Arizona Law Journal of Emerging Technologies

Volume 6 *Article 4*

4-2023

REVISITING DISINFORMATION LAWS IN THE AGE OF SOCIAL MEDIA

Cheng-Chi (Kirin) Chang
Tsinghua University

2022-2023 **https://azlawjet.com** **Vol. 6, Art. 4**

Arizona Law Journal of Emerging Technologies

REVISITING DISINFORMATION LAWS IN THE AGE OF SOCIAL MEDIA

Cheng-Chi (Kirin) Chang, Cyber & Data Rule of Law Research Assistant

Table of Contents

REVISITING DISINFORMATION LAWS IN THE AGE OF SOCIAL MEDIA

Cheng-Chi (Kirin) Chang*

I. Abstract

Misinformation on social media is a significant issue today, necessitating a thorough analysis of the problem and potential solutions. To protect free speech while holding those who spread fake news accountable, the US legal system must make certain adjustments. While the First Amendment safeguards online freedom of expression, exceptions exist. Social media firms are exempt from responsibility under § 230 of the Communications Decency Act of 1996 (CDA) since they only host, rather than create, false news. These firms may withdraw content without legal repercussions because of this immunity, which creates the possibility of censorship. To obtain a comprehensive understanding of how social media-driven disinformation is handled, it is also important to compare the US legal model to those of the European Union and certain Asian nations. This essay will investigate the issue of fake news disseminated through social media and propose solutions to the issue while also recommending modifications to misinformation laws to improve their efficacy and prevent censorship risks.

II. Introduction

In recent years, policymakers and lawyers have given increased attention and scrutiny to disinformation on social media, due to the potentially drastic consequences of false information about political, economic, and medical topics. This topic has gained importance since the COVID-19 pandemic because of the potential for illness and death to spread unnecessarily due to bad information from social media sources.[1] This essay will explore how the legal systems in the United States, the European Union, and Asian countries deal with disinformation on social media, and how they must do so without sliding into censorship or political repression.

Another issue countries must face involves a prominent case of social media disinformation in the news. In June 2022, entrepreneur Elon Musk canceled his proposed $44 billion takeover of social media giant Twitter, accusing the company of hiding the true amount of "bot" and "spam" accounts.[2] Twitter refused to release information necessary to analyze the number of these

* Cyber & Data Rule of Law Research Assistant, Institute for Studies on AI and Law, Tsinghua University; J.D. Candidate, University of Florida Levin College of Law; LL.M., University of Arizona James E. Rogers College of Law; LL.B., National Chung Hsing University. Thanks to Michael Sayle and the fellow editors at the Arizona Law Journal of Emerging Technologies for remarkable feedback throughout the entire editing process.

[1] Samuli Laato et al., *What drives unverified information sharing and cyberchondria during the COVID-19 pandemic?*, 29 EUR. J. INF. SYST. 288 (2020).

[2] Irina Ivanova, *Elon Musk threatens to cancel Twitter buyout*, CBSNEWS.COM (2022), https://www.cbsnews.com/news/elon-musk-threatens-to-cancel-twitter-buyout/ [https://perma.cc/K8AP-9PTQ].

accounts, causing Musk to allege bad faith and a material breach on Twitter's part.[3] Musk stated that fake accounts perpetrating scams and false information are a problem for Twitter users and for advertisers who use the platform to reach out to potential consumers.[4] Elon Musk canceling the Twitter buyout is a real-world example of the importance of disinformation on social media, as the prevalence of disinformation and fraudulent accounts on a prominent social media platform was a major issue in negotiations between the parties.[5] This essay will examine how legal systems throughout the world police disinformation on social media so that problems with fake accounts such as that alleged by Musk during the Twitter deal are adequately dealt with.

III. United States Laws on Disinformation

a. *First Amendment Protection*

Protection of the freedom of speech has long been one of the bedrock protections afforded to American citizens by the First Amendment of the U.S. Constitution. This is because people must have the means to form their own opinions about public policy for a democratic government to properly respond to public opinion.[6] However, with the rise of mass media, including social media, and the decline of traditional media outlets such as newspapers, "fake news" has begun to be a problem, with the dissemination of false information skyrocketing in recent years.[7] One of the problems of social media disinformation is that the spread of this disinformation undermines the ability of the electorate to choose qualified candidates for political office.[8] Research on misinformation has shown that fake news peaked on Twitter during the 2012 and 2016 presidential elections in the United States, and a bipartisan U.S. Senate committee found that the Russian government used social media platforms such as Facebook, Instagram, and Twitter to spread conspiracy theories and fake news and to sow chaos in the United States.[9] A single Russian firm with fewer than a hundred agents generated fake news content that reached over 150 million Facebook users on the Russian government's behalf.[10] A 2018 study found that false rumors spread faster and wider than factual information, with false statements being 70 percent more likely to be retweeted on Twitter and to reach 1,500 people six times faster.[11]

i. Political Misinformation on Social Media

Political misinformation is nothing new, however, social media has made its dissemination much easier and faster. Social media allows politicians and political groups to overwhelm users and disrupt their sense of reality by disseminating news and political ads at unprecedented speeds,

---

[3] *Id.*
[4] *Id.*
[5] Stacy Rosenberg, *Publics in Emerging Economies Worry Social Media Sow Division, Even as They Offer New Chances for Political Engagement*, PEW RESEARCH CENTER: INTERNET, SCIENCE & TECH (2019), https://www.pewresearch.org/internet/2019/05/13/publics-in-emerging-economies-worry-social-media-sow-division-even-as-they-offer-new-chances-for-political-engagement/ (last visited Feb 27, 2023).
[6] Daniela Manzi, *Managing the Misinformation Marketplace: The First Amendment and the Fight Against Fake News*, 87 FORDHAM LAW REV. 2623, 2627 (2019).
[7] *Id.*
[8] *Id.* at 2627–2628.
[9] Sara Brown, *MIT Sloan research about social media, misinformation, and elections*, MIT SLOAN SCHOOL OF MANAGEMENT (2020), https://mitsloan.mit.edu/ideas-made-to-matter/mit-sloan-research-about-social-media-misinformation-and-elections [https://perma.cc/Z2VG-97V6].
[10] Manzi, *supra* note 6 at 2630.
[11] Brown, *supra* note 9.

with extreme or outrageous stories more likely to be widely shared on social media.[12] Studies have found that those who share false information are more likely to be distracted or lazy, instead of biased or actively malicious, and in some cases, people appreciate a candidate who tells obvious lies, with the candidate being seen as "authentic."[13] For example, in the 2016 presidential election, false pro-Trump articles were shared 30.3 million times and false pro-Clinton articles were shared 7.6 million times on Facebook.[14] In addition, between September 16 and October 21 of 2016, about 20 percent of all tweets about the presidential election were generated by bot accounts.[15] Disinformation on social media is harmful to the political system because it undermines people's trust in the media, and thus enables the government to operate without effective scrutiny by the news media.[16]

ii. Bot Accounts

Bot accounts operate through a process called "flooding," in which they distribute the same false stories through diverse sources to trick readers into believing that the story is widely accepted, causing the viewers to not know what to believe and to shun both credible and non-credible news sources.[17] These bot accounts take advantage of social media algorithms to promote content favored by their viewers, creating echo chambers where users are exposed to the same information constantly, with no exposure to any contradictory information.[18] In addition, politicians can use social media platforms to spread false information about their opponents without being corrected by legitimate media sources.[19] Fake news on social media platforms is a major problem in the United States today, distorting the political discourse by allowing lies to flourish.

iii. False Speech, First Amendment and the Marketplace of Ideas

The problem is compounded by the fact that false speech is protected by the First Amendment, which prohibits Congress from making any law "abridging the freedom of speech."[20] The Supreme Court has further held that false statements of fact have no constitutional value, stating in *Hustler Magazine, Inc. v. Falwell* that "false statements of fact are particularly valueless because they interfere with the truth-seeking function of the marketplace of ideas," and in *Gertz v. Robert Welch, Inc.* that "neither the intentional lie nor the careless error materially advances society's interest in 'uninhibited, robust, and wide-open' debate on public issues."[21] The Supreme Court's most recent case dealing with false statements comes in *United States v. Alvarez*, in which the Court struck down a law criminalizing lying about receiving military medals, holding that a statement's falsity does not put it outside the realm of First Amendment

[12] Manzi, *supra* note 6 at 2628.
[13] Brown, *supra* note 9.
[14] Manzi, *supra* note 6 at 2628.
[15] *Id.* at 2630.
[16] *Id.* at 2632.
[17] *Id.* at 2630.
[18] *Id.* at 2631.
[19] *Id.*
[20] *Id.* at 2627 (quoting U.S. Const. Amend. I).
[21] *Id.* at 2634; 485 U.S. (1988); 418 U.S. 323 (1974).

protection.[22] The Court ruled that the Act could not be constitutionally justified because the government did not prove that the false claims of military valor caused provable harm and that the law would create a dangerous precedent for an overly broad regulation of false speech without judicial oversight.[23] The Court stated that this broad regulation of falsehood would lead to censorship and suppression of free speech because people would fear selective prosecution by the government.[24] The solution to this problem, according to the Supreme Court, is that the decision about truth should be hashed out in the "marketplace of ideas."[25] False statements may be valuable for allowing people to challenge widely held beliefs without fear of potential prosecution.[26]

iv. The Act of Liking and Sharing on Social Media

The act of liking and sharing information on social media platforms such as Facebook and Twitter constitutes a form of free speech protected by the First Amendment. A "like" on a platform such as Facebook is a way for a user to acknowledge content that he or she finds relevant and interesting with his or her friends on the platform.[27] Similarly, a "share" on a social media platform allows for content to be posted on the user's wall, as opposed to a "like" that adds the content to the interests section of the user's profile. [28] Generally, the First Amendment protects social media posts, however, the social media companies themselves have the power to police content shared on their platforms.[29] This is because courts have found social media companies to be private actors who are not performing the same functions as governmental entities.[30] However, this is controversial because many have argued that social media companies should not exercise such control over the content on their platforms because they function as a "public forum" where free speech cannot be regulated under the First Amendment.[31]

A July 2019 study found that Facebook's like and share features could be targeted for the spread of misinformation through "invite whales," people who sent invitations out to others to join a private group.[32] These individuals would then spam the groups with posts featuring harmful or misleading content, and in many cases, pages with tens of thousands of followers were sold by

[22] Sharon Beckstrand, *Does Free Speech Protect COVID-19 Vaccine Misinformation?*, STAN. L. SCH.: LEGAL AGGREGATE (2022), https://law.stanford.edu/2022/04/22/does-free-speech-protect-covid-19-vaccine-misinformation/ [https://perma.cc/8PHZ-ATL8]; 567 U.S. 709 (2012).
[23] Manzi, *supra* note 6, at 2634.
[24] *Id.* at 2635.
[25] Beckstrand, *supra* note 22.
[26] Id.
[27] TJ McCue, *Facebook Likes and Shares: What They Mean for Your Website*, AMERICAN EXPRESS (2011), https://www.americanexpress.com/en-us/business/trends-and-insights/articles/facebook-likes-and-shares-what-they-mean-for-your-website-1/ [https://perma.cc/EXQ5-5CT8].
[28] First Digital, *Facebook "Likes", "Shares" and "Like With Comments". Confused?*, FIRST DIGITAL (2011), https://firstdigital.co.nz/blog/social-media-marketing/facebook-likes-shares-and-like-with-comments-confused/.
[29] Brett Pinkus, *The Limits of Free Speech in Social Media*, ACCESSIBLE LAW (2021), https://untdlaw.wixsite.com/accessible-law/post/the-limits-of-free-speech-in-social-media.
[30] *Id.*
[31] *Id.*
[32] Mike Isaac, *Facebook Wrestles With the Features It Used to Define Social Networking*, N.Y. TIMES, (Oct. 25, 2021), https://www.nytimes.com/2021/10/25/technology/facebook-like-share-buttons.html [https://perma.cc/P4Z2-JGMV].

the founders, and the buyers used the page to show misinformation of politically inflammatory content.[33] Facebook has also experimented with limiting the use of the like and share buttons to stem the spread of misinformation on the site, but it did not remove the features entirely.[34] Overall, the like and share buttons made users three times more likely to share any kind of content from reshare aggregators, which led to fears that bad content such as bullying, nudity, or "hate bait," content promoting racial, religious, or ethnic violence would be more aggressively disseminated through Facebook.[35]

Overall, "liking" and "sharing" posts on Facebook and other social media are forms of speech protected by the First Amendment. This means that policing misinformation is more difficult because people's opinions on important subjects often differ, and it is not always clear that a news or political group or page is spreading disinformation or merely their own opinion. However, social media companies themselves can remove content that features misinformation or posts inciting hatred or discrimination, which is not protected by the First Amendment.[36] The United States court system has repeatedly refused to consider social media platforms a public forum where restrictions on speech are impermissible under the First Amendment.[37] The Supreme Court ruled in *Heffron v. Int'l Soc. for Krishna Consciousness, Inc.* that the First Amendment "does not guarantee the right to communicate one's views at all times and places or in any manner that may be desired."[38] Therefore, the best solution to the problem of misinformation is policing by social media platforms themselves.

b. Legal Punishments for Spreading Misinformation

Under United States law, there are a variety of ways to punish those who spread misinformation through social media. Posters may be criminally liable through defamation, and they may also be civilly liable for violating other rights as well, such as the right to privacy. In addition, they may be held liable under copyright laws and the California Conversation Robot Disclosure Act. The United States legal system has multiple avenues to punish those who spread misinformation on the internet. It's worth noting that the U.S. government and legislature began studying how to regulate the problems caused by "DEEP FAKES" in 2019, but there has been no progress in developing relevant legal statutes.[39]

i. Defamation

One aspect of criminal liability for misinformation concerns the criminal act of defamation. Defamation consists of two separate categories: slander, or spoken words, and libel, or written or otherwise permanently fixed words.[40] Defamation has three key elements: an individual whose

[33] *Id.*
[34] *Id.*
[35] *Id.*
[36] Pinkus, *supra* note 29.
[37] *Id.*
[38] *Id.*; 452 U.S. 640, 647 (1981).
[39] Tiffany Hsu, *As Deepfakes Flourish, Countries Struggle With Response*, N.Y. TIMES, (Jan. 22, 2023), https://www.nytimes.com/2023/01/22/business/media/deepfake-regulation-difficulty.html.
[40] Tommaso Tani, *Legal Responsibility For False News*, 8 J INTL MEDIA ENT L 229, 261 (2020).

reputation is harmed, an actor, and a statement that causes damage to the first individual's reputation.[41] False news steps over the line of mere misinformation into defamation if it names a specific person and does significant reputational damage.[42] For example, during the 2016 presidential campaign, a conspiracy theory circulated on social media that presidential candidate Hillary Clinton and other prominent Democrats were coordinating a child trafficking ring out of a Washington D.C. pizza restaurant called Comet Ping Pong.[43] In another example, 20th Century Fox worked with a fake news publisher to create websites imitating traditional online news sources that published false information about public figures and controversial political and public interest subjects to publicize the film "A Cure for Wellness."[44] However, defamation is only a criminal act in 15 states, a remnant of 17th century British common law, and criminal libel charges have been dropped or reduced when the defendants questioned the charges' constitutionality.[45] The key for misinformation to be labeled as defamation is a recognizable victim and reputational damage to said victim.[46]

ii. False Reporting Laws

Another way that those who spread misinformation online may be held criminally liable is through false reporting laws. These statutes impose criminal liability on anyone who initiates or circulates a report or warning of an impending crime or catastrophe with the knowledge that the report/warning is false, and that it is likely to cause public alarm or inconvenience.[47] For example, New York has a statute that prohibits circulating reports of emergencies or natural disasters that the speaker knows to be false or baseless that are "not unlikely" to cause "public alarm or inconvenience."[48] The statute imposes a penalty of one year in prison and a $1,000 fine.[49] Courts have generally upheld convictions under such statutes. For example, the Ohio Supreme Court upheld the 2012 conviction for "inducing panic" of a teenager who called for "another mass murder" in the wake of the Sandy Hook school shooting.[50] Generally, false reporting statutes have not been used in relation to social media yet, but there is a strong probability that they will given the prominence of social media. As an example, the New York state legislature has introduced a bill to increase the severity of the offenses in the false reporting statute, caused by the possibility of unique harm of disinformation spread by social media.[51]

iii.Deep Fakes and Law

"Deep Fakes" are fake videos or images created using AI-powered software that can superimpose someone's face onto another person's body, making it look authentic. They can be used to

[41] *Id.*
[42] *Id.* at 258–260.
[43] *Id.* at 262.
[44] David Klein & Joshua Wueller, *Fake News: A Legal Perspective*, J. INTERNET LAW 9, 1, 6 (2017).
[45] *Id.* at 9.
[46] Tani, *supra* note 40, at 261–262.
[47] Louis W. Tompros et al., *The Constitutionality of Criminalizing False Speech Made on Social Networking Sites in a Post-Alvarez, Social Media-Obsessed World*, 31 HARV. J. LAW TECHNOL. HARV. JOLT 65, 83 (2017).
[48] *Id.* at 69.
[49] *Id.*
[50] *Id.* at 82.
[51] *Id.*

manipulate public perception, spread false information, and cause social unrest.[52] This concern started when users on the social media website Reddit began posting pornographic clips with celebrities' faces imposed on performers' bodies, and many expressed concerns that this process could influence the democratic process by manipulating candidates' images.[53] There are concerns that the use of "Deep Fakes" may increase in the future, and there are currently no specific laws or regulations in place to address this issue.[54] Former President Trump signed a law in 2019 that requires a report on foreign use of deep fakes, and for the government to notify Congress if foreign deep fake disinformation targets US elections. A "Deepfakes Prize" competition was also established to encourage the development of deep fake detection technologies.[55] Some lawmakers have introduced bills to address these concerns, but none have yet become law. Senator Rob Portman introduced a bill called the Deepfake Report Act of 2019, which would require the Department of Homeland Security to report on the development of deep fake technology, but the House has not passed it.[56] Representative Yvette Clarke introduced the DEEP FAKES Accountability Act in 2021, which would require individuals to disclose if they post an altered deep fake, and imposed fines and imprisonment if the post was intended to harm or commit fraud. However, this act has not been passed.[57]

Although the Deep Fakes Accountability Act did not pass, it still provides valuable insights into the use of deep fake content. The act proposed federal restrictions on the distribution of deep fakes and criminal liability for their creators and distributors.[58] The act required that any deep fake content intended for sharing contain a disclaimer that the depicted content was false.[59] The act made sharing pornographic content, content that incites violence, content that interferes with an official proceeding such as an election, content that contains content amounting to foreign interference in domestic affairs, or contact which aids in an act of fraud subject to criminal liability.[60] The act also provided that the person depicted in a deep fake video would receive at least $50,000 per deep fake image shared or altered, and courts could issue an injunction for a disclaimer to be added to the content.[61]

It excluded government-created records, parodies, and film or television productions where the person depicted had given their consent. Additionally, the act created a task force within the Department of Homeland Security to find new deep fake detection technologies to share with

[52] Rebecca Delfino, *Deepfakes on Trial: A Call To Expand the Trial Judge's Gatekeeping Role To Protect Legal Proceedings from Technological Fakery*, 74 HASTINGS LAW J. 293, 298 (2023).
[53] Danielle Citron & Robert Chesney, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CALIF. LAW REV. 1753, 1763 (2019).
[54] Hsu, *supra* note 39.
[55] Natalie Lussier, *NONCONSENSUAL DEEPFAKES: DETECTING AND REGULATING THE RISING THREAT TO PRIVACY*, 58 IDA. LAW REV., 366 (2022), https://digitalcommons.law.uidaho.edu/idaho-law-review/vol58/iss2/6.
[56] S. Rept. 116-93 - DEEPFAKE REPORT ACT OF 2019, http://www.congress.gov/ (last visited Feb 28, 2023).
[57] Yvette D. [D-NY-9 Rep. Clarke, *Text - H.R.2395 - 117th Congress (2021-2022): DEEP FAKES Accountability Act*, (2021), http://www.congress.gov/ (last visited Feb 28, 2023).
[58] Zachary Schapiro, *Deep Fakes Accountability Act: Overbroad and Ineffective*, INTELLECT. PROP. TECHNOL. FORUM 16 (2020).
[59] *Id.*
[60] *Id.* at 3.
[61] *Id.*

social media companies.[62] The Deep Fakes Accountability Act sought to counter defamation by requiring deep fakes to cause perceptible individual or societal harm. Although its effectiveness in combating disinformation on social media is uncertain, it highlighted the need for legal remedies for the spread of disinformation, such as copyright law, criminal liability for revenge porn, or civil law claims for defamation. [63] Thus, referring to the content of the act remains relevant for understanding the complex legal and ethical issues surrounding deep fakes.

iv. Civil Liability

In addition to criminal liability for spreading misinformation on social media, there is the option of civil liability against offenders. For particular people who have suffered reputational damage as a result of false information, the tort of defamation is available because the act in question involves communicating false statements to the public, and courts in the United States have been willing to apply defamation law to online statements.[64] Under defamation law, anybody who repeated the defamatory statement may be held liable, in addition to the person who originated it, a concept known as the "republication rule."[65] However, the First Amendment's free speech guarantee means that the plaintiff in such cases, if he or she is a public figure, must prove that the defendant acted with "actual malice," and must prove that the defendant acted negligently if the plaintiff is a private individual.[66] In addition, states have enacted anti-SLAPP (Strategic Lawsuits Against Public Participation) laws that allow defendants to file for early dismissal of such cases if they can prove that their speech was a lawful exercise of their First Amendment rights.[67] These laws protect against frivolous lawsuits based on intentional or negligent infliction of emotional distress, as well as other common law claims that would limit speech on matters of public interest.[68]

Intentional infliction of emotional distress is another tort impacted by the spread of false information on social media. Intentional infliction of emotional distress occurs when a defendant engages in extreme or outrageous behavior that causes another person to suffer severe emotional distress.[69] However, under United States law, there is a high burden for a plaintiff to clear for an intentional infliction of distress, as the statement must be published with "actual malice."[70] For example, the Supreme Court ruled in *Hustler v. Falwell* that statements that are parodies that are not reasonably believable are not published with actual malice.[71] However, fake news publications that publish extreme content may be held liable under intentional infliction of emotional distress. For example, in 2014, a federal district court in Virginia found that a man who took nude photographs of an aspiring male model and photoshopped sexual elements into

[62] *Id.*
[63] *Id.* at 5.
[64] *Id.* at 8.
[65] What Legal Recourse Do Victims Of Fake News Stories Have?, NPR, Dec. 7, 2016, https://www.npr.org/2016/12/07/504723649/what-legal-recourse-do-victims-of-fake-news-stories-have [https://perma.cc/PEH3-ZA92].
[66] Klein and Wueller, *supra* note 44.
[67] *Id.* at 7.
[68] *Id.*
[69] *Id.*
[70] *Id.* at 8.
[71] 485 U.S. 46, 56-57 (1988).

the photographs and published them on various websites explicitly naming the male model was liable for intentional infliction of emotional distress.[72]

v. Copyright Law

Tort law offers another way to hold disinformation purveyors accountable, and intellectual property rights are a key aspect. Copyright law, in particular, can be useful in combatting the spread of "deep fakes" by utilizing the fair use doctrine and the nature of a person's likeness.[73] Copyright can protect the visual components used in creating a deep fake that belong to the original creators.[74] However, even though someone's likeness has been used to create a deep fake, if there is no agreement in place with the photographer or videographer, the subject depicted cannot claim copyright infringement. The person who produced the material holds the copyright and is the rightful owner unless they have specifically given permission to use their work. This means that the subject in the deep fake cannot claim ownership of the material used to create the deep fake.[75] In addition, when a court is reviewing a case involving possible copyright infringement, it is crucial to consider fair use. To determine whether the use of copyrighted material is legal, the court must consider four critical factors: the purpose and character of the use, the type of copyrighted work, the amount of the material used in comparison to the whole work, and the potential impact on the future market for the copyrighted work. This analysis is known as fair use.[76] The question of whether producers of deep fake content can be held accountable for copyright violations hinges on the nature and intention of the use, and its potential impact on the market. Utilization of deep fakes for political aims may violate copyright protection laws, while applications within the domains of scholarship or comedy, for example, are more likely to be considered permissible.[77]

Furthermore, it's important to note that both federal and state laws, such as the Lanham Act and unfair competition laws, prohibit the unauthorized use of trademarks and false representations of the quality or nature of someone else's products, services, or business activities. This implies that publishers of misleading news stories who utilize third-party brands for endorsements or promotions may be subject to legal action.[78] Creators of written text, photographs, artwork, and other original works of authorship have exclusive rights under federal copyright law to reproduce, distribute, display, and create derivative works from their works of authorship, so permission from the content owners is necessary to avoid copyright infringement claims.[79] Intellectual property rights also exist for a person’s name and likeness, and 47 states have acknowledged a “right of publicity” giving individuals the right to control the commercial use of their public images.[80]

[72] Klein and Wueller, *supra* note 44, at 8.
[73] Schapiro, *supra* note 58, at 12.
[74] *Id.* at 12–13.
[75] *Id.* at 13.
[76] *Id.*
[77] *Id.*
[78] Klein and Wueller, *supra* note 44, at 8.
[79] *Id.*
[80] *Id.*

vi. B.O.T. Bill

Publishers of false news can face legal action for various reasons, including defamation, violation of intellectual property rights, and intentional infliction of emotional distress. Additionally, they may be held accountable for false light invasions of privacy, fraud, tortious interference, or unfair and deceptive trade practices, depending on the specifics of each case.[81] California has taken a new approach to combat fake news on social media by targeting "bots." The California legislature passed the Bolstering Online Transparency bill, or B.O.T. bill, which requires all bots attempting to influence California citizens' voting or purchasing behaviors to conspicuously label themselves. The bill came into effect in July 2019.[82] This law assigns the task of labeling bot accounts entirely to their owner or creator, and the platform is not held accountable for this responsibility.[83] "Bot" accounts have been used to mislead social media users, artificially inflate follower counts, likes, and retweets, and manufacture a false consensus on controversial political issues.[84] In 2015 and 2016, Twitter had a significant number of bot accounts. These accounts were frequently utilized to make certain topics appear more popular than they were, leading to the manipulation of information and the creation of a false impression that some news stories were more important than they were.[85]

However, the California law has some flaws. The law is ambiguous because it fails to adequately define what a "bot" is, meaning that Twitter bots and customer service chatbots could be lumped together in the same category, and the law does not provide a definition for the term "substantially."[86] This is an issue because such bot accounts are often automated for a certain period of time so that the account can be "aged" so that it does not appear to have been created for a particular purpose, and the account is taken over by a human operator for an influence operation.[87] Russia's Internet Research Agency accounts, which are responsible for spreading a great deal of misinformation on social media and are often credited with interfering in the 2016 presidential election, operate in this manner.[88] In addition, the law fails to address what exactly constitutes "influencing a vote in an election," leading to the possibility that sharing legitimate news stories or voting locations may be classified as improper "influencing."[89]

Another issue with the law is that it does not mandate platform responsibility, such as a "bot" badge similar to Twitter's "verified" checkmark assigned to accounts on sites such as Facebook and Twitter.[90] Many feared that such frameworks would not be feasible for startups and small

---

[81] *Id.* at 7–8.
[82] Renee DiResta, *A New Law Makes Bots Identify Themselves—That's the Problem*, WIRED, https://www.wired.com/story/law-makes-bots-identify-themselves/ [https://perma.cc/RT8B-RCNV].
[83] *Id.*
[84] *Id.*
[85] *Id.*; Amelia M. Jamison, David A. Broniatowski & Sandra Crouse Quinn, *Malicious Actors on Twitter: A Guide for Public Health Researchers*, 109 AM. J. PUBLIC HEALTH 688 (2019).
[86] DiResta, *supra* note 82.
[87] *Id.*
[88] *Id.*; Massimo Stella, Emilio Ferrara & Manlio De Domenico, *Bots increase exposure to negative and inflammatory content in online social systems*, 115 PROC. NATL. ACAD. SCI. 12435, 115 (2018); Hatim Rahman, *Why Are Social Media Platforms Still So Bad at Combating Misinformation?*, KELLOGG INSIGHT (2020), https://insight.kellogg.northwestern.edu/article/social-media-platforms-combating-misinformation.
[89] DiResta, *supra* note 82.
[90] *Id.*

platforms to implement and that only large platforms such as Twitter could successfully implement these frameworks.[91] Thus, the responsibility of labeling the account falls solely on the account owner, and critics have noted that it is unrealistic to expect that such account owners voluntarily identify themselves.[92]

Finally, the California law has ambiguous enforcement mechanisms, with the state government lacking the ability to effectively identify bots, meaning that the overly-broad language of the statute could give steep fines of up to $2,500 per violation to small sites using automated marketing tools such as chatbots. This will lead to California penalizing law-abiding businesses.[93] California's Robot Disclosure Act is an imperfect law because it does not effectively define key terms and it lacks a mechanism for enforcement. However, it is a step in the direction to combat the spread of disinformation through social media sources.
Overall, one key aspect of fighting against the proliferation of disinformation is to hold its creators accountable either through criminal or civil liability. When speech is clearly false or misleading or potentially harmful, the law, however imperfect, may step in to penalize those spreading disinformation. However, the lines between misinformation and legitimate opinion are not always clear, so policymakers should use caution in applying criminal or civil liability.

c. Responsibilities of Social Media Companies

Social media corporations can be crucial allies in the effort to combat the dissemination of fraudulent news and disinformation on the internet. The CDA contains the most pertinent legal provisions for this purpose, particularly § 230. The impetus for this legal statute was a court decision which ruled that Prodigy, an internet service provider, could be viewed as a "publisher" of libelous statements that a third party had posted on an online message board maintained and moderated by Prodigy, and therefore, could face civil charges for defamation.[94] Congress passed the Act out of concern that the ruling meant that internet service providers could be prevented from removing objectionable material such as pornography from websites frequented by minors due to the possibility of being subjected to publisher liability.[95] Under § 230, a purveyor of an interactive computer service is immune from accountability for disseminating or releasing an aspersive assertion generated by a third party, provided that the service provider did not participate in making or spreading the contentious material.[96]

The CDA has affected the judiciary, with the Court of Appeals for the Ninth Circuit holding in *Batzel v. Smith* that a website or listserv provider can be immune from publisher liability.[97] Furthermore, state courts have found that internet service providers could not be held responsible

[91] *Id.*
[92] *Id.*
[93] *Id.*
[94] Pinkus, *supra* note 29.
[95] Ryan King, *Online Defamation: Bringing the Communications Decency Act of 1996 in Line With Sound Public Policy*, 2 DUKE LAW TECHNOL. REV. 1, ¶4 (2003).
[96] *Id.* at ¶5.
[97] Id. at ¶7; *Batzel v. Smith*, 333 F.3d 1018 (9th Cir. 2003), overruled on other grounds by *Breazeale v. Victim Servs.*, 878 F.3d 759 (9th Cir. 2017).

for the content posted on their websites, using federal cases to support their rulings.[98] In addition, state courts, when asserting immunity from contract claims, rely on federal cases rather than state cases. Additionally, state courts have frequently recognized that § 230 covers claims of negligence, unfair competition, premises liability, contract claims, and criminal liability.[99] For example, in *Morrison v. America Online, Inc*., the court found that claims to be a third-party beneficiary of a service provider's member agreement with chat-room users did not constitute § 230 immunity.[100] In *People v. Gourlay*, the court ruled that "the prosecution of defendant for distributing or promoting images of child sexually abusive material, as tried and argued before the jury, was based on defendant's active involvement in the creation of [multiple] websites," and thus the defendant could be prosecuted under § 230 of the C.D.A.[101]

Furthermore, federal courts have also held that § 230 immunizes internet service providers against federal causes of action.[102] For example, in *Fair Housing v. Roommates.com*, the court ruled that the website was not a passive publisher of content but instead was an information content provider due to the questionnaires it developed to elicit information from users, and was thus not immune from liability under § 230.[103] Furthermore, courts have held that § 230 immunity does not apply to federal intellectual property claims, with the court in *Gucci Am., Inc. v. Hall & Assocs.*, claiming that immunizing the defendant from trademark claims would limit the reach of intellectual property laws.[104] Courts in the United States have almost unanimously held that the law applies to more than traditional internet service providers, but that diverse services such as online booksellers, online matchmakers, a public library, and chat room creators are all immune from civil liability, with the court in *Carafano* stating that § 230 protection is "quite robust, adopting a relatively expansive definition of 'interactive computer service.'[105]

However, a minority of courts have ruled that distributor liability is not covered by § 230 immunity, with the California Supreme Court ruling *Barrett v. Rosenthal* that the CDA's immunity only applied to publisher liability and not to distributors for defamatory statements.[106] In addition, courts have found that § 230 is an affirmative defense, and thus defendants cannot move for dismissal of such claims.[107] For example, the court in *Novak v. Overture Servs., Inc*. found that "invocation of Section 230(c) immunity constitutes an affirmative defense[, a]s the parties are not required to plead around affirmative defenses, such an affirmative defense is generally not fodder for a Rule 12(b)(6) motion," although the court did grant the defendant's motion to dismiss under § 230.[108] Courts have found a wide variety of interpretations of the

[98] Electronic Frontier Foundation, *Defamation: CDA Cases*, INTERNET LAW TREATISE, https://ilt.eff.org/Defamation__CDA_Cases.html (last visited Feb 28, 2023) [https://perma.cc/9A6F-DJD2].; see *Carafano v. Metrosplash.com, Inc.*, 339 F.3d 1119 (9th Cir. 2003); *Gentry v. eBay, Inc.*, 99 Cal. App. 4th 816 (2002); *Universal Commc'ns Sys., Inc. v. Lycos, Inc.*, 478 F.3d 413 (1st Cir. 2007).
[99] *Id.*
[100] 153 F. Supp. 2d 930 (N.D. Ind. 2001).
[101] 2009 WL 529216 (Mich. App. Ct. Mar. 3, 2009).
[102] King, *supra* note 95 at ¶9.
[103] Electronic Frontier Foundation, *supra* note 98; 489 F.3d 921 (9th Cir. 2007).
[104] *Id.*; 135 F. Supp. 2d 409 (S.D.N.Y. 2001).
[105] Electronic Frontier Foundation, *supra* note 98; *Carafano*, 339 F.3d at 1123.
[106] Electronic Frontier Foundation, *supra* note 98; 146 P.3d 510 (Cal. 2006).
[107] Electronic Frontier Foundation, *supra* note 98.
[108] *Id.*; 309 F.Supp.2d 446, (E.D.N.Y. 2004).

CDA, and it has generally protected internet service providers, including social media networks, from liability under § 230.

Social media companies' immunity under § 230 has allowed social media companies to remove or restrict access to "obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable content."[109] Social media companies accordingly set their policies and Terms of Conditions to state that they may remove such content, with platforms such as Facebook, Twitter, and YouTube banning terrorist groups posting material promoting violence, and have banned groups such as ISIS, Al Qaeda, and Hezbollah from access due to their status as U.S-designated terrorist organizations.[110] Also, platforms such as Facebook, Twitter, Snapchat, YouTube, Reddit, and Twitch suspended President Donald Trump's accounts and some of his supporters under § 230 for posting misinformation about the 2020 presidential election.[111] Thus, the CDA provides the main mechanism for social media companies to remove or restrict disinformation from their platforms without triggering the First Amendment, as these are considered private, rather than public, forums.

d. Current Regulatory Model in the United States

i. The First Amendment and Section 230

The Supreme Court established in *Alvarez* that there is a 'right to lie' under the First Amendment, meaning that false speech can be protected by the First Amendment.[112] However, private social media companies are given broad latitude to police content on their platforms through CDA § 230, and this law protects interactive computer service providers from liability from defamatory statements if the service provider did not develop the statement itself.[113] While social media companies themselves are not liable for the spread of fake news and misinformation, they nevertheless have taken proactive steps to combat their spread. For example, social media platforms may prohibit users from posting content that violates copyright law or solicits illegal activity, and may prohibit users from posting information that it deems to be false under its Terms and Conditions.[114]

ii. Controversies Surrounding Section 230

However, the protections for internet providers from liability for removing content have become controversial, with many arguing that the moderation on social media websites amounts to censorship. The main reform suggested is to amend CDA § 230 to remove the immunity provisions from the act.[115] Indeed, in September 2020, the Justice Department sent draft legislation to Congress to reform § 230 by narrowing the scope of liability protection.[116]

[109] Pinkus, *supra* note 29.
[110] *Id*.
[111] *Id*.
[112] Tompros et al., *supra* note 48, at 93.
[113] Pinkus, *supra* note 29.
[114] Jason A. Gallo & Clare Y. Cho, *Social Media: Misinformation and Content Moderation Issues for Congress*, 32 (2021), https://crsreports.congress.gov/product/pdf/R/R46662.
[115] *Id*.
[116] *Id*. at 19.

However, this draft legislation was not enacted.[117] This is in response to the dominance of major technology firms such as Twitter, Facebook, and Google over social media and their role as gatekeepers to other media, as policymakers were concerned that these companies would use their content moderation policies to censor stories that they did not want published.[118] However, there have been objections from critics of such efforts that removing social media companies' § 230 immunity would lead to such companies allowing objectionable or obscene content on their platforms due to their inability to restrict such content without being subject to liability.[119]

iii. Proposed Amendments to the CDA

It has also been proposed that the CDA be amended to include a broad definition of "development" and a "take-down and put-back" provision. This would hold an interactive computer service provider responsible for publishing content if they deliberately selected the information posted on their service, and for distributing content if they refused to delete a statement they knew was defamatory or had received a formal complaint from an affected party.[120] Under the broad definition, internet service providers would be held accountable for publisher liability in the event of their active selection of posted information. Additionally, the take-down and put-back provisions would render the service provider liable for distributor liability if they decline to eliminate defamatory content from their service.[121]

iv. Challenges and the Need for Clarity

It is clear that the current regulatory framework for dealing with misinformation on social media is very imperfect, although it has been effective in policing truly objectionable content. The line between what is defamatory and false and what is a difference of opinion needs to be more clearly enunciated.

IV. European and Asian Laws on Disinformation

a. European Union Laws for Controlling Disinformation

Europe's approach to disinformation differs greatly from that of the United States because its experiences have been shaped by the history of totalitarian dictatorships that arose in the 1920s and 1930s, which featured antisemitic rhetoric and actions, and the persecution and genocide of minority groups.[122] For this reason, European countries do not have any equivalent to the First Amendment, with laws in many nations restricting the freedom of speech to prevent the

[117] *Id*.
[118] *Id*. at 19–20; Shannon Bond, *Trump Warns Social Media As Justice Department Aims To Weaken Tech's Legal Shield*, NPR, Sep. 23, 2020, https://www.npr.org/2020/09/23/916096008/justice-department-proposes-weakening-social-medias-legal-shield.
[119] Gallo and Cho, *supra* note 114.
[120] King, *supra* note 95, at 5–10.
[121] King, *supra* note 95.
[122] Tani, *supra* note 40, at 241.

recurrence of racism.[123] For example, Germany, Austria, and France have laws against Holocaust denial.[124]

The main law for the European Union on freedom of speech is Article 10 of the European Convention on Human Rights (ECHR), which sets out limits for the freedom of expression as follows:

> The exercise of these freedoms, since it carries with it duties and responsibilities, may be subject to such formalities, conditions, restrictions or penalties as are prescribed by law and are necessary in a democratic society, in the interests of national security, territorial integrity or public safety, for the prevention of disorder or crime, for the protection of health or morals, for the protection of the reputation or rights of others, for preventing the disclosure of information received in confidence, or for maintaining the authority and impartiality of the judiciary.[125]

Under the legal framework established by the European Union, the expression of the fundamental freedom of speech is subject to intricate evaluation, and restrictions are permissible in instances where safeguarding an individual's entitlement to privacy, standing, or the preservation of national security is imperative.[126]

The ECHR has interpreted Article 10 of the European Convention on Human Rights in various cases, including NIT S.R.L. v. the Republic of Moldova (2022) and, Lehideux and Isorni v. France (1998). In NIT S.R.L. v. the Republic of Moldova, the ECHR found that withdrawing a television station's license in Moldova for failing to provide balanced political coverage and biased support of a political party did not violate the right to freedom of expression and information, as the Moldovan authorities had struck a fair balance between protecting media pluralism and the right to freedom of expression.[127] In Lehideux and Isorni v. France, the court held that ECHR Article 10 may still protect the expression of ideas and information which shock, offend, or disturb, but the justification of a pro-Nazi policy cannot enjoy the protection afforded by the Convention as it is an abuse of ECHR rights.[128] The European Union's approach to disinformation is that there is a legitimate state interest in limiting it, a sharp contrast to the First Amendment's protections of some false speech in *Alvarez*.

Moreover, the European Commission established the EU Code of Practice on Disinformation, an unprecedented example of self-regulatory legislation aimed at encouraging private companies to

[123] *Id.*
[124] *Id.*
[125] Council of Eur., *Eur. Conv. for the Protection of H.R. and Fundamental Freedoms*, Art. 10.
[126] Tani, *supra* note 41, at 242.
[127] NIT S.R.L. v. the Republic of Moldova, GLOBAL FREEDOM OF EXPRESSION, https://globalfreedomofexpression.columbia.edu/cases/nit-s-r-l-v-the-republic-of-moldova/ (last visited Mar 1, 2023).
[128] Lehideux and Isorni v. France, GLOBAL FREEDOM OF EXPRESSION, https://globalfreedomofexpression.columbia.edu/cases/lehideux-and-isorni-v-france/ (last visited Mar 1, 2023).

collaborate on finding solutions to the problem of disinformation.[129] As a testament to its effectiveness, 21 companies committed to adhering to the Code, which produced a range of concrete measures that addressed the issue of disinformation. Notably, Facebook, Google, and Twitter developed ad libraries to organize political advertisers, Twitter published its takedown data, while several companies augmented their fact-checking capabilities, implemented misinformation labels, and introduced educational programs to promote journalism and media literacy.[130] Despite the Code's efforts, detractors argued that the lack of a mechanism to report qualitative outcomes made it challenging to gauge its impact.[131] To address this issue, a diverse group of 34 stakeholders, including private companies, trade and industry associations, and international organizations, reviewed and revised the EC Code.[132]

On June 16, 2022, the Commission released the amended and strengthened Code, which makes 44 commitments and includes 128 specific measures.[133] In the updated 2022 Code, there are certain obligations for companies to disclose whether they have accepted particular commitments or to provide a clear rationale for their choice not to do so.[134] Moreover, the Code contained provisions aimed at imposing more stringent criteria for qualification, establishing mechanisms for scrutinizing content monetization, and implementing ad revenue-sharing initiatives. Furthermore, the Code provided for impartial auditors to have equitable access to the companies' data and services.[135] The updated Code now incorporates regulations that mandate more transparency in political advertising. Additionally, it establishes improved mechanisms for cooperating in removing automated bots, systems, or other entities engaged in manipulative behavior.[136]

The European Union has a more rigorous enforcement system than the United States for the dissemination of disinformation. This is because the European Union lacks a legal provision similar to the First Amendment of the United States, and its courts have ruled that preventing misinformation is more important than the unrestricted freedom of speech. Such strict regulation may not be accepted in the United States, as it does not have Europe's history of harmful speech leading to totalitarian regimes. Therefore, it is important to keep in mind these essential differences in the regulatory approach when searching for ways to address misinformation in the United States.

b. Laws of Asian-Pacific Countries for Controlling Disinformation

i. The Consequences of Disinformation in Asian-Pacific Countries

Many Asian countries are in urgent need of government action to combat the spread of disinformation. The circulation of false information has resulted in severe consequences for

[129] Brooke Tanner, *EU Code of Practice on Disinformation*, BROOKINGS (2022), https://www.brookings.edu/blog/techtank/2022/08/05/eu-code-of-practice-on-disinformation/ [https://perma.cc/H6AM-TQEJ].
[130] *Id.*
[131] *Id.*
[132] *Id.*
[133] *Id.*
[134] *Id.*
[135] *Id.*
[136] *Id.*

numerous nations, such as the incitement of violence against ethnic and religious minorities in Myanmar,[137] India, Sri Lanka,[138] and Bangladesh.[139] Social media platforms are often used to disseminate false claims and propaganda, which exacerbates the problem.[140] In addition, misinformation about the risks of burying COVID-19 victims created prejudice against the Muslim community in Sri Lanka, resulting in mandatory cremations.[141]

Additionally, it has been alleged that China is conducting a widespread campaign of spreading false information globally, as part of a larger plan to maintain power, disrupt societal order, demonize opponents, and undermine democratic governance. The campaign's scale and complexity is significant enough for social media giants like Twitter and Facebook to remove thousands of accounts linked to the Chinese government.[142] According to a survey conducted by the International Federation of Journalists in 2021, China's growing influence has also significantly impacted the media landscape worldwide, affecting journalism unions and individuals.[143] Also, there is a rising concern regarding the possible foreign involvement in Taiwan's elections, specifically from mainland China. The dissemination of false or misleading information has been recognized as a potential method that may be used to sway election results. The matter carries significant importance in Taiwan, and measures are being taken to devise strategies to thwart any interference in the democratic process.[144]

ii. Efforts to Combat Disinformation on Social Media Platforms

To combat the widespread circulation of false information, Asian countries have taken proactive steps. This is crucial because these nations have some of the world's largest and fastest-growing online communities, with over 1.2 billion mobile internet users (excluding China). Additionally, Asia is home to seven of the largest Facebook markets, seven of the largest Twitter markets, and

[137] Paul Mozur, *A Genocide Incited on Facebook, With Posts From Myanmar's Military*, THE NEW YORK TIMES, Oct. 15, 2018, https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html; A Year Later, Myanmar's Fact-checkers Try to Catch Up with COVID-19 Infodemic, REPORTING ASEAN - VOICES AND VIEWS FROM WITHIN SOUTHEAST ASIA (2020), https://www.reportingasean.net/a-year-later-fact-checkers-try-to-catch-up-with-covid-19-infodemic/.

[138] AFP, *Misleading claim circulates that Muslims ignored COVID-19 curfew at Sri Lankan mosque*, AFP FACT CHECK (2020), https://factcheck.afp.com/misleading-claim-circulates-muslims-ignored-covid-19-curfew-sri-lankan-mosque.

[139] Qadaruddin Shishir, *Bangladeshi media's troubling tango with fake news*, THE DAILY STAR (2022), https://www.thedailystar.net/views/opinion/news/bangladeshi-medias-troubling-tango-fake-news-2985186.

[140] GSMA, *Exploring online misinformation and disinformation in Asia Pacific*, 22 (2021), https://www.gsma.com/asia-pacific/wp-content/uploads/2021/07/190721-Exploring-misinformation-in-Asia-Pacific-1.pdf.

[141] Matt Field, *Old hatreds fuel online misinformation about COVID-19 in South Asia*, BULLETIN OF THE ATOMIC SCIENTISTS (2020), https://thebulletin.org/2020/11/old-hatreds-fuel-online-misinformation-about-covid-19-in-south-asia/.

[142] Kari Paul, *Twitter and Facebook crack down on accounts linked to Chinese campaign against Hong Kong*, THE GUARDIAN, Aug. 20, 2019, https://www.theguardian.com/technology/2019/aug/19/twitter-china-hong-kong-accounts.

[143] Raksha Kumar, *How China uses the news media as a weapon in its propaganda war against the West*, REUTERS INSTITUTE FOR THE STUDY OF JOURNALISM (2021), https://reutersinstitute.politics.ox.ac.uk/news/how-china-uses-news-media-weapon-its-propaganda-war-against-west.

[144] Jude Blanchette et al., *Protecting Democracy in an Age of Disinformation: Lessons from Taiwan*, 35, 5–12 (2021).

six of the largest Instagram markets, which highlights the region's significant impact on the digital world. Therefore, these nations consider it essential to address the dissemination of misinformation on social media platforms.[145] Pakistan and Thailand have both taken steps to address the spread of false information related to COVID-19. Pakistan has launched a website to prevent the dissemination of misleading information, while Thailand has established an Anti-Fake News Centre supervised by the Ministry of Digital Economy and Society. The Centre has identified around 500 instances of false news related to COVID-19 between January and June 2020.[146] It's important to highlight that private businesses have stepped forward to address this challenge. Telenor and Ooredoo have introduced an extensive digital school program and digital literacy campaign in Myanmar. Their goal is to promote digital literacy among the general population.[147] In Indonesia, the government and media partners have collaborated with nine social media and messaging apps to detect and eliminate user-generated content that contains false or misleading information.[148]

iii. Legal Measures Against Disinformation in Asian-Pacific Countries

Numerous Asian governments are taking decisive measures against the spread of false information on social media. In Indonesia, for instance, the responsibility of combating misinformation falls mainly on the communication ministry, but other organizations such as the National Cyber and Encryption Agency (BSSN) can be called upon for support.[149] In addition, the agency has established a special unit to monitor the spread of disinformation on social media.[150] In Bangladesh, the government has launched a campaign called "Asol Chini" to counter social media disinformation.[151]

Several Asian countries are taking legal measures to combat the spread of disinformation on social media in addition to government actions. For example, Indonesia is using the Information and Electronic Transactions Law (ITE Law) for this purpose,[152] while Pakistan has introduced the Removal and Blocking of Unlawful Online Content Rules 2020 to enforce the Prevention of

[145] GSMA, *supra* note 140, at 8.
[146] *Id.* at 9.
[147] *Id.* at 10; Joseph Waring, *Ooredoo pushes digital literacy with Facebook, GSMA*, MOBILE WORLD LIVE (2019), https://www.mobileworldlive.com/asia/asia-news/ooredoo-pushes-digital-literacy-with-facebook-gsma/ (last visited Feb 14, 2023); Telenor Myanmar's Digital School programme tops 100,000 participants, https://www.telecompaper.com/news/telenor-myanmars-digital-school-programme-tops-100-000-participants--1326088 (last visited Feb 14, 2023); Ellen Alarilla, *ICT in education – perspectives from Myanmar*, https://www.ericsson.com/en/blog/2018/6/ict-in-education--perspectives-from-myanmar (last visited Feb 14, 2023).
[148] Safrin La Batu, *Nine social media, messaging apps agree to combat fake news*, THE JAKARTA POST (2018), https://www.thejakartapost.com/news/2018/01/31/nine-social-media-messaging-apps-agree-to-combat-fake-news.html; GSMA, *supra* note 140, at 11–12.
[149] Freedom House, *Indonesia: Freedom on the Net 2021 Country Report*, FREEDOM HOUSE, https://freedomhouse.org/country/indonesia/freedom-net/2021 (last visited Mar 1, 2023); GSMA, *supra* note 140, at 13.
[150] Freedom House, *supra* note 147; GSMA, *supra* note 138 at 13.
[151] Bangladesh launches campaign against fake information and rumours, DD NEWS, Sep. 9, 2020, https://ddnews.gov.in/international/bangladesh-launches-campaign-against-fake-information-and-rumours.
[152] GSMA, *supra* note 138 at 14; Petra Mahy, Monika Winarnita & Nicholas Herriman, *Influencing the influencers: Regulating the morality of online conduct in Indonesia*, 14 POLICY INTERNET 574 (2022).

Electronic Crimes Act of 2016.[153] Bangladesh has also enacted the Digital Security Act (DSA) to address conflicts arising from the dissemination of false information on Facebook. In response to false information regarding COVID-19, law enforcement agencies in Bangladesh requested the Bangladesh Telecommunication Regulatory Commission (BTRC) to block access to 50 websites and 82 Facebook pages in April 2020.[154] In February 2021, India updated its rules for social media, video streaming, and digital news sites. The changes call for social media firms to set up a complaint resolution mechanism and name executives to collaborate with law enforcement.[155] The guidelines also mandate that social media firms with more than 5 million users should enable the tracking of end-to-end encrypted messages.[156] Furthermore, the ASEAN countries have been working together for over five years to tackle disinformation.[157] And the Taiwanese government and an impartial group called the Taiwan FactCheck Centre help debunk false information.[158] The government clarifies false information quickly by making meme images related to the department where the false information originated and spreading them through social media. On the legal side, Taiwan has two primary legal frameworks for combating disinformation. The first is a set of 13 laws that address disinformation in elections and online platforms' obligations to remove false information. Seven of these laws have been passed, focusing on punishing individuals who spread rumors and false information.[159] Meanwhile, the Taiwanese government currently relies on an existing law called the Social Order Maintenance Act rather than the newly revised laws for prosecution. The second legal framework is the Anti-Infiltration Act, which was passed in late 2019. This law aims to prevent foreign forces from interfering in domestic elections and specifically targets mainland China.[160]

iv. Human Rights and Civil Liberties in Disinformation Fight

However, in pursuing their goal to eliminate disinformation, many Asian governments have violated the human rights and civil liberties of their people. For instance, Cambodia has prohibited reporting in COVID-19 "red zones," while Malaysian authorities expelled a Bangladeshi national for criticizing the government's management of migrant laborers. Additionally, Malaysia has criminalized the dissemination of inaccurate COVID-19 information, imposing prison sentences of up to three years.[161] Likewise, in Taiwan, there is a law called the Social Order Maintenance Act that may be used to penalize individuals who spread false

[153] GSMA, *supra* note 138 at 14; Pakistan: Repeal Amendment to Draconian Cyber Law, HUMAN RIGHTS WATCH (2022), https://www.hrw.org/news/2022/02/28/pakistan-repeal-amendment-draconian-cyber-law (last visited Mar 1, 2023).
[154] GSMA, *supra* note 138 at 14–15; Saad Hammadi, *Bangladesh's 'Fake News' Law Is Used to Stifle Dissent*, (2021), https://thediplomat.com/2021/08/bangladeshs-fake-news-law-is-used-to-stifle-dissent/.
[155] GSMA, *supra* note 138 at 15.
[156] *Id.*; Namrata Maheshwari & Greg Nojeim, *Part 2: New Intermediary Rules in India Imperil Free Expression, Privacy and Security*, CENTER FOR DEMOCRACY AND TECHNOLOGY (2021), https://cdt.org/insights/part-2-new-intermediary-rules-in-india-imperil-free-expression-privacy-and-security/.
[157] GSMA, *supra* note 140, at 16.
[158] Chih-Chieh Yang, *Fighting Against Disinformation from Foreign Forces? Or Suppressing Criticism from Domestic Opposition Parties?: Taiwan's Experience*, 24 ASIA-PAC. J. HUM. RIGHTS LAW 43, 46 (2023).
[159] *Id.* at 44.
[160] *Id.* at 45.
[161] Karen Lee & Andreyka Natalegawa, *Fake News Crackdowns Do Damage Across Southeast Asia During Pandemic*, CSIS (2021), https://www.csis.org/blogs/new-perspectives-asia/fake-news-crackdowns-do-damage-across-southeast-asia-during-pandemic [https://perma.cc/Z949-JZ67].

information through the internet. It has been reported that this law is commonly used to punish those who criticize the government, especially the ruling party. Most cases involve individuals who criticize the government or spread false information about issues such as violence, epidemics, candidates, and unfair election administration. The responsibility for enforcing this law rests with the police, who are under government control, and may be more likely to investigate cases that involve criticism of the ruling party rather than those directed at opposition parties. Although the courts may not always impose penalties on the individuals under investigation, the fact that they are summoned to the police station or court may deter people from criticizing the government.[162] In addition, it is possible that the ruling party may use the Anti-Infiltration Act, a law in Taiwan, to classify certain forms of political engagement as criminal and suppress opposing views.[163]

Also, Singapore combats COVID-19 misinformation with laws like POFMA, allowing ministers to remove inaccurate online information. The health minister issued a correction order to a Facebook page in Feb 2020, while the government directed Facebook and Twitter to display cautionary messages in May 2021 after India blamed Singapore for a COVID-19 strain. However, these measures may erode trust, impede compliance, and heighten vaccine hesitancy.[164] In March 2021, a prominent politician in Thailand was detained after publicly expressing dissent about the government's decision to rely heavily on a company owned by the king for the production and distribution of vaccines to Thailand and neighboring countries. This decision was made even though the company had no prior experience in vaccine creation.[165] Furthermore, in Thailand, the Minister of Digital Economy and Society made an unusual decision by directing internet service providers to shut down eight Facebook accounts run by pro-democracy activists and political commentators. The reason given was that these accounts were spreading false information on social media.[166] These examples show that several Asian governments are using fake news and misinformation to hide negative information and to punish individuals who express differing viewpoints.

These particular occurrences serve to highlight the significance of the First Amendment in the United States, since the absence of such an amendment or the guidelines implemented by the European Convention of Human Rights would result in governments utilizing the guise of countering fake news to engage in political repression of dissenting individuals. While combating fabricated news and disinformation is a commendable objective, it must not be employed as a tool to quash dissent and alternative perspectives, a course of action that has been observed in Thailand, Malaysia, and Singapore. Without the guarantees of freedom of expression, moderation of social media content can rapidly transform into censorship.

[162] Yang, *supra* note 158, at 58–62.
[163] *Id.* at 62–73.
[164] Lee and Natalegawa, *supra* note 161.
[165] Reuters, *Thai politician indicted for royal insult over vaccine speech*, REUTERS, Apr. 11, 2022, https://www.reuters.com/world/asia-pacific/thai-politician-indicted-royal-insult-over-vaccine-speech-2022-04-11/.
[166] Thailand: Proposed initiatives to combat 'fake news' undermine freedom of expression, ARTICLE 19 (2021), https://www.article19.org/resources/thailand-fake-news-undermine-freedom-of-expression/ (last visited Mar 1, 2023); Komsan Tortermvasana, *Court orders ISPs to shut down 8 internet political commentators*, BANGKOK POST, Jun. 2, 2021, https://www.bangkokpost.com/thailand/general/2125675/court-orders-isps-to-shut-down-8-internet-political-commentators.

V. Conclusion

a. Need to Revise the Past Control Model

The current control model needs to be revised, as laws such as the California Conversation Robot Disclosure Act are overly broad and vague, and contain inadequate enforcement mechanisms. In addition, the dominance of companies such as Twitter and Facebook in the social media sphere poses a problem, as those companies can suppress speech and affect public discourse. This is concerning, as their dominance may squelch honest debate on important or controversial issues. The line between legitimate debate and false information needs to be clearly enunciated to avoid cases of censorship by social media companies.

There is a clear need to revise the past control model, because there is a danger that legitimate public discourse on controversial topics, such as the pandemic and presidential elections, will be restricted under the pretext of combatting "fake news." Such things have already happened in several Asian countries, so it is far from out of the imagination that social media companies or unscrupulous government officials would use fighting misinformation to repress critical speech. Thus, the lines between misinformation and legitimate debate need to be clearly drawn in legislation dealing with fake news, and there needs to be consistent and fair enforcement, so that neither the political left nor the political right is repressed.

b. Correction of Direction

A potential solution to the issue of false information on social media involves a comprehensive review of current legislation and its redrafting to eliminate any ambiguity or excessive breadth. Furthermore, it is imperative to revise CDA § 230, as it presents a significant flaw by providing immunity to social media corporations against censorship allegations. Given that these entities function as public forums for people to express their thoughts and perspectives, this act should not grant them immunity. The regulation proposed by the Trump administration in 2020 serves as a positive initiative, as it disempowers giant social media corporations from acting as gatekeepers. It is also worth noting that these corporations impose their own Terms and Conditions that prohibit the posting of content that is genuinely offensive, such as violent or hateful speech. Consequently, the elimination of § 230 would not impede social media companies' ability to monitor content on their platforms, but would prevent tech giants like Facebook and Twitter from monopolizing and repressing opposing views. Such reforms would create clarity on the issue of fake news by limiting liability to the removal of specific types of content.[167]

It has been predicted that up to 90% of online content could be artificially generated in the next few years. Despite attempts to create a federal task force to examine deep fake technology in the United States, progress has stalled.[168] Representative Yvette D. Clarke, a New York Democrat, proposed the DEEP FAKES Accountability Act in both 2019 and 2021, but the bill has not yet been voted on. Clarke plans to reintroduce the bill in 2023.[169] The proposed DEEP FAKES

[167] Gallo and Cho, *supra* note 114.
[168] Hsu, *supra* note 39.
[169] *Id.*

Accountability Act is one potential solution to the issue of deep fakes. However, some experts suggest that it may not be the most effective way to address the problem. One potential drawback of the act is the creation of a new federal bureaucracy, which would increase the size and influence of the US government. Additionally, it is uncertain whether the legislation would effectively combat the use of deep fakes, as it does not specify the onus of proof necessary to establish that a post is a deep fake rather than a valid parody. The act also fails to specify which party holds responsibility for bearing this burden of proof.[170] Furthermore, revoking § 230 immunity allocated to social media platforms may serve as a catalyst for social media companies to proactively eliminate any inappropriate content. The reason is that these companies function as passive agents, distributing the content rather than creating it. Consequently, any objectionable material is the social media platform's responsibility.[171]

Existing laws also provide a blueprint for defeating fake news, with the tort of defamation, intellectual property laws, criminal laws against "revenge porn," and state tort laws that can apply to "deep fakes."[172] Presently enforced legal statutes hold significant potential in impeding the dissemination of falsified media, including "deep fakes" and erroneous information shared through social media channels. These existing laws furnish a means for legally holding individuals who propagate such misinformation responsible for their actions. As such, it is unnecessary to enact new legislation that may fail in achieving the objective of countering the spread of falsehoods.

To effectively address the issue of misinformation on social media platforms, several key measures ought to be taken. One such measure involves narrowing § 230 immunity, which would restrict social media companies' protection from legal liability to the removal of only certain types of content. By limiting the scope of this immunity, governments can avoid censoring alternative or dissenting viewpoints, while still removing genuinely objectionable material. Another essential aspect of addressing misinformation on social media platforms is the need for increased clarity and precision in laws regarding this matter. By ensuring that there is no ambiguity in defining fake news or misinformation, governments can avoid confusion and misinterpretation. This would further help prevent instances of censorship and preserve the fundamental principle of freedom of speech. Taken together, these measures represent a comprehensive and effective solution to the challenge of combating misinformation on social media platforms. By striking a careful balance between protecting against harmful content and safeguarding the fundamental right to express one's opinions, this approach has the potential to benefit both individuals and society.

[170] Schapiro, *supra* note 58.
[171] *Id.*
[172] *Id.*