Athens Journal of Law - Volume 7, Issue 3, July 2021 – Pages 335-358

Fake News and the Convention on Cybercrime[1]

By Robert Smith[*] & Mark Perry[±]

The COVID-19 pandemic and the recent term of the United States President, Donald Trump, brought the term "fake news" to the attention of the broader community. Some jurisdictions have developed anti-fake news legislation, whilst others have used existing cybercrime legislation. A significant deficiency is the lack of a clear definition of fake news. Just because a person calls something "fake news" does not mean that it is indeed false. Especially during pandemics, the primary aim should be to have misinformation and disinformation removed quickly from the web rather than prosecute offenders. The most widely accepted international anti-cybercrime treaty is the Convention on Cybercrime developed by the Council of Europe, which is silent on fake news, the propagation of which may be a cybercrime. There is an Additional Protocol that deals with hate speech, which the authors consider to be a subset of fake news. Using examples from Southeast Asia, the paper develops a comprehensive definition of what constitutes fake news. It ensures that it covers the various flavours of fake news that have been adopted in various jurisdictions. Hate speech can be considered a subset of fake news and is defined as the publication or distribution of fake news with the intention to incite hatred or violence against ethnic, religious, political, and other groups in society. The paper proposes some offences, including those that should be applied to platform service providers. The recommendations could be easily adapted for inclusion in the Convention on Cybercrime or other regional conventions. Such an approach is desirable as cybercrime, including propagating fake news, is not a respecter of national borders, and has widespread deleterious effects.

Keywords: *Fake news; hate speech; Convention on Cybercrime; draft legislation*

Introduction

Social media is becoming pervasive worldwide, especially as the cost of smartphones makes it more readily available to all levels of society. The widespread use of social media has opened up more opportunities for computer-related crime, where crimes that once required a real presence can be undertaken

[*] PhD (Civil Engineering); MPhil in law, Postgraduate Student, University of New England, Armidale NSW, Australia. Corresponding author. E-mail: r.b.smith@unswalumni.com or robert@aecconsultants.asia.

[±] Professor of Law, School of Law, University of New England, Armidale NSW, Australia.
E-mall: mperry21@une.edu.au.

[1] This research is supported by an Australian Government Research Training Program (RTP) Scholarship.

https://doi.org/10.30958/ajl.7-3-4 doi=10.30958/ajl.7-3-4

in the virtual world;[2] hence the appellation – *cybercrime*. The simplest definition of cybercrime is a computer-related crime that uses a computer network.[3]

Cybercrime is a fast-growing area of crime. More and more criminals are exploiting the speed, convenience and anonymity of the Internet to commit a diverse range of criminal activities that know no borders, either physical or virtual, cause serious harm and pose genuine threats to victims worldwide.[4] The total value at risk from cybercrime worldwide in 2019-2023 is estimated to be as high as USD 5.2 trillion.[5] Worldwide spending on information security was estimated to be USD 114 billion in 2018,[6] with global spending on security exceeding USD 1 trillion cumulatively for 2017 to 2021.[7]

The most widespread misuse of social media is probably participation in the cybercrime of publishing or spreading misinformation or disinformation, commonly called *fake news*. The posting and sharing of fake news can be trivial[8] or extreme. It can lead to racial, sectarian, and political tension resulting in property damage, violence, and even death.[9] During the current (2020-2021) COVID-19 pandemic, the posting of fake news can have dire consequences where the promotion of unproven therapies and conspiracy theories can lead to needless victims and an increasing death toll. The difficulty for many social media users, especially in the developing world, is differentiating between the truth and fake news. There are also claims that fake news has been used to inflame ethnic tensions in Myanmar and Indonesia; and influence elections in countries such as Indonesia[10] and the Philippines[11]and countries in the developed world.[12]

In 2021 the Special Rapporteur for the United Nations Human Rights Council noted:

> Disinformation is a complex, multifaceted phenomenon with serious consequences. It destroys people's trust in democratic institutions. It thrives where public information regimes are weak and independent investigative journalism is constrained. It disempowers individuals, robbing them of their autonomy to search, receive and share information and form opinions. In the platform world, individuals are regarded as users, not as rights holders with agency.[13]

Smith and Perry argued that the aim of fake news legislation should be to stop the promulgation of fake news, have the information removed with a correction and an apology issued.[14] Only serious cases should be prosecuted.[15] Publication of

[2] World Bank and United Nations (2017) at 18.

[3] Gercke (2014) at 11.

[4] Interpol (2019).

[5] Accenture Security (2019).

[6] Cybercrime Ventures (2019).

[7] Ibid.

[8] Smith & Perry (2020).

[9] Temby (2019) at 6.

[10] Ibid.

[11] See e.g. Ong & Cabañes (2019); Tapsell & Curato (2019).

[12] See e.g. Allcott & Gentzkow (2017); Cantarella, Fraccaroli & Volpe (2019).

[13] Khan (2021).

[14] Smith & Perry (2020) at 259.

the charges in the press tends to spread the 'fake news' rather than suppress it.[16] Removing fake news and requiring an apology rather than making it a criminal offence would also help deny the internet trolls their hobby of reporting to police the posting of views that do not agree with their views or those of the monarchy or the government.[17] The action of such trolls is tantamount to persecution. Furthermore, conducting misinformation campaigns by government authorities, including the military, should be explicitly prohibited.[18]

Whilst referring specifically to the Association of Southeast Asian Nations (ASEAN), Smith and Perry considered that there should be a uniform approach to addressing fake news:

a. First and foremost is to develop a common understanding of what constitutes *fake/hoax news*;
b. Develop a 'model' anti-misinformation/fake news law, treaty or declaration that has a comprehensive definition of fake news with the focus being on the instigators and not the propagators. For fake news, civil penalties should be considered with the focus on removal of fake news from the Internet, corrections, and apologies rather than severe prison sentences. It should include a specific offence of spreading misinformation by government personnel, including the military, whether in an official or unofficial capacity. In addition, it should, if possible, grade the severity of possible offences; and
c. Provide guidelines on regulating the fake news industry.

This study will develop a comprehensive definition of fake news and test its utility. International borders do not constrain the posting and spreading of fake news in cyberspace, so it is critical to have a widely accepted definition. The difficulty arises with enforcement where the evidence moves from the physical to the electronic, requiring a different skills base for investigators and prosecutors.[19] The study will then investigate how the offence of fake news might be introduced into a treaty, the Convention on Cybercrime.[20]

It should be noted that misinformation and disinformation have similar meanings with authors preferring one term over the other.

[15] Ibid.
[16] Ibid.
[17] Ibid.
[18] Ibid at 260.
[19] World Bank & United Nations (2017) p 19.
[20] Convention on Cybercrime.

Background

Convention on Cybercrime

The Council of Europe developed the *Convention on Cybercrim*e.[21] As of 16 May 2021, the *Convention* had been ratified by 66 countries, including 46 of the 47 members of the Council of Europe.[22] The Philippines is the only ASEAN member to have signed the *Convention*.[23] The other parties to the treaty are Argentina, Australia, Canada, Cape Verde, Chile, Colombia, Costa Rica, Dominican Republic, Ghana, Israel, Japan, Mauritius, Morocco, Panama. Paraguay, Peru, Senegal, Sri Lanka, Tonga, and the United States.[24] The *Convention* mandates what national-level legislative and other measures are required under the domestic law of the signatories.[25] Whilst the *Convention* is not unique, it is the most recognised and has the most member states as parties.

The *Convention* groups the offences into five categories.

Offences against the confidentiality, integrity and availability of computer data and systems are:

a) Illegal access;[26]
b) Illegal interception;[27]
c) Data interference;[28]
d) System interference;[29] and
e) Misuse of devices.[30]

Computer-related offences are:

a) Computer-related forgery;[31] and
b) Computer-related fraud.[32]

Content related offences are all related to child pornography.[33]

As can be seen from the above, computer-based crime, as defined in the *Convention*, is wide-ranging. Nevertheless, there is one obvious omission, namely any reference to *fake news*. The potential for social mass media was in its infancy, and the focus was on criminal activity moving from the physical to the digital

[21] Reservations and Declarations for Treaty No.185.
[22] Ibid.
[23] Ibid.
[24] Ibid.
[25] Convention on Cybercrime, ch II.
[26] Ibid art 2.
[27] Ibid art 3.
[28] Ibid art 4.
[29] Ibid art 5.
[30] Ibid art 6.
[31] Ibid art 7.
[32] Ibid art 8.
[33] Ibid art 9.

space. *Facebook*, the American social media conglomerate, which initiated the worldwide rise of computer-based mass social media and social networking, was founded in February 2004.[34] This was, over two years after the *Convention* was opened for signature and eight months before the Convention entered into force.[35] In its first-quarter report for 2021, Facebook reported that the number of daily active users (DAUs) was, on average, 1.88 billion for March 2020, and there were 2.85 billion monthly active users (MAUs) as of 31 March 2021.[36] This was an increase year on year of 8% for DAUs and 10% for MAUs.[37] With this almost exponential growth has come Facebook's use as a tool to propagate false news.[38]

On 1 March 2006, an *Additional Protocol*[39] on hate speech entered into force following five ratifications from members of the Convention.[40] As of 21 May 2021, the Protocol had been ratified by 33 countries, including 30 members of the Council of Europe.[41] Twelve parties had signed but not ratified the treaty.[42]

The Protocol defines *racist and xenophobic material* as any written material, any image or any other representation of ideas or theories, which advocates, promotes or incites hatred, discrimination, or violence, against any individual or group of individuals, based on race, colour, descent or national or ethnic origin, as well as religion if used as a pretext for any of these factors.[43]

A party is required to make it a criminal offence under domestic law to distribute or otherwise make available when committed intentionally and without right, or "racist and xenophobic material to the public through a computer system".[44] Provided that other effective remedies rather other than criminal liability are available such liability need not apply where a person "advocates, promotes or incites discrimination that is not associated with hatred or violence".[45] A party may reserve the right not to apply criminal sanctions in cases of discrimination "for which, due to established principles in its national legal system concerning freedom of expression, it cannot provide for effective remedies" under art 3(2).[46]

It also requires a party to make it a criminal offence under domestic law to use a computer system intentionally and without right, to make a racist and xenophobic motivated threat[47] and to use a computer system to make a racist and xenophobic motivated insult.[48] This offence occurs when the accused threatens or insults publicly through a computer system "persons for the reason that they

[34] Facebook (2008).
[35] Convention on Cybercrime.
[36] Facebook (2021).
[37] Ibid.
[38] See e.g. Etter (2017); Mozur (2018); Stecklow (2018).
[39] Additional Protocol to the Convention on Cybercrime.
[40] Chart of signatures and ratifications of Treaty 189.
[41] Ibid.
[42] Ibid.
[43] Additional Protocol to the Convention on Cybercrime.
[44] Ibid art 3.1.
[45] Ibid art 3.2.
[46] Ibid art 3.3.
[47] Ibid art 4(1).
[48] Ibid art 5(1).

belong to a group, distinguished by race, colour, descent or national or ethnic origin, as well as religion, if used as a pretext for any of these factors, or a group of persons which is distinguished by any of these characteristics".[49] In the case of insults, a party may reserve the right not to apply the article[50] or may require that the insult actually exposed the person or group of persons to hatred, contempt or ridicule.[51]

Denial, gross minimisation, approval or justification of genocide or crimes against humanity should also be established as a criminal offence.[52] Again, the party may reserve the right not to apply the article or require that the action be committed with intent.[53] When committed intentionally and without right, the aiding or abetting the commission of any of the offences in the Additional Protocol must also be established as a criminal offence.[54]

The Additional Protocol is different from the Convention in that each of the Articles allows for limited reservations. Nevertheless, despite the parties' ability to register reservations, there has been reluctance by a number of the parties to the Convention to become parties to the Additional Protocol. During negotiations for the Convention, the resistance of the United States resulted in provisions designed to eliminate racist ("hate speech") internet sites not being included in the Convention.[55] The United States is not a party to the Additional Protocol.

The United States Department of Justice wrote to the Council of Europe committee of experts drafting the Additional Protocol. It noted that the United States "deplores racism and xenophobia, and the violence and other harmful conduct that racist and xenophobic groups often seek to foster".[56] The United States does not criminalise or prohibit racist speech per se. In addition, it is severely constrained by the First Amendment to the United States Constitution on state action restraining or punishing speech based on its content.[57] "Nor are the restrictions of the First Amendment necessarily limited to materials originating in the United States or to publications with a purely American audience".[58]

In January 2018, the European Commission appointed a high-level group of experts (HLEG) to advise it on countering the spread of disinformation online.[59] The Committee preferred the word "disinformation" over "fake news" and "includes all forms of false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit".[60] Rather than recommending a legislative response, they recommended responses based on five pillars: enhance the transparency of online news; promote media and

[49] Ibid art 4(1) and art 5(1)

[50] Ibid art 5(2)(b).

[51] Ibid art 5(3)(a).

[52] Ibid art 6(1).

[53] Ibid art 6(2)(a).

[54] Ibid art 7.

[55] Murphy (2002) p 973.

[56] Boyd & Chertoff (2001).

[57] Ibid.

[58] Ibid.

[59] European Commission (2018) ch 3, p 2.

[60] Ibid.

information literacy; develop tools for empowering users and journalists; safeguard the diversity and sustainability of the European news media ecosystem; promote continued research on the impact of disinformation.[61]

Analysis

Developing a Comprehensive Definition of Fake News

Definition

There seems to be no accepted definition of "fake news" or a threshold for defining the offences. Often an item is posted in good faith, believing it to be true. At first glance, the definition of fake news proposed by Gelfert, namely the "*deliberate* presentation of (typically) false or misleading claims as news, where the claims are *misleading by design",*[62] would seem to be applicable. Even in cases of intercommunal tension, as was the case, in, for example, Indonesia[63] and Myanmar[64] it is highly likely that many of the transmitters of hate posts believe them to be true. Often the "news" in fake news does not appear to meet the standard dictionary definition — "information about a recently changed situation or a recent event".[65] Singapore, which is the only member of ASEAN with specific fakes news legislation extant,[66] has included the following two definitions in its *Protection from Online Falsehoods and Manipulation Act 2019*[67]:

a) a statement of fact is a statement which a reasonable person seeing, hearing or otherwise perceiving it would consider to be a representation of fact;[68] and
b) a statement is false if it is false or misleading, whether wholly or in part, and whether on its own or in the context in which it appears.[69]

A definition including "a reasonable person" poses some difficulties as it leaves the meaning wide open to interpretation.

For instance, Tapsell posed the following questions that need to be considered, especially in a country like Indonesia, which is seismically active as well as prone to monsoon-induced natural disasters:[70]

[61] Ibid p 5.
[62] Gelfert (2018) at 108 (emphasis added).
[63] Temby (2019).
[64] See e.g. Douek (2018); Gleicher (2018); Mozur (2018); Stecklow (2018); United Nations Human Rights Council (2018).
[65] Collins English Dictionary (2020).
[66] Al Jazeera (2019).
[67] Protection from Online Falsehoods and Manipulation Act 2019.
[68] Ibid s (1)(a).
[69] Ibid s (2)(b).
[70] Tapsell (2019) ar 7.

a) Was the sharing of information malicious, knowing it was fake, or did they consider it was real?
b) Should those who post-disaster warnings wait until an official warning is posted in situations where time is the essence? and
c) In political hoax cases, is posting it a crime or was it a non-chargeable error of judgement?

In the light of these questions, it appears there should be a minimum threshold in either the definition of fake news or the penalty to be applied. Therefore, a more detailed definition is preferred as it is considered that it would assist in developing the minimum threshold to be applied for the prosecution of fake news. Therefore, as a first step, it is proposed to adopt the definition of Gelfert[71] as the lower threshold, namely: *fake news is the deliberate presentation of (typically) false or misleading claims as news, where the claims are misleading by design.*

This definition is too simplistic, so it is proposed to amplify it by introducing the expanded definition of fake news by introducing the terms of the matrix developed by Wardle.[72]

The 'fake news' matrix consists of seven types of misinformation and disinformation:[73]

a) Satire or parody — no intention to cause harm but with potential to fool;
b) Misleading content — misleading use of information to frame an issue or individual;
c) Imposter content — where genuine sources are impersonated;
d) Fabricated content — content is 100% false and designed to deceive and harm;
e) False connection — headlines, visuals or captions do not support the content;
f) False context — genuine content is shared with false contextual information; and
g) Manipulated content — genuine information or imagery is manipulated to deceive.

The definition was then tested against examples of what was considered to be fake news in three of the ASEAN ten Asian nations, namely Indonesia, the Philippines and Thailand.

Indonesia

The following types of fake news in Indonesia have been identified by Smith (2020):[74]

[71] Gelfert (2018) at 108 (emphasis added).
[72] Wardle (2017).
[73] Ibid.
[74] Smith (2020) ch 3.

a) Saracen, a commercial fake news factory[75] claimed by police to own more than 2,000 online media and around 800,000 social media accounts[76] - fabricated content.
b) Jonru, the blogger, social media star and provocateur,[77] was arrested for 'inciting hatred against ethnic, religious or other groups in society'[78] - fabricated content.
c) Ratna the hoaxer[79] – misleading content.
d) Sharing hoax material about an earthquake and/or tsunami – not deliberate and not misleading by design.[80]
e) Sharing hoax material about children being kidnapped[81] – not deliberate and misleading by design.
f) Sharing fake news that a student protest was taking place in Jakarta[82] – fabricated content.
g) The administrator of an Instagram account spreading material stating President Jakowi was a communist[83] – fabricated content.
h) Uploading a hoax video that said the TNI[84] arrested Chinese citizens for donations to the presidential campaign[85] – fabricated content.
i) The editor of a video that had President Jokowi's running mate, dressed as Santa Claus[86] – fabricated content.
j) Spreading fake news that containers of already punched election ballot papers had arrived in Indonesia before the Presidential election[87] – fabricated content.
k) Sharing claim that President Jakowi had a fake degree[88] – fabricated content.
l) Claim that Election Commission computers had been programmed to automatically determine a President Jakowi win[89] – fabricated content.
m) False claims that Chinese PLA soldiers had infiltrated the Indonesian police[90] – fabricated content.
n) False posts that the Al-Makmur mosque in Jakarta had been attacked[91] – fabricated content.

[75] Chan (2017).
[76] Alfanisa (2017).
[77] Saraswati (2017).
[78] Ibid.
[79] Souisa (2018).
[80] Tapsell (2019) at 2.
[81] Ibid.
[82] Ibid.
[83] Ibid.
[84] Tentara Nasional Indonesia (Indonesian Armed Forces).
[85] Tapsell (2019).
[86] Ibid at 2.
[87] Ibid.
[88] Ibid.
[89] Ibid.
[90] Ibid.
[91] Ibid.

Philippines

The Philippines National Police (PNP) has not disaggregated its data. Smith (2020) noted that the cases investigated by the PNP either online libel-related[92] or COVID-19 Pandemic related.[93] Between January and November 2019, 1,166 cases were investigated by the Philippines National Police.[94] The authors assume that the COVID-19 cases are either not deliberate or misleading by design or fabricated content.

Thailand

a) Animals have been abandoned at Phuket Zoo[95] - not deliberate and misleading by design and possibly alternative offences possibly committed such as soliciting funds or trespassing;
b) Reports of imminent dissolution of House of Representatives[96] - satire or parody;
c) An accusation that the wife of Prime Minister was not Buddhist as she claimed[97] - fabricated content;
d) The government was to impose a tax on feminine hygiene products[98] - not deliberate and misleading by design;
e) A woman claimed that someone had 'returned' her lost wallet[99] - satire or parody;
f) Cups of coffee were selling for Thai baht 12,000[100] – not deliberate and misleading by design.
g) Report of a 'death' at Phuket airport[101] – not deliberate and misleading by design.
h) Report that a person collapsed 'with COVID-19'[102] – misleading content;
i) Suppression of a 'death' in Pattaya hospital'[103] – misleading content;
j) Releasing and sharing *fake news* about COVID-19[104] – misleading content;
k) Spreading *fake news* about COVID-19 spreading in Chiang Mai[105] - fabricated content.
l) False report of 40 COVID-19 cases in neighbourhood[106] - satire or parody;
m) False report of police COVID-19 checkpoint[107] - satire or parody;

[92] Anti Cybercrime Group (2019).
[93] Caliwan (2020).
[94] Anti Cybercrime Group (2019).
[95] Thongtub (2020).
[96] National News Bureau of Thailand (2020).
[97] Thai PBS World (2019).
[98] Ibid.
[99] The Nation (2019a).
[100] The Nation (2019b).
[101] Phuket News (2020).
[102] Boonbandit (2020).
[103] Naew Na (2020).
[104] Nation Thailand (2020a).
[105] Thai PBS World (2020).
[106] Nation Thailand (2020b).
[107] Thongtub (2020b).

n) 44 fake COVID-19 websites[108] - imposter content;
o) 'False' report of no COVID-19 screening at Bangkok Airport[109] – not deliberate and misleading by design; and.
p) Five negative posts on the cash handout scheme[110] - misleading content.

Classification of Fake News

To test the robustness of this definition, it is necessary to ascertain whether all of the offences identified above fill one of the matrix cells based on the available details. The analysis is tabulated summarised in Table 1.

Table 1. *Classification of Fake News Cases Reported Above*

Jurisdiction	Not deliberate and misleading by design[111]	Satire or parody[112]	Misleading content[113]	Imposter content[114]	Fabricated content[115]	False connection[116]	False context[117]	Manipulated content[118]	None of the above
Indonesia	2	0	1	0	12	0	0	0	0
Philippines	X[119]				X[120]				
Thailand	5	4	4	1	2	1	0	0	0

Source: Analysis by the authors.

Whilst this analysis is very preliminary and needs extensive verification, it shows that such a paradigm covers the variety of cases of fake news as discussed above. However, what is required is some refinement of the headings used by

[108] Thai Visa News (2020a).
[109] Human Rights Watch (2020); Thai Visa News (2020).
[110] Nation Thailand (2020c).
[111] The corollary of Gelfert's definition of fake news 'deliberate presentation of (typically) false or misleading claims as news, where the claims are misleading by design' (Gelfert (2018) p 108 (emphasis added)).
[112] 'no intention to cause harm but with potential to fool'- Wardle (2017).
[113] 'misleading use of information to frame an issue or individual' - Wardle (2017).
[114] 'where genuine sources are impersonated' – Wardle.(2017).
[115] 'content is 100% false and designed to deceive and harm' - Waedle (2017).
[116] 'headlines, visuals or captions do not support the content'- Wardle (2017).
[117] 'genuine content is shared with false contextual information' – Wardle (2017).
[118] 'genuine information or imagery is manipulated to deceive' -Wardle (2017).
[119] Caliwan (2020). No details were available on the full range of the 47 offences. It is likely that some were unintentional whilst some may have been deliberate fabrications.
[120] Anti Cybercrime Group (2019). Between January and November 2019, 1,166 cases of online libel were investigated by Philippines National Police.

Wardle, although her more detailed descriptions provide some clarity. Rather than providing a definition of fake news and *truth* which are essentially impossible to define, it is proposed to define the offence incorporating critical components of the Wardle matrix.

Satire or parody is at the very minor end of fake news and certainly does meet the threshold for initiating criminal proceedings. Nor is it likely, in most cases, to meet the threshold for successful actions in civil proceedings for libel or defamation. It has, therefore, been removed from the list of activities that constitute fake news. The definition of fake news becomes:

Fake news is the deliberate publication or distribution of material that contains disinformation or misinformation that was misleading by design and:

a. the material is used to defame an individual; and/or
b. genuine sources are imitated; and/or
c. content is false and meant to deceive or harm; and/or
d. headlines, visuals, or captions do not support the content; and/or
e. genuine content is shared with false contextual information; and/or
f. genuine information or imagery is manipulated to deceive.

This definition does not explicitly cover hate speech which, based on the experience of Indonesia described and to a less extent with that of the Philippines as described above, is a serious issue in both countries. It is related to fake news in that fake news is used in a manner to incite hatred or violence against ethnic, religious, political and other groups in society. Therefore, it is proposed to also develop a definition of hate speech as a particular use of fake news. The definition of hate speech becomes:

Hate speech is the publication or distribution of *fake news* with the intention of inciting hatred or violence against ethnic, religious, political and other groups in society.

These definitions are independent of the mode of publication and distribution. In this case, the proposed convention is restricted to the publication and distribution of fake news in cyberspace. As a result, this requires a definition of a "computer system". In this case, the definition from the *Convention on Cybercrime* would be applicable[121] i.e. a *computer system* means any device or a group of interconnected or related devices, one or more of which, pursuant to a program, performs automatic processing of data;

Recently there has been an increase of persons misrepresenting themselves on social media, using fake names or conduct coordinated campaigns to manipulate public debate.[122] Facebook classifies inauthentic behaviour as people acting so as "to misrepresent themselves on Facebook, use fake accounts, artificially boost the popularity of content or engage in behaviours designed to enable other violations"

[121] Convention on Cybercrime art 1(a).

[122] See e.g. Reuters (2019); Douek (2018); Etter (2017); Gleicher (2019a); Gleicher (2020); Gleicher (2018); Gleicher (2019b); Irving (2018); Mozur (2018); Nyi Nyi Kyaw (2019); Sakim (2020); Stecklow (2018); Tanakasempipat (2021).

of their Community Standards.[123] "This policy is intended to protect the security of user accounts and our services and create a space where people can trust the people and communities they interact with".[124] Behaviour that Facebook prohibits includes:

a. Use of multiple Facebook accounts [to mislead other users];
b. Misleading users about ownership or control of a page;
c. Use inauthentic content about identity, purpose or origin of the entity they represent, source or origin of content or the popularity of the content, to mislead people or Facebook;
d. Use of multiple Facebook accounts working in concert to mislead as defined in dot points as to c above where fake accounts are central to the operation (i.e. coordinated inauthentic behaviour); or
e. Coordinated inauthentic behaviour conducted on behalf of a foreign or government actor.[125]

Whilst this behaviour is undoubtedly anti-social, it would be challenging to legislate it as a criminal offence. A much more appropriate approach is to pass responsibility to the platform service providers.

Fake news requires a distribution channel. In this case, the publishing platform is social media such as Facebook and Twitter and mobile phone messaging applications such as Line, Messenger and WhatsApp. The questions that need to be addressed are: are they neutral players, or are they complicit? What can be enshrined in law to ensure that they remain neutral players?

Role of the Platform Provider

Indonesian GR 80 2019[126] stipulates that a platform provider is not accountable for the content or the consequences due to the existence of illegal electronic information provided it takes immediate action to erase electronic links and/or illegal electronic information immediately becoming aware of its existence.[127] The Asian Internet Coalition (AIC)[128], whose members include Amazon, Facebook, Google and Twitter, suggested that, as their members had various timeframes to undertake removal action, they preferred that the timeframe be altered to "as soon as possible".[129] They recommended the UK. *Electronic Commerce (EC Directive) Regulations 2002* as best practice.[130] The wording of the *Regulations* is much clearer than that of *GR 80*. The hosting article of the *Regulations* reads:

[123] Facebook (2020) s 20.
[124] Ibid.
[125] Ibid.
[126] Government Regulation Number 80 (Indonesia).
[127] Ibid art 22(1)-(2).
[128] Asian Internet Coalition (2020).
[129] Paine (2020).
[130] Ibid.

Where an information society service is provided, which consists of the storage of information provided by a recipient of the service, the service provider (if he otherwise would) shall not be liable for damages or for any other pecuniary remedy or for any criminal sanction as a result of that storage where:

a) the service provider:
   i. does not have actual knowledge of unlawful activity or information and, where a claim for damages is made, is not aware of facts or circumstances from which it would have been apparent to the service provider that the activity or information was unlawful; or
   ii. upon obtaining such knowledge or awareness, acts expeditiously to remove or to disable access to the information, and

b) the recipient of the service was not acting under the authority or the control of the service provider.[131]

As was seen when discussing the comparison between the Indonesian *Amendment to Electronic Information and Transactions Law*[132] and the Facebook Community Standards[133] showed that Indonesian requirements were stricter than those in the Facebook Community Standards. Whilst this was no doubt mainly because Indonesia is a conservative society. Nevertheless, it does show a need to regulate further the activities of platform service providers, especially concerning the removal of harmful content.

Reservations Clause

Clough considered that to be effective, harmonisation must seek to accommodate and reconcile differences between the parties.[134] He suggested that the inclusion of a reservations clause could facilitate this. In the case of fake news, this is even more critical, bearing in mind the right to free speech as, for instance, prescribed in First Amendment to the United States Constitution.[135] As discussed earlier, the United States is severely constrained on undertaking state action restraining or punishing speech based on its content.[136] This constraint is not limited to materials originating within the country or to publications restricted to a purely American audience.[137]

Offences

It is considered that the penalties should be harsher when the fake news is published and/or promulgated by a state actor or consists of hate speech.

[131] Electronic Commerce (EC Directive) Regulations 2002.

[132] Law Concerning Electronic Information and Transactions 2008.

[133] Facebook (2020).

[134] Clough (2014).

[135] Boyd & Chertoff (2001).

[136] Ibid.

[137] Ibid.

Draft Clauses

Definitions

1. *Fake news* is the deliberate publication or distribution of material that contains disinformation or misinformation that was misleading by design and:

   (a) the material is used to defame an individual; and/or
   (b) genuine sources are imitated; and/or
   (c) content is false and meant to deceive or harm; and/or
   (d) headlines, visuals or captions do not support the content; and/or
   (e) genuine content is shared with false contextual information; and/or
   (f) genuine information or imagery is manipulated to deceive.

2. *Hate speech* is the publication or distribution of *fake news* with the intention of inciting hatred or violence against ethnic, religious, political and other groups in society.
3. A *computer system* means any device or a group of interconnected or related devices, one or more of which, pursuant to a program, performs automatic processing of data.

Offences Related to using a Computer System to Publish or Distribute Fake News

1. Each party shall adopt such legislative measures as may be necessary to establish as a criminal offence under its domestic law, when committed intentionally and without right, using a computer system to publish or distribute fake news.[138]
2. Each party shall adopt such legislative or other measures as may be appropriate so that offenders are liable to higher penalties than usual if any of the following aggravating circumstances are present:[139]
   (a) Where the offence includes the publication or distribution of hate speech; and/or
   (b) Where the offence is committed by a public official in the performance of his or her duties.[140]

Corporate Liability[141]

1. Each Party shall adopt such legislative and other measures as may be necessary to ensure that legal persons can be held liable for a criminal offence established under this Convention, committed for their benefit by

[138] The wording up to and including "without right" is taken verbatim from Convention on Cybercrime art 9.
[139] The wording is taken verbatim from ASEAN Convention Against Trafficking art 5(3).
[140] The wording is taken verbatim from ASEAN Convention Against Trafficking art 5(3) (g).
[141] The wording is taken verbatim from Convention on Cybercrime art 12.

any natural person, acting either individually or as part of an organ of the legal person, who has a leading position within it, based on:

(a) power of representation of the legal person;
(b) an authority to make decisions on behalf of the legal person;
(c) an authority to exercise control within the legal person.

2. In addition to the cases already provided for in paragraph 1 of this article, each Party shall take the measures necessary to ensure that a legal person can be held liable where the lack of supervision or control by a natural person referred to in paragraph 1 has made possible the commission of a criminal offence established under this Convention for the benefit of that legal person by a natural person acting under its authority.
3. Subject to the legal principles of the Party, the liability of a legal person may be criminal, civil or administrative.
4. Such liability shall be without prejudice to the criminal liability of the natural persons who have committed the offence.

Role of the Platform Service Provider[142]

1. Liability from Financial Remedies or Criminal Sanctions:

Where a service is provided which consists of the storage of information provided by a recipient of a service, the service provider (if it otherwise would) shall not be liable for damages or for any other financial remedy or for any criminal sanction as a result of that storage where:

(a) the service provider:
 i. does not have actual knowledge of unlawful activity or information and, where a claim for damages is made, is not aware of facts or circumstances from which it would have been apparent to the service provider that the activity or information was unlawful; or
 ii. upon obtaining such knowledge or awareness, acts expeditiously to remove or to disable access to the information, and

(b) the recipient of the service was not acting under the authority or the control of the service provider.

2. Inauthentic Behaviour by Platform Users:

Upon obtaining knowledge or awareness of the following activities being undertaken by users of the platform, the Platform Service Provider shall act expeditiously to remove or to disable access to the information:

(a) use of multiple accounts on the platform, with the intention to mislead other users;
(b) misleading users about ownership or control of a page;

[142] The wording is slightly modified from that in Electronic Commerce Regulations 2002 art 19 by deleting the words "information society" and adding a new clause on inauthentic behaviour.

(c) use of inauthentic content about identity, purpose or origin of the entity they represent, source or origin of content or the popularity of the content, to mislead people or the platform service provider;

(d) coordinated inauthentic behaviour by using multiple accounts working in concert to mislead as defined in dot points (a) to (c) above; where fake accounts are central to the operation;

(e) Coordinated inauthentic behaviour conducted on behalf of a foreign or government actor.

Discussion

It is increasingly clear, especially since the commencement of the COVID-19 pandemic in early 2020 that fake news is becoming more pervasive. As it has become more pervasive it has also become more dangerous as misinformation and disinformation has thrived during the pandemic. A global response is required. The *Convention on Cybercrime*[143] as it stands does not cover fake news. Hate speech but no other forms of fake news are the subject of the Additional Protocol.[144] In light of the fact that significantly less of the parties to the *Convention* are also parties to the *Additional Protocol* it may not be possible to negotiate and then obtain a sufficient number of parties for a protocol on fake news to enter into force.

A more appropriate approach could be to prepare it as a Model Law that would be available for jurisdictions to adopt as they feel appropriate. The advantage of this approach is that it would provide comprehensive definition of fake news as defined above, namely: *Fake news is the deliberate publication or distribution of material that contains disinformation or misinformation that was misleading by design*. It has two essential features that it is *deliberate* and is *misleading by design*. Prosecution should only apply when both these elements are present.

The other feature that the Model Law should also address is inauthentic behaviour on social media. This should include provisions that require platform service providers to take action concerning in authentic behaviour. Without solid support from the platform service providers, it will be impossible to adequately monitor and remove inauthentic content which does not come under the definition of fake news. Government monitoring sites can impact the spreading of fake news that they can identify or is reported to them. However, identification of inauthentic behaviour is a joint responsibility with the platform service provider taking the lead role.

There have been several non-legislative responses recommended in response to the spread of fake news. As mentioned, earlier the European Commission appointed a high-level group of experts (HLEG) recommended a non-legislative response to fake news. Such an approach was also developed by the Ministers Responsible for Information of the ASEAN Member States in their framework to

[143] Convention on Cybercrime.

[144] Additional Protocol to the Convention on Cybercrime.

minimise the harmful effects of fake news.[145] The framework included the following elements:[146]

a) Cooperation in the fields of information and media, and support development of socially responsible media in ASEAN
b) Capitalise on the potential of online and social media and ensure that the Internet remains a safe space;
c) Raise awareness on the potential problems posed by fake news;
d) Improve digital literacy;
e) Strengthen national capacity to detect and respond to fake news;
f) Encourage stakeholders to build on industry norms and guidelines against fake news;
g) Share best practices and experiences on responses to the challenge of fake news
h) Encourage all ASEAN partners and relevant stakeholders to cooperate and join hands in the implementation of this Framework.

Such an approach is undoubtedly required, but as fake news is so pervasive, it is unlikely to have a short-term impact. It may have an impact on the less sophisticated users of the Internet, including those who essentially only use social media. On the other hand, it will have a limited impact on the fake news production industry, such as the fake news factories or farms.

Conclusion

A draft protocol on fake news makes a valuable tool in the fight against the worst excesses of fake news. Even if not part of a protocol, the draft could act as a model for consideration for other jurisdictions, especially those that use non-specific cybercrime legislation to prosecute purveyors of fake news.

Probably the most crucial aspect of the paper has been to develop a robust definition of fake news. The definition must show intent to deceive. It is considered that an appropriate definition is: "*deliberate* presentation of (typically) false or misleading claims as news, where the claims are *misleading by design*".[147] Such a definition would preclude trivial offences and allow law enforcement agencies to focus on the extreme end of the spectrum.

It is also critical to enlist the resources of the platform service providers to control inauthentic behaviour. The approach of Facebook is a good start, but it needs to allocate adequate resources to monitor users.

[145] Framework and Declaration to Minimise Harmful Effects of Fake News.
[146] Ibid.
[147] Gelfert (2018).

References

Accenture Security (2019). 'Ninth Annual Cost of Cybercrime Study', (Published Report, Accenture Security, 6 March 2019). https://www.accenture.com/us-en/insights/security/cost-cybercrime-study.

Al Jazeera (2019). 'Malaysia Parliament Scraps Law Criminalising Fake News', (online, 10 October). https://www.aljazeera.com/news/2019/10/malaysia-parliament-scraps-law-criminalising-fake news-191010024414267.html

Alfanisa, E.W. (2017). 'Saracen and the Dilemma Between Hate Speech and Freedom of Speech in Indonesia', Center for Digital Society (Blog Post, 2 October 2017). https://cfds.fisipol.ugm.ac.id/article/177.

Allcott, H. & M. Gentzkow (2017). 'Social Media and Fake News in the 2016 Election' in *Journal of Economic Perspectives* 31(2):211-236.

Anti Cybercrime Group (2019). 'PNP ACG Operational Accomplishment Report (January to November)' (Manila, Philippines National Police Anti-Cybercrime Group. https://drive.google.com/file/d/1qdk0oPvI48heju6CVvWnoa9HFtkEP4xk/view

Asian Internet Coalition (2020). About (web page) https://aicasia.org/about/.

Boonbandit, T. (2020). 'Two Arrested for Spreading Coronavirus Fake News', *Khaosod English* (online, 30 January 2020) https://www.khaosodenglish.com/politics/2020/01/30/two-arrested-for-spreading-coronavirus-fake news

Boyd, R.F. & M. Chertoff (2001). Letter from U.S. Dep't of Justice Assistant Attorney Generals Ralph F. Boyd, Jr., and Michael Chertoff to Chairman of the Council of Europe PC-RX Committee, 13 December 2001 *cited* in Murphy S.D. (2002).

Caliwan, C.L. (2019). 'PNP Nabs 47 Covid-19 Fake News Peddlers', *Philippines News Agency* (online, 15 April 2020). https://www.pna.gov.ph/articles/1099910.

Cantarella, M., Fraccaroli, N. & R. Volpe (2019) 'Does Fake News Affect Voting Behaviour?' (DEMB Working Paper Series No 146, UNIMORE Dipartimento di Economia Marco Biagi, 6 June 2019) https://ideas.repec.org/p/mod/depeco/0146.html.

Chan, F. (2017). 'Indonesian Police Uncover 'Fake News Factory'', *The Straits Times* (online, 17 September 2017). https://www.straitstimes.com/asia/se-asia/indonesian-police-uncover-fake news-factory

Chart of signatures and ratifications of Treaty 189 - Additional Protocol to the Convention on Cybercrime, concerning the criminalisation of acts of a racist and xenophobic nature committed through computer systems (Chart of Signatures, 21 May 2021). https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/189/signatures?p_auth=2flYzJmy.

Clough, J. (2014). ‘A World of Difference: the Budapest Convention on Cybercrime and the Challenges of Harmonisation’ in *Monash University Law Review* 40:698-736.

Collins English Dictionary (online on 21 June 2020) (online ed, 'News' (def 1)).

Cybersecurity Ventures (2019). '2019 Official Annual Cybercrime Report', (Published Report,) https://www.herjavecgroup.com/wp-content/uploads/2018/12/CV-HG-2019-Official-Annual-Cybercrime-Report.pdf.

Douek, E. (2018). 'Facebook’s Role in the Genocide in Myanmar: New Reporting Complicates the Narrative', *LawFare* (Blog Post, 22 October 2018). https://www.lawfareblog.com/facebooks-role-genocide-myanmar-new-reporting-complicates-narrative

Etter, L. (2017). 'In the Philippines Facebook is a Weapon' in (11 December 2017) *Bloomberg Businessweek*: 54-59.

European Commission (2018) A Multi-dimensional Approach to Disinformation: Report of the Independent High Level Group on Fake News and Online Disinformation.

Facebook (2008). 'Company Timeline', Press Room (28 February 2008) https://web.archive.org/web/20080228004941/http://www.facebook.com/press/info.php?timeline

Facebook (2019). 'Facebook Removes Hundreds of Indonesian Accounts Linked to Fake News and Hate Speech', *Reuters* (online, 1 February 2019) https://www.abc.net.au/news/2019-02-01/facebook-removes-indonesian-accounts-fake-news-hate-speech/10772968

Facebook (2020). 'Community Standards', Facebook (Policy Statement, 2020) https://www.facebook.com/communitystandards/.

Facebook (2021). 'Facebook Reports First Quarter 2021 Results', (Investor News, Facebook Inc, 28 April 2021) https://investor.fb.com/investor-news/press-release-details/2021/Facebook-Reports-First-Quarter-2021-Results/default.aspx

Gelfert, A. (2018). 'Fake News: A Definition' in *Informal Logic* 38(1):84-117.

Gercke, M. (2014). 'Understanding Cybercrime: Phenomena, Challenges and Legal Response' (Published Report, International Telecommunication Union, November 2014). https://www.itu.int/en/ITU-D/Cybersecurity/Documents/Cybercrime2014_E.pdf

Gleicher, N. (2018). 'Removing Myanmar Military Officials From Facebook' (Media Release, Facebook, 28 August 2018 updated 18 December 2018) https://about.fb.com/news/2018/08/removing-myanmar-officials/.

Gleicher, N. (2019a). 'Taking Down Coordinated Inauthentic Behavior in Indonesia [Update]' (Press Release, Facebook, 11 April 2019) https://about.fb.com/news/2019/01/taking-down-coordinated-inauthentic-behavior-in-indonesia

Gleicher, N. (2019b). 'Removing Coordinated Inauthentic Behavior in UAE, Nigeria, Indonesia and Egypt' (Press Release, Facebook, 3 October 2019). https://about.fb.com/news/2019/10/removing-coordinated-inauthentic-behavior-in-uae-nigeria-indonesia-and-egypt

Gleicher, N. (2020) 'Removing Coordinated Inauthentic Behavior from Russia, Iran, Vietnam and Myanmar' (Media Release, Facebook, 12 February 2020). https://about.fb.com/news/2020/02/removing-coordinated-inauthentic-behavior.

Human Rights Watch (2020). 'Thailand: COVID-19 Clampdown on Free Speech', (25 March 2020). https://www.hrw.org/news/2020/03/25/thailand-covid-19-clampdown-free-speech.

Interpol (2019). 'Crime Areas: Cybercrime', International Criminal Police Organization (Web Page, 24 February 2019). https://www.interpol.int/Crime-areas/Cybercrime/Cybercrime.

Irving, E. (2018). ‘The Role of Social Media is Significant: Facebook and the Fact Finding Mission on Myanmar', *OpinioJuris* (Blog Post, 7 September 2018). http://opinionjuris.org/2018/09/07/the-role-of-social-media-is-significant-facebook-and-the-fact-finding-mission-on-myanmar

Khan, I. (2021). Special Rapporteur, Disinformation and Freedom of Opinion and Expression: Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, 47th session, Agenda Item 3, UN Doc A/HRC/47/25 (13 April 2021).

Mozur, P. (2018). 'A Genocide Incited on Facebook, with Posts from Myanmar’s Military', *The New York Times* (online, 15 October 2018) https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html

Murphy, S.D. (2002) 'Hate-Speech Protocol to Cybercrime Convention' in *The American Journal of International Law* 96(4):973-975.

Naew Na (2020). 'Fake News! Thai Woman Arrested for Saying Someone Died in Pattaya from Coronavirus', Naew Na (translated by Thaivisa.com) (online, 6 February 2020). https://forum.thaivisa.com/topic/1146940-fake-news-thai-woman-arrested-for-saying-someone-died-in-pattaya-from-coronavirus.

Nation (2019a). 'Woman May Face Computer Act Charges Over 'Found' Wallet Story' in The Nation (online, 23 June 2019). http://www.nationmultimedia.com/detail/national/30371520.

Nation (2019b). 'Future Forward Deputy Meets with Police Over 3-Minute Fake Post'. The Nation (online, 11 March 2019). http://www.nationmultimedia.com/detail/breakingnews/30365579

Nation Thailand (2020a). 'Govt to Crack Down on Dissemination of Fake News About Coronavirus', The Nation Thailand (online, 6 February 2020). https://www.nationthailand.com/news/30381677

Nation Thailand (2020b). 'Minister Pushes for Prosecution of Woman Who Put Out 'False Message' on Covid-19', The Nation Thailand (online, 3 March 2020). https://www.nationthailand.com/news/30383253

Nation Thailand (2020c). 'IT Crime Taskforce Looking into Negative Social Media Posts on Cash Handout Scheme', The Nation Thailand (online, 11 April 2020) https://www.nationthailand.com/news/30385786

NNBT (2020). 'Government Dismisses Rumors of House Dissolution to Avert Censure Debate', National News Bureau of Thailand (online, 20 January 2020). https://forum.thaivisa.com/topic/1144117-government-dismisses-rumors-of-house-dissolution-to-avert-censure-debate

Nyi Kyaw, (2019). 'Facebooking in Myanmar: From Hate Speech to Fake News to Partisan Political Communication' (Perspective No 36, ISEAS Yusof Ishak Institute, 9 May 2019). https://www.iseas.edu.sg/images/pdf/ISEAS_Perspective_2019_36.pdf

Ong, J.C. & J.V.A. Cabañes (2019). 'Politics and Profit in the Fake News Factory: Four Work Models of Political Trolling in the Philippines' (Research Report, NATO Strategic Communications Centre of Excellence, November 2019). https://nysean.org/blog/2019/12/12/politics-and-profit-in-the-fake-news-factory-four-work-models-of-political-trolling-in-the-philippines

Ong, J.C., Tapsell, R. & N. Curato (2019). 'Tracking Digital Disinformation in the 2019 Philippine Midterm Election' (Report, New Mandela, August 2019). https://www.newmandala.org/wp-content/uploads/2019/08/Digital-Disinformation-2019-Midterms.pdf

Paine, J. (2020) 'Industry Submission on Government Regulation No 80 Year 2019 on eCommerce ("GR 80") from Jeff Paine to Minister of Trade of Republic of Indonesia', 25 February 2020 in Minister of Trade of Republic of Indonesia. https://aicasia.org/wp-content/uploads/2020/04/AIC_Industry-submission-on-Government-Regulation-no.-80-Year-2019-on-e-Commerce-%E2%80%9CGR-80%E2%80%9D-MOT-1.pdf

Phuket News (2020). 'Phuket Virus Death Fake News Poster Acknowledges Computer Crimes Act Charge', *Phuket News* (online, 17 February 2020) https://www.thephuketnews.com/phuket-virus-death-fake news-poster-acknowledges-computer-crimes-act-charge-74864.php

Reservations and Declarations for Treaty No.185 - Convention on Cybercrime Nature of Declarations: Reservations (as of 7 June 2020) https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/185/signatures?p_auth=YVsvM0qq.

Reuters (2019). 'Facebook Removes Hundreds of Indonesian Accounts Linked to Fake News and Hate Speech', *Reuters* (online, 1 February 2019) https://www.abc.net.au/

news/2019-02-01/facebook-removes-indonesian-accounts-fake-news-hate-speech/10772968.

Sakim, N. (2020). 'Facebook had role in 2017 Myanmar violence', *Anadolu Agency* (online, 23 August 2020) https://www.aa.com.tr/en/asia-pacific/facebook-had-role-in-2017-myanmar-violence/1950942.

Saraswati, M. S. (2017). '‘Buzzer’ Jonru Muzzled at Last?', Indonesia at Melbourne (Blog Post, 10 October 2017) https://indonesiaatmelbourne.unimelb.edu.au/buzzer-jonru-muzzled-at-last/.

Smith, R.B. (2020), 'Fake News in ASEAN' (Master of Philosophy Thesis, University of New England, Australia).

Smith, R.B. and Perry, M (2020). '“Fake News” Legislation in Thailand: The Good, the Bad and the Ugly' in *Athens Journal of Law* 6(3):243-264.

Souisa, H. (2018). 'Misinformation, Ratna the Hoaxer, and 1965', Indonesia at Melbourne (Blog Post, 8 October 2018) https://indonesiaatmelbourne.unimelb.edu.au/fake news-ratna-the-hoaxer-and-1965/.

Stecklow, S. (2018). 'Inside Facebook’s Myanmar Operation: Hatebook - A Reuters Special Report', *Reuters Investigates* (online, 15 August 2018) https://www.reuters.com/investigates/special-report/myanmar-facebook-hate

Tanakasempipat, P. (2021). 'Facebook Removes Thai Military-linked Information Influencing Accounts', *Reuters* (online, 4 March 2021). https://www.reuters.com/article/us-facebook-thailand-idUSKBN2AV252

Tapsell, R. (2019) 'Indonesia’s Policing of Hoax News Increasingly Politicised' in *ISEAS Perspective* No 75.

Temby, Q. (2019). 'Disinformation, Violence, and anti-Chinese Sentiment in Indonesia’s 2019 Elections' in *ISEAS Perspective* No 67.

ThaiPBS World (2019). 'PM Complains of Fake News About His Wife’s Religion', Thai PBS World (online, 17 December 2019) https://www.thaipbsworld.com/pm-complains-of-fake news-about-his-wifes-religion

Thai PBS World (2020). 'Four People in Custody Accused of Spreading Fake News about Coronavirus', Thai PBS World (online, 17 February 2020) https://www.thaipbsworld.com/four-people-in-custody-accused-of-spreading-fake news-about-coronavirus

Thai Visa (2020a). 'Covid-19: Tech Police go after Dozens of Fake Websites - Thais Trick Thais Slipping up in Their Own Language!', *thaivisa* News (online, 30 March 2020). https://forum.thaivisa.com/topic/1156655-covid-19-tech-police-go-after-dozens-of-fake-websites-thais-trick-thais-slipping-up-in-their-own-language

Thai Visa (2020b). 'Watch Yourself! Court Orders Thai Held for 12 Days After "Fake" Covid-19 Post - He is Later Given Bail', *thaivisa* (online, 25 March 2020) https://forum.thaivisa.com/topic/1155731-watch-yourself-court-orders-thai-held-for-12-days-after-fake-covid-19-post-he-is-later-given-bail

Thongtub, E. (2020a) 'Australian, American Among Four Charged for ‘Abandoned Phuket Zoo’ Post', *Phuket News* (online, 23 April 2020). https://www.thephuketnews.com/australian-american-among-four-charged-for-abandoned-phuket-zoo-post-75810.php

Thongtub, E. (2020b) 'Two Charged for Fake News Posts of Police Levying B200 Fines for not Wearing Face Masks', *Phuket News* (Online, 27 March 2020) https://www.thephuketnews.com/two-charged-for-fake-news-posts-of-police-levying-b200-fines-for-not-wearing-face-masks-75474.php

United Nations Human Rights Council (2018). Report of the Detailed Findings of the Independent International Fact-Finding Mission on Myanmar, A/HRC/39/CRP.2 Comm, UN Doc A/HRC/39/CRP.2 (18 September 2018).

Wardle, C. (2017). 'Fake News. It's Complicated', First Draft (Research Report, 16 February 2017). https://firstdraftnews.org/latest/fake news-complicated

World Bank and United Nations (2017). 'Combatting Cybercrime: Tools and Capacity Building for Emerging Economies', (Published Report, World Bank and United Nations, 2017). https://www.itu.int/en/ITU-D/Cybersecurity/Documents/worldbank-combating-cybercrime-toolkit.pdf

Treaties & Legislation

Additional Protocol to the Convention on Cybercrime, Concerning the Criminalisation of Acts of a Racist and Xenophobic Nature Committed Through Computer Systems, opened for signature 28 January 2003, UNTS No A-40916 (entered into force 1 March 2006).

ASEAN Convention Against Trafficking in Persons, Especially Women and Children, opened for signature 21 November 2015 (entered into force 8 March 2016).

Convention on Cybercrime, opened for signature 23 November 2001, UNTS No 40916 (entered into force 1 July 2004).

Electronic Commerce (EC Directive) Regulations 2002 (up to date as of 15 July 2020) (United Kingdom).

Framework and Joint Declaration to Minimise the Harmful Effects of Fake News (14th Conference of the ASEAN Ministers Responsible for Information (AMRI), signed 10 May 2018.

Government Regulation Number 80 Year 2019 on Trade Through Electronic System (Indonesia) [tr Minister of State Secretariat].

Law Concerning Electronic Information and Transactions 2008 (as amended by Law 19 of 2016) (Indonesia) [tr Wishnu Basuki].

Protection from Online Falsehoods and Manipulation Act 2019 (current version as of 21 June 2020) (Singapore).