformation on matters of public health must then be reassessed to address the urgent need for factual information on COVID-19.

## II. FREEDOM OF EXPRESSION AND THE RIGHT TO HEALTH

Freedom of speech and expression is enshrined in the Constitution of the Philippines. Section 4, Article III provides that “[n]o law shall be passed abridging the freedom of speech, of expression, or of the press, or the right of the people peaceably to assemble and petition the government for redress of grievances.”<sup>9</sup> The right of people to express their thoughts through words and action is thus heavily protected by the highest law of the land.

Several international legal instruments likewise embody freedom of expression. For one, the Universal Declaration of Human Rights (“UDHR”)—to which the Philippines is a signatory—contains several articles pertaining to its protection. In particular, Article 19 of the UDHR states that “[e]veryone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas *through any media* and regardless of frontiers.”<sup>10</sup> The inclusion of the words “through any media” makes it clearly applicable to expression through the Internet.<sup>11</sup>

---

<sup>8</sup> Ethan Shattock, *Is it time for Europe to reassess internet intermediary liability in light of coronavirus misinformation*, EUROPEAN LAW BLOG WEBSITE, Apr. 20, 2020, at <https://europeanlawblog.eu/2020/04/20/is-it-time-for-europe-to-reassess-internet-intermediary-liability-in-light-of-coronavirus-misinformation>

<sup>9</sup> CONST. art. III, § 4.

<sup>10</sup> United Nations General Assembly, Universal Declaration of Human Rights [hereinafter “UDHR”], art. 19, U.N. Doc. A/RES/217(III)A (Dec. 10, 1948). (Emphasis supplied.)

<sup>11</sup> CENTER FOR DEMOCRACY & TECHNOLOGY, “REGARDLESS OF FRONTIERS:” THE INTERNATIONAL RIGHT TO FREEDOM OF EXPRESSION IN THE DIGITAL AGE 3 (2011), *available at* [https://cdt.org/wp-content/uploads/pdfs/CDT-Regardless\\_of\\_Frontiers\\_v0.5.pdf](https://cdt.org/wp-content/uploads/pdfs/CDT-Regardless_of_Frontiers_v0.5.pdf)

Furthermore, Article 27<sup>12</sup> of the UDHR reinforces the right to seek, receive, and impart information embodied in Article 19. The right to “seek” information can be linked to browsing and searching the Internet through search engines and portals. The right to “impart” information, on the other hand, can be applied to blogging or posting information through social network sites or ISPs. Lastly, the right to “receive” information relates to the exchanging of e-mails, the reading of information through Web pages, and the downloading of information.<sup>13</sup>

While the UDHR is not a treaty, it has become a normative instrument that creates “some legal and moral obligations for Member States of the UN.”<sup>14</sup> In any case, the Philippines, as signatory, has also adopted its declarations through its Constitution’s Bill of Rights. As previously mentioned, Section 4 of Article III in particular protects the same freedom of expression embodied in the UDHR in that its dual aspects are recognized, namely freedom from censorship or prior restraint and freedom from subsequent punishment.<sup>15</sup>

Like any other right, however, freedom of speech also admits certain exceptions. The UDHR, for instance, provides that “everyone shall be subject only to such limitations as are determined by law solely for the purpose of securing due recognition and respect for the rights and freedoms of others and of meeting the just requirements of morality, public order and the general welfare in a democratic society.”<sup>16</sup> In the Philippine jurisdiction, valid government interference to freedom of expression may be allowed if the subject expression passes the clear and present danger rule,<sup>17</sup> the dangerous tendency rule,<sup>18</sup> or the balancing of interests rule,<sup>19</sup> whichever is applicable.

---

<sup>12</sup> UDHR, art. 27.

<sup>13</sup> Center for Democracy & Technology, *supra* note 11.

<sup>14</sup> *Id.*

<sup>15</sup> JOAQUIN BERNAS, *THE 1987 CONSTITUTION OF THE PHILIPPINES: A COMMENTARY* 248 (2003 ed.).

<sup>16</sup> UDHR, art. 29.

<sup>17</sup> *Cabansag v. Fernandez*, 102 Phil. 152 (1957). The clear and present danger rule inquires on whether words are used in such circumstance and of such nature as to create a clear and present danger that will bring about the substantive evil that the State has a right to prevent.

<sup>18</sup> *Id.* The dangerous tendency rule states that a person could be punished for words uttered or for ideas expressed which create a dangerous tendency, or which will cause or bring about a substantive evil which the State has a right to prevent.

<sup>19</sup> CARLO CRUZ & ISAGANI CRUZ, *CONSTITUTIONAL LAW* 522 (2015). The balancing of interests rule requires a Court to consider the circumstances in each particular case, and thereafter, it shall settle the issue of which right demands greater protection.

The right to health<sup>20</sup> is undoubtedly linked to the rights to freedom of expression and information.<sup>21</sup> This is especially apparent during the COVID-19 pandemic, where inaccurate or false information regarding the disease may lead to the loss of lives. It may then be argued that inaccurate statements with regard to public health may be categorized under any of the three aforementioned exceptions to freedom of expression. The United Nations (“UN”) Committee on Economic, Social and Cultural Rights has emphasized that “information accessibility” is a vital component of the right to health.<sup>22</sup> Inaccurate information involving public health is therefore not an ordinary kind of speech. There does remain, however, the legal question on whether such speech is protected. The current regime of intermediary liability in the Philippines may provide the solution in regulating such kind of speech. Applying the same, false and misleading information may then be regulated without necessarily suppressing freedom of speech.

### III. THE ROLE OF INTERMEDIARIES IN REGULATING COVID-19 FALSE CONTENT IN THE PHILIPPINES

#### A. Regulating Fake News in the Philippines

The proliferation of fake news is not a new phenomenon in the digital age. In the Philippines, an Anti-False Content Bill<sup>23</sup> was introduced in 2019, recognizing the increasing need to combat the existence of fake news. As of May 2020, the bill is currently pending in Congress.<sup>24</sup> Under this bill, a government agency would be made arbiter of permissible online material.<sup>25</sup> International group Human Rights Watch opposed the proposed law, saying

---

<sup>20</sup> CONST. art. III, § 15.

<sup>21</sup> Article 19, *Viral Lies: Misinformation and the Coronavirus*, at 9 (Mar. 2020), available at <https://www.article19.org/wp-content/uploads/2020/03/Coronavirus-briefing.pdf>. See Article 19, *A Healthy knowledge: Right to information and the right to health*, ARTICLE 19 WEBSITE, Sept. 27, 2012, at <https://www.article19.org/resources/healthy-knowledge-right-information-right-health>

<sup>22</sup> Committee on Economic, Social and Cultural Rights, General Comment No. 14: The Right to the Highest Attainable Standard of Health (Art. 12), ¶ 12(b), U.N. Doc. E/C.12/2000/4 (Aug. 11, 2000).

<sup>23</sup> S. No. 9, 18<sup>th</sup> Congress, 1<sup>st</sup> Sess. (2019).

<sup>24</sup> See *Anti-False Content Act (Leg. History)*, SENATE OF THE PHILIPPINES WEBSITE, at [https://senate.gov.ph/lis/bill\\_res.aspx?congress=18&q=SBN-9](https://senate.gov.ph/lis/bill_res.aspx?congress=18&q=SBN-9)

<sup>25</sup> S. No. 9, 18<sup>th</sup> Congress, 1<sup>st</sup> Sess., § 5 (2019).

that the Bill is “sweepingly broad and threatens to stifle discussion on websites worldwide” and “would excessively restrict online freedom of speech.”<sup>26</sup>

This opposition forwards the argument that false content is still protected speech. This is in line with the principle that protections under the right to freedom of expression are not limited to truthful statements or information.<sup>27</sup> Indeed, untruthful statements may likewise enjoy protection under the right to freedom of expression.<sup>28</sup>

False information about COVID-19 has been distributed more widely than information from authoritative sources such as the World Health Organization (“WHO”) and the Centers for Disease Control and Prevention.<sup>29</sup> The public health issue caused by COVID-19 thus necessarily brings to light the importance of stopping the proliferation of false information regarding the said disease.

The Philippine government recognized this need by including in the Bayanihan to Heal as One Act (R.A. No. 11469)<sup>30</sup> a provision penalizing the act of spreading false and alarming information.<sup>31</sup> Section 6(f) of the said law penalizes the following acts:

(f) Individuals or groups *creating, perpetrating, or spreading false information* regarding the COVID-19 crisis on social media and other platforms, such information having no valid or beneficial effect on the population, and are clearly geared to promote chaos, panic, anarchy, fear, or confusion; and those participating in cyber incidents that make use or take advantage of the current crisis situation to prey on the public through scams, phishing fraudulent emails, or other similar acts.<sup>32</sup>

While the Bayanihan to Heal as One Act is only valid for three months<sup>33</sup> or until June 24, 2020, this piece of legislation shows the stance of

---

<sup>26</sup> Human Rights Watch, *Philippines: Reject Sweeping ‘Fake News’ Bill*, HUMAN RIGHTS WATCH WEBSITE, July 25, 2019, at <https://www.hrw.org/news/2019/07/25/philippines-reject-sweeping-fake-news-bill/>

<sup>27</sup> Erwin Chemerinsky, *False Speech and the First Amendment*, 71 OKLA. L. REV. 1, 5-6 (2018).

<sup>28</sup> Article 19, *supra* note 21.

<sup>29</sup> John Gregory, *The coronavirus ‘infodemic’ is real. We rated the websites responsible for it*, STAT NEWS WEBSITE, Feb. 28, 2020, at <https://www.statnews.com/2020/02/28/websites-spreading-coronavirus-misinformation-infodemic>

<sup>30</sup> Rep. Act. No. 11469 (2020).

<sup>31</sup> § 6(f).

<sup>32</sup> § 6(f). (Emphasis supplied.)

<sup>33</sup> § 9.

lawmakers against the proliferation of false information. It is also worthy to note that the persons liable under this provision are those who create, perpetrate, or spread false information. The Author proposes that these provisions only apply to social media users and not to the ISP itself, given that ISPs are given a safe harbor<sup>34</sup> under the CPA.

Criticisms against the law emphasize that penalizing the acts in Section 6(f) would violate the constitutional right to free speech.<sup>35</sup> Moreover, there is no law that criminalizes or defines fake news as a crime, thereby making the provision vague and overbroad.<sup>36</sup> The acts of “perpetrating, or spreading false information” are also akin to the “aiding and abetting” provision<sup>37</sup> in the Cybercrime Prevention Act, a provision that has already been struck down by the Supreme Court for being unconstitutional.<sup>38</sup> It can therefore be seen how government regulation over content related to COVID-19 is fraught with enforcement and legal issues.

Policy-wise, ensuring that accurate and truthful information is available to the public also means that their sources must not be stifled into silence. It is unfortunate, however, that the supposed importance of stopping the spread of fake news has actually been used by governments all over the world as an excuse to craft repressive and overbroad laws to target misinformation concerning COVID-19.<sup>39</sup> Even worse, some states have begun to spread disinformation for the sake of propaganda. The increase in the scrutiny over government actions during the COVID-19 pandemic has

---

<sup>34</sup> Gemmo Fernandez & Raphael Lorenzo Pangalangan, *Spaces and Responsibilities: A Review of Foreign Laws and an Analysis of Philippine Laws on Intermediary Liability*, 89 PHIL. L.J. 761, 771-772 (2015). Under the safe harbor regime, intermediaries can only be held liable for defamatory or illegal content if they had knowledge that their platform contained illegal content.

<sup>35</sup> See Lian Buan, *Bayanihan Act's sanction vs 'false' info the 'most dangerous'*, RAPPLER, Mar. 29, 2020, at <https://www.rappler.com/nation/256256-sanctions-fake-news-bayanihan-act-most-dangerous>

<sup>36</sup> Lian Buan, *Duterte's special powers bill punishes fake news by jail time, up to P1-M fine*, RAPPLER, Mar. 24, 2020, at <https://www.rappler.com/nation/255753-duterte-special-powers-bill-coronavirus-fines-fake-news>

<sup>37</sup> Rep. Act. No. 10175 (2012), § 5.

<sup>38</sup> *Disini v. Sec'y of Justice*, G.R. No. 203335, 716 SCRA 237, Feb. 11, 2014.

<sup>39</sup> Article 19, *supra* note 21. See Article 19, Thailand: Computer Crime Act (Jan. 2017), available at <https://www.article19.org/data/files/medialibrary/38615/Analysis-Thailand-Computer-Crime-Act-31-Jan-17.pdf>; Article 19, Singapore: New law on “online falsehoods” a grave threat to freedom of expression, ARTICLE 19 WEBSITE, May 9, 2019, at <https://www.article19.org/resources/singapore-new-law-on-online-falsehoods-a-grave-threat-to-freedom-of-expression>; Article 19, Malaysia: Communications and Multimedia Act must be urgently revised, ARTICLE 19 WEBSITE, Mar. 24, 2017, at <https://www.article19.org/resources/malaysia-communications-and-multimedia-act-must-be-urgently-revised>

given states the incentive to control narratives and public perception.<sup>40</sup> It can be argued that the Bayanihan to Heal as One Act and the Anti-Terrorism Bill<sup>41</sup> are examples of such acts.

## **B. Intermediary Liability in Regulating Fake News**

It is apparent that governments themselves may be the source of false or misleading information regarding COVID-19. For instance, the governments of both the United States and China have been accused of spreading misleading characterizations of COVID-19, as evidenced by their own policy responses.<sup>42</sup> This kind of state-sponsored misinformation is particularly dangerous due to the influence of governments over its people. Further, it obliterates the trust in public authorities and instead encourages misguided responses by the public and health officials.<sup>43</sup> The proposed and existing laws<sup>44</sup> on prohibiting the spread of false information in the Philippines focus on having the government as arbiters of permissible information.<sup>45</sup> While it is undeniable that there is a need to regulate false information, problems will arise if the regulators themselves have a vested interest in limiting the available accurate information. The Author proposes that the most viable answer to these threats is hinged on utilizing the unique legal protections given to ISPs.

### *1. General Immunity under Section 230 of the Communications Decency Act*

In the United States (“U.S.”), intermediaries enjoy general immunity under Section 230 of the Communications Decency Act (“CDA”). The said provision has been the backbone of free speech on the Internet since its passage in 1996.<sup>46</sup> Section 230 was intended to protect computer “Good Samaritans” or those ISPs who make the effort to regulate its content. Also,

---

<sup>40</sup> Jane Dalton, *Mike Pompeo refuses to deny conspiracy theory that coronavirus is ‘hoax created to damage Trump,’* THE INDEPENDENT, Feb. 29, 2020, at <https://www.independent.co.uk/news/world/americas/coronavirus-mike-pompeo-trump-hoax-outbreak-us-cases-ted-lieu-a9366516.html>; James Palmer, *Beijing Knows Who to Blame for the Virus: America*, FOREIGN POLICY, Mar. 2, 2020, available at <https://foreignpolicy.com/2020/03/02/china-blames-united-states-coronavirus>

<sup>41</sup> As of writing, the bill has passed its third and final reading.

<sup>42</sup> Dalton, *supra* note 40.

<sup>43</sup> Article 19, *supra* note 21.

<sup>44</sup> These laws are the Bayanihan to Heal as One Act and the Anti-False Content Bill.

<sup>45</sup> S. No. 9, 18<sup>th</sup> Congress, 1<sup>st</sup> Sess., § 5 (2019).

<sup>46</sup> Andrew Bolson, *Flawed but Fixable: Section 230 of the Communications Decency Act at 20*, 42 RUTGERS COMPUT. & TECH. L.J. 1, 5 (2016).



the statute was intended to allow Internet companies to grow without the fear of crippling regulation.<sup>47</sup> The case of *Stratton Oakmont v. Prodigy Services Co.*<sup>48</sup> was the catalyst to this provision's passage. In that case, it became apparent that courts would punish websites that tried to screen out offensive content far more harshly than those websites that did not filter their content at all.<sup>49</sup> With the passage of Section 230, websites would be free from liability regardless of whether it failed to block enough content or for blocking too much content.<sup>50</sup>

Over the years, Section 230 faced criticism because it allegedly fostered a “do-nothing” approach among ISPs.<sup>51</sup> In other words, while the disincentive seen in *Stratton* may be eliminated, there remains no actual incentive to self-regulation by ISPs. Despite this, however, most ISPs have taken steps to regulate third-party content posted using their platforms. For instance, Facebook,<sup>52</sup> Google,<sup>53</sup> and Twitter<sup>54</sup> have their respective sets of content moderation policies.

The efforts to moderate content have increased over the years due to the rapid growth of social media platforms' reach and influence.<sup>55</sup> To grasp the magnitude of the Internet's effects to society, it would be useful to look at how “Facebook was used to spread misinformation, hate speech, and incitement to violence in the lead-up to and during the violence in Myanmar.”<sup>56</sup> In the Philippines, the effects of the Internet and social media were most prevalent during the 2016 presidential elections, which was widely considered as the first “social media election” in the Philippines.<sup>57</sup> Based on a

---

<sup>47</sup> 141 Cong. Rec. H8468 (daily ed. Aug. 4, 1995) (statement of Rep. Cox).

<sup>48</sup> 1995 WL 323710 (N.Y. Sup. Ct. May 24, 1995).

<sup>49</sup> Ryan French, *Picking up the Pieces: Finding Unity after the Communications Decency Act Section 230 Jurisprudential Clash*, 72 LA. L. REV. 443, 447 (2012).

<sup>50</sup> *Id.* at 450.

<sup>51</sup> Andrew Sevanian, *Section 230 of the Communications Decency Act: A “Good Samaritan” Law Without the Requirement of Acting as a “Good Samaritan,”* 21 UCLA ENT. L. REV. 121, 136 (2014).

<sup>52</sup> Facebook, Facebook Community Standards, FACEBOOK WEBSITE, at <https://www.facebook.com/communitystandards>

<sup>53</sup> Google, Google Help Communities Content Policy, GOOGLE COMMUNITIES HELP WEBSITE, at <https://support.google.com/communities/answer/7425194?hl=en>

<sup>54</sup> Twitter, The Twitter Rules, TWITTER HELP CENTER WEBSITE, at <https://help.twitter.com/en/rules-and-policies/twitter-rules>

<sup>55</sup> Terry Lee, *The global rise of “fake news” and the threat to democratic elections in the USA*, 22 PUB. ADMIN. & POLICY: AN ASIA-PAC. J. 15 (2019).

<sup>56</sup> Emma Irving, *Suppressing Atrocity Speech on Social Media*, 113 AJIL UNBOUND 256, 256 (2019).

<sup>57</sup> Aim Sinpeng, Dimitar Gueorguiev & Aries Arugay, *Strong Fans, Weak Campaign: Social Media and Duterte in the 2016 Presidential Election*, J. E. ASIAN STUD. 1 (2020).

quantitative and qualitative study, the winner of the 2016 elections had the most active, engaged, and networked advocates in social media.<sup>58</sup>

In the unique context of the COVID-19 pandemic, social media platforms have reinforced their content moderation policies, recognizing the heightened importance of suppressing inaccurate and misleading information. Last March 2020, Twitter revised its terms of service in order to stop the spread of virus-related misinformation, saying that it would remove posts if such went “against guidance from authoritative sources of global and public health information.”<sup>59</sup> Facebook has also taken steps in “ensuring that everyone has access to accurate information and removing harmful content”<sup>60</sup> by “connecting people to credible information on Facebook, Messenger, Instagram, and WhatsApp,”<sup>61</sup> by combating COVID-19 misinformation across its apps,<sup>62</sup> “investing \$100 million in the news industry and supporting fact-checkers,”<sup>63</sup> and “prohibiting exploitative tactics in ads and banning ads for medical face masks, hand sanitizer, disinfecting wipes, and COVID-19 test kits.”<sup>64</sup>

Despite the seemingly united effort in combating the dissemination of false and misleading information, social media companies still have varying stances on content moderation. Such differences are most apparent in their respective policies regarding the moderation of political speech. In May 2020, Twitter fact-checked tweets of U.S. President Donald Trump due to alleged violations of Twitter’s Civic Integrity Policy.<sup>65</sup> President Trump posted similar content on Facebook, but the platform held firm in its stance that it cannot be made “arbiters of truth” and that President Trump’s posts did not violate its own Community Standards.<sup>66</sup>

---

<sup>58</sup> *Id.*

<sup>59</sup> Vijaya Gadde & Matt Derella, *An update on our continuity strategy during COVID-19*, TWITTER BLOG WEBSITE, Mar. 16, 2020, at [https://blog.twitter.com/en\\_us/topics/company/2020/An-update-on-our-continuity-strategy-during-COVID-19.html](https://blog.twitter.com/en_us/topics/company/2020/An-update-on-our-continuity-strategy-during-COVID-19.html)

<sup>60</sup> Kang-Xing Jin, *Keeping People Safe and Informed About the Coronavirus*, ABOUT FACEBOOK WEBSITE, July 16, 2020, at <https://about.fb.com/news/2020/06/coronavirus/#misinformation-update>

<sup>61</sup> *Id.*

<sup>62</sup> *Id.*

<sup>63</sup> *Id.*

<sup>64</sup> *Id.*

<sup>65</sup> Twitter, *Civic integrity policy*, TWITTER HELP CENTER WEBSITE, May 2020, at <https://help.twitter.com/en/rules-and-policies/election-integrity-policy>

<sup>66</sup> Donie O’Sullivan, *Facebook and Twitter clash over fact-checking as Trump threats intensify*, CNN, May 28, 2020, at <https://edition.cnn.com/2020/05/28/media/jack-dorsey-donald-trump-twitter/index.html>

The similarities and differences in content moderation point to the fact that social media companies essentially have free rein in regulating its content due to the protection provided to them by Section 230 of the CDA. This protection is also why President Trump's response to Twitter's fact-checking was to invoke this particular section. Legally, social media platforms are under no obligation to write policies regulating content posted on its platform. The current global situation, however, points to the increasing need to moderate false and misleading information.

## *2. Safe Harbor under the Cybercrime Prevention Act*

In the Philippines, intermediary liability is regulated by the Cybercrime Prevention Act and its IRR. Under the said law, a service provider that “willfully abets or aids in the commission of any of the [offenses] enumerated in this Act shall be held liable.”<sup>67</sup> Liability is also imposed if service providers fail to “preserve computer data within a specified period”<sup>68</sup> or to “disclose such traffic data and subscriber information after being compelled to do so by authorities.”<sup>69</sup> The provisions on the liability of service providers seem straightforward in that the aiding and abetting must be done willfully in order to be punishable under the law.

The IRR of the Cybercrime Prevention Act, however, expounded on the liability of a service provider. Section 20<sup>70</sup> in particular provides a general immunity from liability for service providers, subject to three exceptions enumeration in paragraph (b). In fine, Section 20 provides the following:

*[N]o person or party shall be subject to any civil or criminal liability in respect of a computer data for which the person or party acting as a service provider merely provides access if such liability is founded on [...] [t]he making, publication, dissemination or distribution of such computer data or any statement made in such computer data, including possible infringement of any right subsisting in or in relation to such computer data: Provided, That:*

1. The service provider does not have actual knowledge, or is not aware of the facts or circumstances from which it is apparent, that the making, publication, dissemination or distribution of such material is unlawful or infringes any rights subsisting in or in relation to such material;

---

<sup>67</sup> Rep. Act. No. 10175 (2012), § 5(a).

<sup>68</sup> § 13(a).

<sup>69</sup> § 14(a).

<sup>70</sup> Rep. Act. No. 10175 (2012) Rules & Regs., § 20.

2. The service provider does not knowingly receive a financial benefit directly attributable to the unlawful or infringing activity; and
3. The service provider does not directly commit any infringement or other unlawful act, does not induce or cause another person or party to commit any infringement or other unlawful act, and/or does not directly benefit financially from the infringing activity or unlawful act of another person or party[.]<sup>71</sup>

Section 20 is a substantial reproduction of Section 30 of the Electronic Commerce Act of 2000.<sup>72</sup> It adheres to the safe harbor regime of intermediary liability, wherein intermediaries are only held liable for defamatory or illegal content if they had knowledge that their platform included content of such nature.<sup>73</sup> Safe harbor laws often include a “notice and takedown mechanism,” which requires intermediaries to remove or disable access to an illegal content upon receiving knowledge of its existence on the platform.<sup>74</sup>

In the context of the COVID-19 pandemic, the ISPs with the biggest reach in the Philippines are also Facebook, Google, and Twitter.<sup>75</sup> As discussed previously, these ISPs have already taken active steps in moderating false and misleading information on their platforms. The proposed and existing laws<sup>76</sup> on regulating fake news in the Philippines penalize the source but not the platform. Further, it is proposed by the Author that the safe harbor regime in the CPA also offers the same protection as that of Section 230 of the CDA. The only difference is in the three enumerated exceptions. In practice, this is already seen in how concerted efforts of ISPs in combating false information related to COVID-19 are also applied in the Philippines without any legal restrictions.

---

<sup>71</sup> § 20. (Emphasis supplied.)

<sup>72</sup> Rep. Act. No. 8792 (2000), § 30.

<sup>73</sup> Fernandez & Pangalangan, *supra* note 34, at 772.

<sup>74</sup> *Id.* at 773.

<sup>75</sup> Digital Marketing Philippines, *Comprehensive Look on the Top 7 Social Media Platforms*, DIGITAL MARKETING PHILIPPINES WEBSITE, at <https://digitalmarketingphilippines.com/comprehensive-look-on-the-top-7-social-media-platforms> (last accessed June 3, 2020).

<sup>76</sup> These laws are the Bayanihan to Heal as One Act and the Anti-False Content Bill.

### 3. Utilizing the Legal Safe Harbor

The main criticism to Twitter's fact-checking of President Trump's tweets centers on its possible suppression of conservative speech.<sup>77</sup> Facebook also emphasized its stance that social media companies cannot be made "arbiters of the truth" and that people should be left to decide for themselves on how they would interpret content available on the platform.<sup>78</sup> Facebook then made a distinction in its approach in combating false and misleading information regarding COVID-19 from its moderation of political speech. In the current global pandemic, however, political speech and COVID-19 information are not entirely separable. Governments control the narrative in the efforts against COVID-19. Meanwhile, social media platforms are not only the most accessible sources of information to the public, but are also the only places where public conversation may occur in this era of social distancing. Social media companies may thus be the only arbiter available, given the unique protection provided by law.

Enforcement of an all-encompassing form of content moderation, however, may pose serious problems. Twitter has already been accused of being partisan because it singled out President Trump.<sup>79</sup> It is worthy to note, however, that the content moderation method employed by Twitter in fact-checking the President did not actually suppress speech. Twitter did not delete the tweets. Instead, it provided additional sources of information that may be used by its users in discerning the truthfulness of the tweets. This method of content moderation is akin to the "right of reply," which is a proposed amendment to Section 230 of the CDA.<sup>80</sup> In utilizing the proposed right of reply, the original statement will remain accessible to Internet users who may then consider both the original content and the reply.<sup>81</sup>

Twitter's fact-checking is therefore an act protected by Section 230 of the CDA. If such content moderation was done to Filipino speech, it would also be protected by the CPA's safe harbor provision considering that such form of content moderation is not within the three exceptions<sup>82</sup> to general immunity. Hence, it is an act within the bounds of law. The partiality or

---

<sup>77</sup> Queenie Wong, *Twitter faces conservative backlash for fact-checking Trump's tweets for the first time*, CNET, May 27, 2020, at <https://www.cnet.com/news/twitter-faces-conservative-backlash-for-fact-checking-trumps-tweets-for-the-first-time>

<sup>78</sup> O'Sullivan, *supra* note 66.

<sup>79</sup> Wong, *supra* note 77.

<sup>80</sup> Michael Scott, *Would a "Right of Reply" Fix Section 230 of the Communications Decency Act*, 4 J. INT'L MEDIA & ENT. L. 57, 67 (2011).

<sup>81</sup> *Id.* at 67.

<sup>82</sup> Rep. Act. No. 10175 (2012) Rules & Regs. § 20(b).

impartiality of such an act, however, is a pressing issue that remains unresolved and must be addressed by the social media companies. In addition, social media companies have the responsibility to be transparent and accountable. This idea is in line with the Manila Principles, with the fifth Principle stating that “[l]aws and content restriction policies and practices must respect due process”<sup>83</sup> and the sixth Principle stating that “[t]ransparency and accountability must be built into laws and content restriction policies and practices.”<sup>84</sup>

#### IV. CONCLUSION

The immunity from liability of ISPs may give them armor that is critical in finding the balance between protecting the fundamental freedom of speech and preventing the proliferation of inaccurate or false information detrimental to public health. Such immunity is present in Section 230 of the U.S. Communications Decency Act and in the Cybercrime Prevention Act of the Philippines. ISPs must take a more active role in this global pandemic by utilizing the safe harbor provision in current intermediary liability laws. It is important to remember, however, that content regulation must always be within the bounds of freedom of expression. False information is not necessarily unprotected speech.

The free rein given to ISPs may result in different techniques in content moderation. It is also evident that ambivalence or a hands-off approach is the easiest path for ISPs to take. The global situation now, however, transcends the issue of liability. It is now a matter of responsibility.

- o0o -

---

<sup>83</sup> Electronic Frontier Foundation, Manila Principles on Intermediary Liability (Mar. 24, 2015), at [www.manilaprinciples.org/principles](http://www.manilaprinciples.org/principles)

<sup>84</sup> *Id.*