MASTER'S THESIS

Prohibiting Atrocity Speech on Social Media Platforms

A research on State obligations under international human rights law with respect to the regulation of social media platforms regarding hate speech leading to breaches of jus cogens

Sajet, R.

Award date:
2022

Awarding institution:
Department of Public Law

Prohibiting Atrocity Speech on Social Media Platforms

A research on State obligations under international human rights law with respect to the regulation of social media platforms regarding hate speech leading to breaches of *jus cogens*

Name student: Ruth Sajet

e-mail: ruth.sajet@gmail.com

Student number: 838367882

Course: Master Thesis International Law (RM9906)

Essay supervisor: prof. dr. Gleider Hernández

Examiner: dr. Eduardo Arenas Catalan

Wordcount: 14822

(excluding title page, table of contents, table of abbreviations, footnotes, bibliography and register)

Date of submission: 26 June 2022

Contents

Abbreviations

ARSIWA	Draft Articles on the Responsibility of States for Internationally Wrongful Acts
CJEU	Court of Justice of the European Union
CPPCG	Convention on the Prevention and Punishment of the Crime of Genocide
DSA	Digital Services Act
ECHR	European Convention for the Protection of Human Rights and Fundamental Freedoms
ECtHR	European Court of Human Rights
EU	European Union
GDPR	General Data Protection Regulation
HRC	Human Rights Council
ICC	International Criminal Court
ICCPR	International Covenant on Civil and Political Rights
ICERD	International Convention on the Elimination of All Forms of Racial Discrimination
ICESR	International Covenant on Economic, Social and Cultural Rights
ICJ	International Court of Justice
ICTR	International Criminal Tribunal for Rwanda
ICTY	International Criminal Tribunal for the former Yugoslavia
IHRL	International Human Rights Law
IIFFMM	Independent International Fact-Finding Mission on Myanmar
OB	(Meta) Oversight Board
SMCs	Social Media Companies
TEU	Treaty on European Union
TFEU	Treaty on the Function of the European Union
TNCs	Transnational Corporations
UDHR	Universal Declaration of Human Rights
UN	United Nations
UNGPs	UN Guiding Principles on Business and Human Rights
US	United States of America
VCLT	Vienna Convention on the Law of Treaties

1. Introduction

Much of the world's dissemination of ideas and information takes place in cyberspace, on social media platforms owned and managed by large technology companies, which are often incorporated under US law. The largest social media platform is Facebook, with nearly 3 billion monthly active users. YouTube is the second largest platform with 2.6 billion monthly users. Facebook is owned by Meta Platforms, Inc. (Meta), which also owns Instagram (1.4 billion users) and Whatsapp (2 billion users). YouTube is owned by Alphabet Inc. (Alphabet), which also owns Google. Another significant platform is Twitter, with 436 million monthly users.[1] Meta, Alphabet and Twitter are private entities incorporated under US law,[2] but offer their services on a global scale. As such, these technology companies operate in more than one country and are considered transnational corporations (TNCs).[3] For the sake of clarity, in this thesis this type of TNC will be referred to as Social Media Company or SMC.

SMCs are private entities with a great commercial interest in keeping users engaged on their platforms.[4] According to the user policies of the SMCs, the platforms are vehicles for the freedom of expression.[5] Billionaire Elan Musk said in a recent statement announcing his acquisition of Twitter: 'Free speech is the bedrock of a functioning democracy, and Twitter is the digital town square where matters vital to the future of humanity are debated.'[6] Indeed, the freedom of expression is an essential human right and as such enshrined in major international human rights treaties.[7] It is not, however, an absolute right. Unfettered expression may become abusive and insulting towards others and incite violence, hatred or discrimination. This type of harmful expression is considered hate speech and may infringe on the right to equality and the prohibition of discrimination.[8] For this reason, international human rights law (IHRL) allows for restrictions on the freedom of expression.[9] In accordance with some international human rights treaties, hate speech may even be prohibited.[10]

It matters that the largest SMCs are incorporated under US law. The freedom of expression is guaranteed by the First Amendment[11] which also protects hate speech.[12] SMCs, as private entities, also enjoy First Amendment rights.[13] Consequently, the US currently does not regulate hate speech on SMC platforms within US territory, and certainly not extraterritorially. Furthermore, SMCs are

[1] Statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users.

[2] See for more information about Meta: About.facebook.com/company-info, for more information about Alphabet: Abc.xyz/investor/#product-and-business-updates, and for more information about Twitter: Investor.twitterinc.com/home/default.aspx.

[3] In the United Nations Guiding Principles on Business and Human Rights, UN Doc. HR/PUB/11/04 (2011) (UNGPs), reference is made to transnational corporations. The term transnational corporations is interchangeable with the term multinational corporations.

[4] Oremus et al., *The Washington Post* 26 October 2021.
See also: Amnesty.org/en/latest/campaigns/2022/02/what-is-big-techs-surveillance-based-business-model.

[5] The relevant policies of Facebook: Transparency.fb.com/en-gb/policies/community-standards, Youtube: Transparencyreport.google.com/youtube-policy/featured-policies/hate-speech?hl=en, and Twitter: Help.twitter.com/en/rules-and-policies/defending-and-respecting-our-users-voice.

[6] Prnewswire.com/news-releases/elon-musk-to-acquire-twitter-301532245.html.

[7] For instance: article 19 of the International Covenant of Civil and Political Rights (ICCPR) and article 10 of the European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR). Freedom of expression can also be found in article 19 of the Universal Declaration of Human Rights (UDHR).

[8] Article 7 UDHR, article 26 ICCPR and article 14 ECHR.

[9] Article 29(2) UDHR, article 19(3) ICCPR and article 10(2) ECHR.

[10] Article 20(2) ICCPR and article 4(a) of the International Convention of the Elimination of all Forms of Racial Discrimination (CERD).

[11] Constitution.congress.gov/constitution/amendment-1.

[12] Ala.org/advocacy/intfreedom/hate. Other content that can be considered problematic, but legal under the US First Amendment, is election misinformation, medical misinformation, porn, graphic violence and terrorist manifestos. See for instance Fried, *Axios.com* 28 April 2022.

[13] Strossen 2018, p. 56.

granted broad protection against legal liability for any content posted by users on their platforms through Section 230 of the US Communications Decency Act.[14] However, SMCs regulate their platforms through user policies that include provisions on harmful content, and Meta has even constituted an Oversight Board (OB) to review its content decisions.[15] Even so, one cannot help but wonder if leaving the regulation of hate speech to the scrutiny of SMCs provides sufficient protection of fundamental human rights, especially where hate speech leads to detrimental consequences such as violations of peremptory norms.

This type of hate speech may be called atrocity speech.[16] An example is direct and public incitement to commit genocide.[17] In case law of international tribunals[18] "public" is defined as a public place or mass media.[19] SMCs provide a new form of mass media[20] as their platforms are widely accessible. In Myanmar, Facebook was the dominant medium for widespread atrocity speech targeting Muslims, in particular Rohingya.[21]

International human rights law (IHRL) creates obligations for the State to ensure human rights in relation to individuals within its territory and subject to its jurisdiction.[22] The State also has a duty to protect human rights against acts committed by private entities.[23] International law does not impose IHRL duties directly on private entities, unless it concerns violations of peremptory norms.[24] When it concerns hate speech on a social media platform, another challenge is posed. The actual abuse is committed by one user against another (group of) user(s) in cyberspace. It is very difficult to pinpoint which State has jurisdiction because of the borderless nature of cyberspace.[25] In this light, the European Union (EU) aims to solve issues regarding the regulation of SMCs and the question of State jurisdiction within the EU. It will soon implement the Digital Services Act (DSA),[26] to create a safer digital space in which the fundamental rights of all users of digital services are protected.[27]

This thesis aims to answer the following question: **to what extent do States have international human rights law obligations with respect to the regulation of social media platforms regarding hate speech, especially when atrocity speech leads to breaches of *jus cogens*?** The primary focus of this thesis are SMCs incorporated under US law and how this affects IHRL obligations with respect to the regulation of their platforms. In chapter two the IHRL framework regarding the freedom of expression will be analysed, including the conditions under which expression should be protected, may be restricted or must be prohibited. In chapter three the applicability of international law in cyberspace will be discussed, as well as the question of extraterritorial jurisdiction in cyberspace. Then State responsibility with a focus on international law and US domestic law regimes will be discussed, as well as the responsibility of the SMCs with respect to the regulation of expression on their platforms. Chapter four will consist of a case study of the role of Facebook as dominant platform in Myanmar for the dissemination of atrocity speech. Finally, in chapter five Meta's OB and the EU's DSA will be discussed as possible solutions for the regulation of social media platforms regarding hate speech.

---

[14] Smith & Van Alstyne, *Harvard Business Review* 12 August 2021, see also 47 U.S.C. § 230.
[15] Oversightboard.com.
[16] Gordon 2017, p. 5.
[17] Article III(c) of the Convention on the Prevention and Punishment of the Crime of Genocide (CPPCG).
[18] Such as the International Military Tribunal at Nuremberg, the International Criminal Tribunal for the former Yugoslavia (ICTY) and the International Criminal Tribunal for Rwanda (ICTR).
[19] ICTR 2 September 1998, Case No. ICTR-96-4-T, (*Prosecutor/Akayesu*), Judgement, para 556.
[20] Oster 2017, p. 1
[21] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1310.
[22] Article 2(1) ICCPR.
[23] UN Human Rights Committee, General Comment No. 31 (2004, CCPR/C/21/Rev.1/Add.13), para 8.
[24] Ruggie 2013, p. 68-71.
[25] Solmone 2020, p. 24.
[26] COM(2020)825 final.
[27] Digital-strategy.ec.europa.eu/en/policies/digital-services-act-package.

2. Hate Speech in International Law

2.1. Introduction

There is no clear definition of hate speech in international law.[28] According to Stephanie Farrior, hate speech refers to forms of expression that are 'abusive, insulting, intimidating, harassing and/or which incite to violence, hatred or discrimination against a group or a person based on race, religion, ethnicity or national origin.'[29] In 2019, UN Secretary General António Guterres stated that hate speech can lay the foundation for violence, and that it has been a precursor to atrocity crimes, including genocide, over the past 75 years.[30] According to Gregory S. Gordon, the development of international hate speech law is fragmented.[31] As the term "international hate speech law" does not adequately capture incitement to breaches of *jus cogens* such as the prohibition of genocide or the prohibition of crimes against humanity, Gordon proposes to designate the body of international rules and jurisprudence regarding incitement to particularly heinous or mass violence as *atrocity speech law*.[32] This distinction suits the purposes of this thesis, as *hate speech* can be considered to cover a much broader spectrum of abusive expression, while *atrocity speech* can be applied to cover expression of such a nature that it may lead to particular breaches of *jus cogens*.

The issue of hate speech should be considered in light of both the human right to freedom of expression and the human right to equality, which includes the prohibition against discrimination. In international human rights law (IHRL), exercising one's freedom of expression can be limited if expressions of hatred violate the rights and freedoms of others. In this chapter, three distinctions will be made between forms of expression; those that should be protected from restriction, those that may be restricted, and those that must be prohibited.[33]

In principle the freedom of expression should not be restricted, even if the expression is considered deeply offensive.[34] Section 2.4 will provide an overview of the freedom of expression and the conditions for restricting this human right. Section 2.5 will explore forms of expression that may be restricted. This could be the case if the expression presents a danger to others and their enjoyment of human rights, and may be considered abusive, and as such hate speech. However, restrictions may only be applied under certain conditions. Finally, section 2.6 will discuss the forms of expression that must be prohibited.[35] This category includes atrocity speech. The chapter will first lay out the general IHRL framework, the major treaties relevant to this thesis and an explanation of peremptory norms of international law, or *jus cogens*.

[28] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1307.
[29] Farrior, *Berkeley Journal of International Law* 1996**,** p. 3.
[30] Un.org/sg/en/content/sg/statement/2019-06-18/secretary-generals-remarks-the-launch-of-the-united-nations-strategy-and-plan-of-action-hate-speech-delivered.
[31] Gordon 2017, p. 3.
[32] Gordon 2017, p. 5.
[33] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1305.
[34] UN Human Rights Committee, General Comment No. 34 (2011, CCPR/C/GC/34), para 11.
[35] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1306.

2.2. General Framework of International Human Rights Law

International human rights law (IHRL) is a specialist regime within international law.[36] Its formal sources coincide with those listed in article 38(1) of the International Court of Justice (ICJ) Statute;[37] treaties, customary international law and general principles are seen as hierarchically equal sources.[38]

Article 1(3) of the UN Charter sets out to promote and encourage respect for human rights and for fundamental freedoms for all without distinction as to race, sex, language, or religion.[39] Article 55(c)[40] of the UN Charter in conjunction with article 56[41] reiterate article 1(3) and add the pledge of Member States to take joint and separate action in this respect. In 1948, the UN General Assembly adopted the Universal Declaration of Human Rights (UDHR), a non-binding resolution.[42] The UDHR is considered a cornerstone of IHRL,[43] as its provisions have served as basis for major multilateral treaties and domestic human rights law. Many provisions have been determined to be customary international law by international courts and tribunals,[44] making those provisions relevant to States that are not parties, or have made reservations, to the major IHRL treaties.[45]

International human rights law lays down obligations which States are bound to respect, to protect and to fulfil,[46] in relation to individuals within their jurisdiction and/or territory.[47] The duty to respect entails a State's negative action, to refrain from violating a human right. The duty to protect requires the State to proactively ensure that persons do not suffer human rights abuses. Under the duty to fulfil States must take positive action to ensure the greater enjoyment of human rights, which includes the requirement of a remedy.[48] States may ensure the respect, protection and

---

[36] Chinkin 2018, p. 63.

[37] Chinkin 2018, p. 65. Article 38(1) ICJ Statute: 'The Court, whose function is to decide in accordance with international law such disputes as are submitted to it, shall apply: a. international conventions, whether general or particular, establishing rules expressly recognized by the contesting states; b. international custom, as evidence of a general practice accepted as law; c. the general principles of law recognized by civilized nations; d. […] judicial decisions and the teachings of the most highly qualified publicists of the various nations, as subsidiary means for the determination of rules of law.'

[38] Hernández 2019, p. 34.

[39] The full text of article 1(3) UN Charter states: 'to achieve international cooperation in solving international problems of an economic, social, cultural, or humanitarian character, and in promoting and encouraging respect for human rights and for fundamental freedoms for all without distinction as to race, sex, language, or religion.'

[40] Article 55(c) of the UN Charter: 'the United Nations shall promote universal respect for, and observance of, human rights and fundamental freedoms for all without distinction as to race, sex, language, or religion'

[41] Article 56 of the UN Charter: 'all Members pledge themselves to take joint and separate action in co-operation with the Organization for the achievement of the purposes set forth in Article 55.'

[42] Chinkin 2018, p. 80.

[43] Gordon 2017, p. 62.

[44] For instance, in para 91 of the ICJ 24 May 1980 *Case Concerning United States Diplomatic and Consular Staff in Tehran (United States/Iran)* the ICJ stated that 'to deprive human beings of their freedom and to subject them to physical constraint in conditions of hardship is in itself manifestly incompatible with the principles of the Charter of the United Nations, as well as with the fundamental principles enunciated in the Universal Declaration of Human Rights.'

[45] Schabas 2021, p. 3.

[46] UN Human Rights Committee, General Comment No. 31 (2004, CCPR/C/21/Rev.1/Add.13), para 5-7. See also Ohchr.org/en/professionalinterest/pages/internationallaw.aspx for a concise explanation of these State obligations.

[47] Mégret 2018, p. 86 and p. 97. See also article 2(1) ICCPR: 'Each State Party to the present Covenant undertakes to respect and to ensure to all individuals within its territory and subject to its jurisdiction the rights recognized in the present Covenant, without distinction of any kind, such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status.'

[48] Mégret 2018, p. 97-99. Article 8 UDHR states that 'everyone has the right to an effective remedy by the competent national tribunals for acts violating the fundamental rights granted him by the constitution or by law.'

fulfilment of human rights obligations by implementing human rights into their domestic legal order.[49]

Modern IHRL finds its basis primarily in treaties.[50] IHRL treaties are different from other treaties. Although States are bound IHRL treaties once they have expressly consented, IHRL treaties are not based on State reciprocity. Instead, they require States to implement human rights within their territory and jurisdiction.[51]

The International Covenant on Civil and Political Rights (ICCPR) is a major treaty that forms the International Bill of Human Rights together with the International Covenant on Economic, Social and Cultural Rights (ICESR)[52] and the UDHR. The ICCPR consists of human rights that require negative action from the State, a duty not to intervene.[53] The Human Rights Committee (HRC) has the competence to consider State complaints and individual complaints regarding violations of the ICCPR. The HRC also regularly publishes general comments with interpretations of the ICCPR provisions.[54] The ICJ also has competence to settle disputes between States with respect to the interpretation of the ICCPR.[55]

The European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR) is an important regional treaty[56] which refers directly to the UDHR,[57] and consists of civil and political rights. The ECHR is distinctive from other international human rights instruments as it establishes the European Court of Human Rights (ECtHR), which has compulsory jurisdiction.[58] Complaints can be lodged by individuals or by member States, and those States are under an obligation to comply with judgements.[59] The ECtHR develops jurisprudence in the European context. However, it interprets the ECHR in conjunction with principles and rules of general international (human rights) law, in accordance with article 31(3)(c) VCLT.[60] ECtHR case law is not only important for national courts, it also contributes to the decisions of international courts and tribunals.[61]

2.3. Jus cogens

Another category of sources, not specified in the ICJ Statute, are peremptory norms of general international law, or *jus cogens*.[62] Peremptory norms are defined in article 53 VCLT:

> […] a peremptory norm of general international law is a norm accepted and recognized by the international community of States as a whole as a norm from which no derogation is permitted and which can be modified only by a subsequent norm of general international law having the same character.

The United Nations International Law Commission (ILC) has produced five reports on *jus cogens*. In the second report, the ILC recognises *jus cogens* to be a category of norms that is hierarchically

[49] Byrnes & Renshaw 2018, p. 484.
[50] Chinkin 2018, p. 65.
[51] Chinkin 2018, p. 66.
[52] Chinkin 2018, p. 66. The ICESCR is also a major IHRL treaty based on the UDHR, however, it contains a catalogue of human rights that is not directly relevant to this thesis.
[53] Van Boven 2018, p. 137.
[54] Ohchr.org/EN/HRBodies/CCPR/Pages/CCPRIntro.aspx.
[55] Article 44 ICCPR.
[56] The American Convention on Human Rights and the African Charter on Human and People's Rights are also regional IHRL treaties, but they fall outside the scope of this thesis.
[57] ECHR, Preamble, second paragraph.
[58] Šturma 2020, para 1.
[59] Echr.coe.int/documents/questions_answers_eng.pdf.
[60] ECtHR 21 November 2001, ECLI:CE:ECHR:2001:1121JUD003576397 (*Al-Adsani/United Kingdom*), para 55. Article 31(3)(c) of the Vienna Convention on the Law of Treaties (VCLT) states that, as a general rule of interpretation of treaties, together with the context, 'any relevant rules of international law applicable in the relations between the parties' shall be taken into account.
[61] Van Aaken, Motoc & Vasel 2018, p. 2.
[62] In this thesis, the terms *jus cogens* and "peremptory norms" will be used interchangeably.

superior to other sources of international law.[63] The reference to the 'international community of States as a whole' in Article 53 VCLT is not to be interpreted as unanimous consensus of all States, but rather that the recognition of the peremptory nature of the norm cannot be vetoed by an individual State.[64] States are bound by *jus cogens* whether they have explicitly consented to these norms or not.[65] Peremptory norms are also referred to as intransgressible, fundamental or cardinal principles of customary international law by the ICJ.[66] Through case law and judgements of national, regional and international courts certain norms have now been recognised as *jus cogens*, which the ILC has specified in its Fourth report: 1) the prohibition of aggression or aggressive force; 2) the prohibition of genocide; 3) the prohibition of slavery; 4) the prohibition of apartheid and racial discrimination; 5) the prohibition of crimes against humanity; 6) the prohibition against torture; 7) the right to self-determination; and 8) basic rules of international humanitarian law.[67] Except for the right to self-determination, these norms are also included in articles 5 - 8 of the Rome Statute of the International Criminal Court.[68] A breach of the above listed prohibitions can lead to criminal liability, States carry a negative obligation to ensure that these rights are not violated.[69]

State obligations in this respect are considered obligations *erga omnes*, or obligations that are directed towards the international community as a whole.[70] The ICJ determined in its Barcelona traction case that 'in view of the importance of the rights involved, all States can be held to have a legal interest in their protection.'[71] According to article 48 ARSIWA,[72] any State other than an injured State is entitled to invoke the responsibility of another State if the obligation breached is owed to the international community as a whole. The ILC refers to examples of *jus cogens* in its commentary to article 48, as it is seen as the legal interest of the international community as a whole to protect human rights.[73] In these matters, universal jurisdiction could be applied,[74] meaning that every State would have the jurisdiction to try suspects of abhorrent offences such as genocide.[75] States also have a duty to prevent certain acts from occurring.

Some peremptory norms have been codified in multilateral treaties. The International Convention on the Elimination of All Forms of Racial Discrimination (ICERD) pertains to the prohibition of racial discrimination. Article 4(a) ICERD pertains to hate speech and requires State parties to criminalise

> all dissemination of ideas based on racial superiority or hatred, incitement to racial discrimination, as well as all acts of violence or incitement to such acts against any race or group of persons of another colour or ethnic origin.

---

[63] ILC, 'Second report on *jus cogens*' (Special Rapporteur Dire Tladi) UN Doc A/CN.4/706 (16 March 2017), para 23-27.

[64] ILC, 'Second report on *jus cogens*' (Special Rapporteur Dire Tladi) UN Doc A/CN.4/706 (16 March 2017), para 67.

[65] Chinkin 2018, p.73.

[66] See for instance ICJ 8 July 1996, *Legality of the Threat or Use of Nuclear Weapons (Advisory Opinion)*, para. 79, and ICJ 27 June 1986, *Case Concerning Military and Paramilitary Activities In and Against Nicaragua (Nicaragua/United States of America)* para 190.

[67] ILC 'Fourth report on *jus cogens*' (Special Rapporteur Dire Tladi) UN Doc A/CN.4/727 (31 January 2019), para 60.

[68] Article 5 of the Rome Statute contains the list of crimes that fall within the jurisdiction of the ICC, article 6 – 8 define the crimes and the acts that fall under these crimes.

[69] De Wet 2013, p 548.

[70] De Wet 2013, p. 554.

[71] ICJ 5 February 1970, *Case Concerning the Barcelona Traction, Light and Power Company, Limited (Belgium/Spain)* (Second Phase), para 33.

[72] The 2001 draft Articles on the Responsibility of States for Internationally Wrongful Acts. Although the ARSIWA are not a convention, they are considered to be a codification of customary international law, and as such enjoy a high level of authority.

[73] *Report of the International Law Commission 2001***,** Commentary on Article 48, para. 9, p. 127.

[74] Orakhelashvili 2006, p. 288.

[75] Hernández 2019, p. 208.

The Convention on the Prevention and Punishment of the Crime of Genocide (CPPCG) defines genocide in its article II as 'an act committed with intent to destroy, in whole or in part, a national, ethnical, racial or religious group.' Article III(c) CPPCG criminalises atrocity speech, being direct and public incitement to commit genocide. In section 2.6 the link between atrocity speech and gross violations of *jus cogens* will be discussed.

2.4. Protecting the Freedom of Expression

The freedom of expression is one of the fundamental freedoms,[76] and part of the collection of civil and political rights. Article 19(2) ICCPR states that

> everyone shall have the right to freedom of expression; this right shall include freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of his choice.

The freedom of expression is also included in article 19 UDHR[77] and article 10(1) ECHR,[78] and is regarded as a first-generation right. This means that the State should refrain from violating this right.[79] However, the majority of human rights are not absolute, under certain circumstances limitations are accepted.[80] Article 29(2) UDHR states that

> in the exercise of his rights and freedoms, everyone shall be subject only to such limitations as are determined by law solely for the purpose of securing due recognition and respect for the rights and freedoms of others and of meeting the just requirements of morality, public order and the general welfare in a democratic society.

The restriction clause in the UDHR has a general application. Although similar formulations are used in the ICCPR and in the ECHR, the restriction clauses are added to specific articles in the respective treaties. The freedom of expression can be limited by means of article 19(3) ICCPR[81] and article 10(2) ECHR.[82] According to these restriction clauses, human rights may be subject to limitations that are in accordance with a three-part test: 1) prescribed by law, which refers to a State's domestic legal system, 2) have a legitimate aim, which refers to a pressing social need within the State's jurisdiction, and 3) are necessary (in a democratic society),[83] which means that the restriction must

[76] The other fundamental freedoms are the freedoms of thought, association and assembly. See: McGoldrick 2018, p. 208.
[77] Article 19 UDHR: 'everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers.'
[78] Article 10(1) ECHR: 'Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. This Article shall not prevent States from requiring the licensing of broadcasting, television or cinema enterprise.'
[79] Hernández 2019, p. 410.
[80] Mégret 2018, p. 99-100.
[81] Article 19(3) ICCPR: 'The exercise of the rights provided for in paragraph 2 of this article carries with it special duties and responsibilities. It may therefore be subject to certain restrictions, but these shall only be such as are provided by law and are necessary: (a) For respect of the rights or reputations of others; (b) For the protection of national security or of public order (*ordre public*), or of public health or morals.'
[82] Article 10(2) ECHR: '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.'
[83] The restriction clauses of articles 8-11 ECHR include the wording "democratic society," whereas ICCPR restriction clauses do not include the wording "democratic society."

be reasonable and proportionate to the goal sought.[84] Tensions can arise between specific human rights, but also between individuals and societal needs. The restriction clauses ensure that these tensions can be weighed.

Freedom of expression includes political and religious discourse, commentary on one's own and public affairs, discussion on human rights, teaching and journalism.[85] The HRC considers a free, uncensored and unhindered press or other media as a cornerstone of a democratic society, and essential to ensure the freedom of (opinion and) expression and the enjoyment of other human rights.[86] Expression may even be regarded as deeply offensive.[87] The ECtHR has famously considered in the *Handyside* judgement that expression may *offend, shock or disturb*:

> [f]reedom of expression constitutes one of the essential foundations of such a [democratic] society, one of the basic conditions for its progress and for the development of every man. Subject to paragraph 2 of Article 10, it is applicable not only to "information" or "ideas" that are favourably received or regarded as inoffensive or as a matter of indifference, but also to those that offend, shock or disturb the State or any sector of the population. Such are the demands of that pluralism, tolerance and broadmindedness without which there is no "democratic society".[88]

In later cases the ECtHR repeats its views with respect to pluralism, tolerance and broadmindedness in a democratic society.[89] The ECtHR does not look favourably to governments suppressing expression,[90] unless it is necessary due to the existence of a pressing social need.[91] This pressing social need may be established if one infringes on the fundamental rights of others by means of hate speech.[92]

2.5. Restricting Expression

The freedom of expression may be restricted if it presents a serious danger for others and for their enjoyment of human rights.[93] An important human right in this respect is the right to equality, which includes the prohibition of discrimination. This right is considered so fundamental that it is enshrined in article 1(2) and 1(3) of the UN Charter,[94] as well as article 7 UDHR, article 26 ICCPR and article 14 ECHR.[95] In its preamble, the ICERD declares that all human beings are equal before the law, and are entitled to protection against any incitement to discrimination.

Article 4 ICERD and article 20(2) ICCPR are regarded as anti-hate speech provisions.[96] Article 4 ICERD requires State Parties

---

[84] Mégret 2018, p. 100-101. See also: ECtHR 7 December 1976, ECLI:CE:ECHR:1976:1207JUD000549372 (*Handyside/United Kingdom*), para 48-49.
[85] UN Human Rights Committee, General Comment No. 34 (2011, CCPR/C/GC/34), para 11.
[86] UN Human Rights Committee, General Comment No. 34 (2011, CCPR/C/GC/34), para 12.
[87] UN Human Rights Committee, General Comment No. 34 (2011, CCPR/C/GC/34), para 11.
[88] ECtHR 7 December 1976, ECLI:CE:ECHR:1976:1207JUD000549372 (*Handyside/United Kingdom*), para 49.
[89] For instance: ECtHR 9 June 1998, ECLI:CE:ECHR:1998:0609JUD002267893 (*Incal/Turkey*), para 46.
[90] Gordon 2017, p. 68, with reference to ECtHR 23 September 1994, ECLI:CE:ECHR:1994:0923JUD001589089 (*Jersild/Denmark*).
[91] ECtHR 26 April 1979, ECLI:CE:ECHR:1979:0426JUD000653874 (*Sunday Times/United Kingdom*), para 67.
[92] ECtHR 8 July 1999, ECLI:CE:ECHR:1999:0708JUD002668295 (*Sürek/Turkey (No. 1)*), para 62.
[93] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1306.
[94] Article 1(1) UN Charter: 'to develop friendly relations among nations based on respect for the *principle of equal rights* and self-determination of peoples, and to take other appropriate measures to strengthen universal peace; and article 1(3) UN Charter: 'to achieve international co-operation in solving international problems of an economic, social, cultural, or humanitarian character, and in promoting and encouraging respect for human rights and *for fundamental freedoms for all without distinction* as to race, sex, language, or religion.'
[95] Article 14 ECHR pertains to the prohibition of non-discrimination, and, although not worded explicitly in the article, encompasses the right to equality.
[96] Gordon 2017, p. 76.

> [to] condemn all propaganda […] based on ideas or theories of one race or group of persons of one colour or ethnic origin, or which attempt to justify or promote racial hatred and discrimination in any form, and undertake to adopt immediate and positive measures designed to eradicate all incitement to, or acts of, such discrimination […].

Article 20(2) ICCPR states that 'any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law.' However, according to a UN report, the threshold of the types of expression that fall under article 20(2) ICCPR is determined by the severity of hatred, the intent of the speaker, the content, form and context of the speech as well as its reach and size of audience. Imminence of the acts called for by the speech and the likelihood of harm occurring are also important determinants.[97] Article 19(3) ICCPR is deemed applicable to restrictions in accordance with article 20(2) ICCPR and article 4 ICERD.[98]

The ECHR does not contain a specific prohibition to hate speech, but notes in case law that

> tolerance and respect for the equal dignity of all human beings constitute the foundations of a democratic, pluralistic society. That being so, as a matter of principle it may be considered necessary in certain democratic societies to sanction or even prevent all forms of expression which spread, incite, promote or justify hatred based on intolerance […], provided that any 'formalities', 'conditions', 'restrictions' or 'penalties' imposed are proportionate to the legitimate aim pursued.[99]

In ECtHR case law, the speech infringement by a State usually passes the first criterion of the three-part test, being prescribed by law, and the second criterion, having a legitimate aim. However, the ECtHR often deems the third criterion, being necessary in democratic society, not satisfied.[100]

For instance, in the *Jersild* case, a Danish journalist was convicted under the Danish Penal Code for aiding and abetting three adolescents skinheads in disseminating hate speech.[101] During an interview conducted by the journalist for a television program, the teens made abusive and derogatory remarks about immigrants and ethnic groups in Denmark.[102] The journalist filed a complaint with the ECtHR following his conviction for violation of article 10 ECHR. In applying the three-part test, the ECtHR determines that the first two criteria were met. However, the ECtHR finds that

> the punishment of a journalist for assisting in the dissemination of statements made by another person in an interview would seriously hamper the contribution of the press to discussion of matters of public interest and should not be envisaged unless there are particularly strong reasons for doing so.[103]

The ECtHR deems the interference of Jersild's right to freedom of expression not "necessary in a democratic society;" in particular the means employed were disproportionate to the aim of protecting "the reputation or rights of others." The ECtHR concludes that the measures taken by the Danish government gave rise to a breach of article 10 ECHR.[104]

An example of a case where the ECtHR finds all three criteria of three-part test met is *Sürek (no. 1)*. Sürek was a major shareholder of a company owning a weekly review.[105] In one of the issues, two reader's letters were published regarding the Kurdish-Turkish conflict, for which Sürek was charged in his capacity as owner of the review for 'disseminating propaganda against the

---

[97] Report of the Special Rapporteur (Frank La Rue) on the promotion and protection of the right to freedom of opinion and expression, UN Doc. A/67/357 (7 September 2012), para 45.
[98] Report of the Special Rapporteur (Frank La Rue) on the promotion and protection of the right to freedom of opinion and expression, UN Doc. A/67/357 (7 September 2012), para 37-41.
[99] ECtHR 6 July 2006, ECLI:CE:ECHR:2006:0706JUD005940500 (*Erbakan/Turkey*), para 56.
[100] Gordon 2017, p. 70.
[101] ECtHR 23 September 1994, ECLI:CE:ECHR:1994:0923JUD001589089 (*Jersild/Denmark*), para 12 and 14.
[102] ECtHR 23 September 1994, ECLI:CE:ECHR:1994:0923JUD001589089 (*Jersild/Denmark*), para 10.
[103] ECtHR 23 September 1994, ECLI:CE:ECHR:1994:0923JUD001589089 (*Jersild/Denmark*), para 35.
[104] ECtHR 23 September 1994, ECLI:CE:ECHR:1994:0923JUD001589089 (*Jersild/Denmark*), para 37.
[105] ECtHR 8 July 1999, ECLI:CE:ECHR:1999:0708JUD002668295 (*Sürek/Turkey (No. 1)*), para 8.

indivisibility of the [Turkish] state and provoking enmity and hatred among the people.'[106] The ECtHR considers with respect to the "necessary in a democratic society" criteria that the adjective "necessary" implies the existence of a "pressing social need."[107] The ECtHR reiterates the essential role of the press in ensuring the proper functioning of democracy,[108] but has particular regard to the words used in the letters and the context in which they were published. It views the content of the letters as capable of inciting to further violence by instilling a deep-seated and irrational hatred against Turkish officials, which are depicted as responsible for alleged atrocities against the Kurds. It finds that the letters are 'an appeal to bloody revenge by stirring up base emotions and hardening already embedded prejudices which have manifested themselves in deadly violence.'[109] The ECtHR concludes that Sürek had the power to shape the editorial direction of the review, and

> was vicariously subject to the "duties and responsibilities" which the review's editorial and journalistic staff undertake in the collection and dissemination of information to the public and which assume an even greater importance in situations of conflict and tension. […] In view of the above considerations the Court concludes that the penalty imposed on the applicant as the owner of the review could reasonably be regarded as answering a "pressing social need."[110]

The ECtHR holds a high regard for the contribution of the press in the function of a democratic society. In *Sürek (no. 1)* the ECtHR restates that the government should not infringe on expression that may offend, shock or disturb. However, the ECtHR finds that speech surpasses this threshold if it entails hate speech and the glorification of violence,[111] in which case the expression may be restricted. These examples show us that ECtHR case law provides a helpful guidance to finding the right balance between freedom of expression and expression that may be restricted,[112] or even prohibited.

2.6. Prohibiting Expression

Atrocity speech is hate speech of such a nature that it may lead to gross violations of *jus cogens*.[113] Article III(c) CPPCG criminalises direct and public incitement to commit genocide. It is important to note that according to case law, general hate speech (inciting racial discrimination or violence) does not fall under article III(c) CPPCG, although it can precede or accompany direct and public incitement to commit genocide.[114] Similar provisions to article III(c) CPPCG can be found in article 4(3)(c) of Statute of the International Criminal Tribunal for the former Yugoslavia (ICTY), article 25(3)(d) of the Rome Statute, as well as article 2(3)(c) of the Statute of the International Criminal Tribunal for Rwanda (ICTR). The focus in this section will be on ICTR case law, as mass media played a pivotal role in the incitement to mass violence.[115]

Direct and public incitement to commit genocide was adjudicated in various cases of the ICTR.[116] In *Akayesu*, the ICTR defines direct and public incitement as

> directly provoking the perpetrator(s) to commit genocide, whether through speeches, shouting or threats uttered in public places or at public gatherings, or through the sale or

[106] ECtHR 8 July 1999, ECLI:CE:ECHR:1999:0708JUD002668295 (*Sürek/Turkey (No. 1)*), para 10-12.
[107] ECtHR 8 July 1999, ECLI:CE:ECHR:1999:0708JUD002668295 (*Sürek/Turkey (No. 1)*), para 58(ii).
[108] ECtHR 8 July 1999, ECLI:CE:ECHR:1999:0708JUD002668295 (*Sürek/Turkey (No. 1)*), para 59.
[109] ECtHR 8 July 1999, ECLI:CE:ECHR:1999:0708JUD002668295 (*Sürek/Turkey (No. 1)*), para 62.
[110] ECtHR 8 July 1999, ECLI:CE:ECHR:1999:0708JUD002668295 (*Sürek/Turkey (No. 1)*), para 63-64.
[111] ECtHR 8 July 1999, ECLI:CE:ECHR:1999:0708JUD002668295 (*Sürek/Turkey (No. 1)*), para 62.
[112] Gordon 2017, p. 67.
[113] Gordon 2017, p. 5
[114] ICTR 3 December 2003, Case No. ICTR-99-52-T (*Prosecutor/Nahimana et al.*), Judgement, para 692.
[115] Gordon 2017, p. 46. See also ICTR 4 September 1998, Case No. ICTR-97-23-S (*Prosecutor/Kambanda*), Judgement, para 39(vii), where former prime minister of Rwanda, Jean Kambanda, acknowledged giving support to Radio Television Libre des Mille Collines (RTLM) with the knowledge that this radio station broadcasted incitement to (among others) killing and persecution of Tutsi and moderate Hutu.
[116] Gordon 2017, p. 165.

> dissemination, offer for sale or display of written material or printed matter in public places or at public gatherings, or through the public display of placards or posters, or through any other means of audiovisual communication.[117]

The ICTR refers to the ILC's definition of "public incitement,"[118] stating that it is

> characterized by a call for criminal action to a number of individuals in a public place or to members of the general public at large by such means as the mass media, for example, radio or television.[119]

The ICTR opines that the "direct" element of incitement 'should be viewed in the light of its cultural and linguistic content.' The cultural aspect refers to the way a particular speech may be perceived from one country to another, and the linguistic content of the speech could be implicit but still be a direct incitement.[120] On a case-by-case basis this can be considered by 'focusing mainly on the issue of whether the persons for whom the message was intended immediately grasped the implication thereof.'[121] The ICTR also posits that the *mens rea* aspect of the crime requires that 'the person who is inciting to commit genocide must have himself the specific intent to commit genocide.'[122] The ICTR concludes that genocide is such a serious crime that direct and public incitement thereto must be punished, even if the incitement failed to produce the result expected by the perpetrator.[123]

The *Media Case* concerns the founders of the radio station *RTLM* and the founder of the newspaper *Kangura*,[124] the role of mass media,[125] and whether the defendants had exercised their right to free speech or infringed on the right to non-discrimination through advocacy of hatred.[126] To distil two criteria for the application of article III(c) CPPCG to mass media, the ICTR reviews international case law, including ECtHR case law,[127] on incitement to discrimination and violence.[128] The first criterion is whether editors and publishers of the media have a legitimate *purpose*, such as historical research, dissemination of news and information, and the public accountability of government authorities.[129] An illegitimate purpose would be the use of inflammatory language or incitement to violence.[130] The second criterion to take into account is *context*.[131] As the ICTR formulates it:

> A statement of ethnic generalization provoking resentment against members of that ethnicity would have a heightened impact in the context of a genocidal environment. It would be more likely to lead to violence. At the same time the environment would be an indicator that incitement to violence was the intent of the statement.[132]

The ICTR also stipulates that when the media disseminate views that constitute ethnic hatred and calls to violence for informative or educational purposes, it is necessary to show clear distancing from these views to ensure that no harm will result from the dissemination. It concludes that the

---

[117] ICTR 2 September 1998, Case No. ICTR-96-4-T, (*Prosecutor/Akayesu*), Judgement, para 559.
[118] International Law Commission Draft Code of Crimes Against the Peace and Security of Mankind, art. 2(3)(f); in: Report of the International Law Commission to the General Assembly, UN Doc A/51/10 (1996).
[119] ICTR 2 September 1998, Case No. ICTR-96-4-T, (*Prosecutor/Akayesu*), Judgement, para 556.
[120] ICTR 2 September 1998, Case No. ICTR-96-4-T, (*Prosecutor/Akayesu*), Judgement, para 557.
[121] ICTR 2 September 1998, Case No. ICTR-96-4-T, (*Prosecutor/Akayesu*), Judgement, para 558.
[122] ICTR 2 September 1998, Case No. ICTR-96-4-T, (*Prosecutor/Akayesu*), Judgement, para 560.
[123] ICTR 2 September 1998, Case No. ICTR-96-4-T, (*Prosecutor/Akayesu*), Judgement, para 562.
[124] ICTR 3 December 2003, Case No. ICTR-99-52-T (*Prosecutor/Nahimana et al.*), Judgement.
[125] ICTR 3 December 2003, Case No. ICTR-99-52-T (*Prosecutor/Nahimana et al.*), Judgement, para 979.
[126] Gordon 2017, p. 154.
[127] Including ECtHR 23 September 1994, ECLI:CE:ECHR:1994:0923JUD001589089 (*Jersild/Denmark*) and ECtHR 8 July 1999, ECLI:CE:ECHR:1999:0708JUD002668295 (*Sürek/Turkey (No. 1)*). See ICTR 3 December 2003, Case No. ICTR-99-52-T (*Prosecutor/Nahimana et al.*), Judgement, paras 981-999.
[128] ICTR 3 December 2003, Case No. ICTR-99-52-T (*Prosecutor/Nahimana et al.*), Judgement, para 1000.
[129] ICTR 3 December 2003, Case No. ICTR-99-52-T (*Prosecutor/Nahimana et al.*), Judgement, para 1001.
[130] ICTR 3 December 2003, Case No. ICTR-99-52-T (*Prosecutor/Nahimana et al.*), Judgement, para 1002.
[131] ICTR 3 December 2003, Case No. ICTR-99-52-T (*Prosecutor/Nahimana et al.*), Judgement, para 1004.
[132] ICTR 3 December 2003, Case No. ICTR-99-52-T (*Prosecutor/Nahimana et al.*), Judgement, para 1022.

defendants did not distance themselves from the message of ethnic hatred, rather they purveyed the message.[133]

The prohibition of crimes against humanity (CAH), a *jus cogens*, has not been codified into a separate treaty.[134] However, CAH is included in the ICTY Statute, ICTR Statute and the Rome Statute.[135] Atrocity speech can qualify as *actus reus* of persecution as a CAH offence.[136]

The requirements for persecution were defined by the ICTY in *Tadic*:

> the occurrence of a persecutory act or omission and a discriminatory basis for that act or omission on one of the listed grounds, specifically race, religion or politics. […] [T]he persecutory act must be intended to cause, and result in, an infringement on an individual's enjoyment of a basic or fundamental right.[137]

In *Ruggiu*, the ICTR determines that direct and public radio broadcasts singling out and attacking an ethnic group on discriminatory grounds can deprive this group of the fundamental rights to life, liberty and basic humanity. The deprivation of these rights can have as aim the death and removal of those persons from the society in which they live, or even from humanity itself.[138] The element of *mens rea* is also required; the accused must know that his act(s) is part of a widespread or systematic attack on a civilian population and pursuant to some kind of policy or plan.[139]

As evidenced in the tribunals' case law, if left unchecked atrocity speech can result in atrocity crimes such as genocide and CAH. For this reason atrocity speech must be prohibited.

2.7. Preliminary Conclusion

It is imperative to curtail certain forms of expression. There is a fine line between expression that may offend, shock or disturb, and expression that may be considered hate speech. The distinction between general hate speech and atrocity speech is also ambiguous. It depends on content and context, intent and imminence of the danger. It also depends whether the speaker targets an individual or an entire racial or ethnic group. The distinction between expression that should not be restricted, may be restricted and must be prohibited is not clear-cut and decided on a case by case basis. This also makes atrocity speech difficult to prevent.

The case law discussed in this chapter all stemmed from the period before the SMC platforms, a form of mass media, came into existence. Chapter three will focus on SMC platforms, which have amplified the scale and reach of expression, whether permitted, restricted or prohibited.

[133] ICTR 3 December 2003, Case No. ICTR-99-52-T (*Prosecutor/Nahimana et al.*), Judgement, para 1024.
[134] Gordon 2017, p. 129.
[135] Article 5(h) ICTY Statute, article 3(h) ICTR Statute and article 7(1)(h) Rome Statute.
[136] Gordon 2017, p. 166.
[137] ICTY 7 May 1997, Case No. IT-94-1-T (*Prosecutor/Tadic*), Judgement, para 715.
[138] ICTR 1 June 2000, Case No. ICTR-97-32-I (*Prosecutor/Ruggiu*), Judgement, para 22.
[139] ICTR 1 June 2000, Case No. ICTR-97-32-I (*Prosecutor/Ruggiu*), Judgement, para 20.

3. Regulating Social Media Platforms

3.1. Introduction

The ECtHR observed that the internet has 'become one of the principal means by which individuals exercise their right to freedom of expression and information, providing as it does essential tools for participation in activities and discussions concerning political issues and issues of general interest.'[140]

Cyberspace is 'the online world of computer networks and especially the internet.'[141] Internet intermediaries are actors that provide broad internet services, such as hosting, servers, and social media platforms.[142] The introduction of this thesis featured three US based multinationals (SMCs) offering social media platform services: Meta, Alphabet and Twitter. Users, such as civilians, commercial entities, news outlets and governments, propagate ideas, information and news on these platforms. The business model of these SMCs is to have as many users as possible, as user engagement enables the SMCs to collect vast amounts of personal data on these users. This data is used for advertising, which is the main source of the SMCs' revenue.[143] The SMCs have developed advanced algorithms to keep users engaged.[144]

The SMC business model has impact on human rights, most notably on privacy[145] and the freedom of expression. This chapter will first set out the general rules of State jurisdiction and State responsibility with respect to IHRL, and whether this applies extraterritorially in cyberspace. Then the chapter will discuss State responsibility with respect to TNCs, including the UNGPs,[146] and whether State duties of due diligence can be exercised extraterritorially regarding SMCs. Thereafter, the chapter will explore the First Amendment of the US constitution as well as 47 U.S. Code, Section 230 and the impact thereof on the SMCs as internet intermediaries. Finally, the chapter will consider the self-regulation of SMCs on the basis of the UNGPs.

3.2. State Jurisdiction in Cyberspace

Article 2(1) of the UN Charter contains the principle of sovereign equality of States. Each State is an equal member of the international community, with equal rights and duties. Sovereignty concerns a State's exclusive competence in regard to its own territory to wield legal authority over its subjects, independent from the commands of another State.[147]

Jurisdiction is a consequence of State sovereignty and can be either prescriptive, executive or adjudicatory. Prescriptive jurisdiction relates to the competence of a State to pass and enact laws. Executive jurisdiction concerns the State's authority to enforce those laws. Adjudicatory jurisdiction refers to the competence of the State's national courts.[148] A State's jurisdiction is protected from intervention by other States through Article 2(7) of the UN Charter, although this protection is not absolute.[149]

---

[140] ECtHR 18 December 2012, ECLI:CE:ECHR:2012:1218JUD000311110 (*Yildirim/Turkey*), para 54.
[141] As defined by the Merriam-Webster dictionary.
[142] Social media platforms are internet websites or apps that publish user content. See: Klonick, *Harvard Law Review* 2018, p. 1604.
[143] Amnesty.org/en/latest/campaigns/2022/02/what-is-big-techs-surveillance-based-business-model.
[144] Oremus et al., *The Washington Post* 26 October 2021.
[145] The right to privacy is outside of the scope of this thesis.
[146] United Nations Guiding Principles on Business and Human Rights (2011, HR/PUB/11/04) (UNGPs).
[147] Arbitral Award 4 April 1928, *Island of Palmas Case* (*Netherlands/United States of America*), p. 838. See also: Hernández 2019, p. 22.
[148] Hernández 2019, p. 194-198.
[149] Article 2(7) of the UN Charter: 'Nothing contained in the present Charter shall authorize the United Nations to intervene in matters which are essentially within the domestic jurisdiction of any state or shall require the Members to submit such matters to settlement under the present Charter; but this principle shall not prejudice the application of enforcement measures under Chapter Vll.'

Territory includes a State's land, internal waters, territorial sea and airspace.[150] Jurisdiction is not confined by territory, as States may claim jurisdiction over acts or breaches outside of their own borders. By doing so, a State would extend its legal power beyond its territory. This would be a claim of extraterritorial jurisdiction, where international law applies.[151]

The principle of State sovereignty applies in cyberspace,[152] as components of the cyber infrastructure are located on a State's territory,[153] and the persons and entities that conduct cyber activities are subject to the jurisdiction of one or more States.[154] This is confirmed in the UN GGE 2021 report,[155] which also declares that international law is applicable in cyberspace.[156]

3.3. State Responsibility and the Extraterritorial Application of IHRL in Cyberspace

It is generally recognised that human rights are protected in cyberspace.[157] State Parties to the ICCPR and the ECHR have the obligation to ensure that individuals can enjoy human rights online. The Tallinn Manual 2.0 finds that both treaty and customary international human rights law apply to cyber activities, including the limitations permitted under IHRL.[158]

While a State's human rights jurisdiction is presumed within that State's territory, extraterritorial human rights jurisdiction is exceptional and has to be established.[159] Article 2(1) ICCPR requires each State Party to respect and ensure the rights of the ICCPR to all individuals within its territory and subject to its jurisdiction. In its *Wall Advisory Opinion*, the ICJ considered that

> while the jurisdiction of States is primarily territorial, it may sometimes be exercised outside of the national territory. Considering the object and purpose of the ICCPR, it would seem natural that, even when such is the case, State parties should be bound to comply with its provisions.

The ICJ concludes that the ICCPR is 'applicable in respect of acts done by a State in the exercise of its jurisdiction outside its own territory.'[160]

The HRC looks at State jurisdiction from the perspective of the relationship between the individual and the State in relation to a violation of any of the rights in the ICCPR, wherever they occurred, rather than the State's relationship to the territory where the violation occurred.[161] The

---

[150] *Report of the International Law Commission 2006,* p. 518.
[151] Hernández 2019, p. 196.
[152] Tallinn Manual 2.0 rule 1, p. 11. The Tallinn manuals are initiatives of the NATO Cooperative Cyber Defence Centre of Excellence. In the 2013 Tallinn Manual, an independent International Group of Experts determined rules in which legal norms are applied to cyber warfare. In the 2017 Tallinn Manual 2.0, a new group of independent International Experts adopted additional rules to include public international law governing cyber operations during peacetime. The Tallinn Manual 2.0 supersedes the original Tallinn Manual. The Tallinn Manual 2.0 is intended as a reflection of the law as it exists, or *lex lata*. See also Tallinn Manual 2.0, p. 1-3.
[153] Tallinn Manual 2.0 rule 1.1, p. 11.
[154] Tallinn Manual 2.0 rule 1.5, p. 12.
[155] UN Report of the Group of Governmental Experts on Advancing Responsible State Behaviour in Cyberspace in the Context of International Security (2021, UN Doc. A/76/135) (UN GGE report 2021), para 71(b). In the UN GGE report 2021 reference is consistently made to information and communication technologies, or ICTs. However, the title of the report refers to cyberspace instead of ICTs.
[156] UN GGE report 2021, para 69.
[157] See, among others, UN Human Rights Council (2016, A/HRC/32/L.20), para 1; EU Human Rights Guidelines on Freedom of Expression Online and Offline (2014), para 6; UN Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security (2015, A/70/174), para 13(e); and UN GGE report 2021, para 36.
[158] Tallinn Manual 2.0, p. 179, para 1.
[159] Besson, *ESIL Reflections* 28 April 2020, p. 3. See also:
ECtHR 12 December 2001, ECLI:CE:ECHR:2001:1212DEC005220799 (*Banković/Belgium*), para 71.
[160] ICJ 6 July 2004, *Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory* (*Advisory Opinion*), para 109-111.
[161] HRC 29 July 1981, Communication No. R.12/52, UN Doc. Supp. No. 40 (*Lopez Burgos/Uruguay*), para 12.2.

HRC indicates that ICCPR rights are 'not only limited to citizens of State Parties, but must also be available to all individuals, regardless of nationality or statelessness […]' and continues to determine that

> [t]his principle also applies to those within the power or effective control of the forces of a State Party acting outside its territory, regardless of the circumstances in which such power or effective control was obtained.[162]

"Effective control" may refer to the "effective control test," which is derived from the ICJ *Nicaragua* judgement,[163] and is a test by which the conduct of a (private) person or group of persons could be attributed to a State. This is the case if that conduct can be considered to be carried out under the instructions of, or under the direction or control of that State.[164] "Effective control" in this context would be exercised over the person or private entity causing the human rights violation, and not the affected person.[165] "Power" refers to a State's jurisdiction when it exercises authority and control over individuals.[166]

The ECHR does not express a territorial limit regarding its application.[167] In its *Banković* judgement, the ECtHR determined that a State's jurisdictional competence is primarily territorial,[168] and that the State must have effective control over a territory. In an earlier judgement, that of *Loizidou*, the ECtHR held that

> Bearing in mind the object and purpose of the Convention, the responsibility of a Contracting Party may also arise when as a consequence of military action—whether lawful or unlawful—it exercises effective control of an area outside its national territory. The obligation to secure, in such an area, the rights and freedoms set out in the Convention, derives from the fact of such control whether it be exercised directly, through its armed forces, or through a subordinate local administration.[169]

The inclusion of "an area" to the requirement of "effective control" adds a territorial aspect to the State's jurisdiction. Whereas the ICJ refers to state control over actors to determine whether their acts can be attributed to that State, the ECtHR refers to state control over territory, for instance through military occupation,[170] to establish a State's jurisdiction. As such, the two concepts of "effective control" are distinct.[171]

In *Issa*, the ECtHR found that a State may also be held accountable for human right violations

> [of] persons who are in the territory of another State but who are found to be under the former State's *authority and control* through its agents operating – whether lawfully or unlawfully - in the latter State.[172]

"Authority and control" in this context is similar to the ICJ's "effective control test."[173] In its *Al-Skeini* judgement, the ECtHR refers to *Banković*, confirms that a State's extraterritorial jurisdiction is exceptional,[174] and adds that

---

[162] UN Human Rights Committee, General Comment No. 31 (2004, CCPR/C/21/Rev.1/Add.13) (GC 31), para 10.
[163] ICJ 27 June 1986, *Case Concerning Military and Paramilitary Activities In and Against Nicaragua (Nicaragua/United States of America),* para 114-116.
[164] Article 8 ARSIWA.
[165] Besson, *ESIL Reflections* 28 April 2020, p. 4.
[166] Milanovic 2011, p. 173.
[167] Joseph & Dipnall 2018, p. 121. See also article 1 ECHR: 'the High Contracting Parties shall secure to everyone within their jurisdiction the rights and freedoms defined […] [the] Convention.'
[168] ECtHR 12 December 2001, ECLI:CE:ECHR:2001:1212DEC005220799 (*Banković/Belgium*), para 59.
[169] ECtHR 18 December 1996, ECLI:CE:ECHR:1996:1218JUD001531889 (*Loizidou/Turkey*), para 52.
[170] ECtHR 12 December 2001, ECLI:CE:ECHR:2001:1212DEC005220799 (*Banković/Belgium*), para 80.
[171] Milanovic 2011, p. 135-137.
[172] ECtHR 16 November 2004, ECLI:CE:ECHR:2004:1116JUD003182196 (*Issa/Turkey*), para 71.
[173] Milanovic, *EJIL* 2012, p. 122.
[174] ECtHR 7 July 2011, ECLI:CE:ECHR:2011:0707JUD005572107 (*Al-Skeini/United Kingdom*), para 149-150.

> The Court does not consider that jurisdiction in the above cases arose solely from the control exercised by the Contracting State over the buildings, aircraft or ship in which the individuals were held. What is decisive in such cases is the exercise of physical power and control over the person in question.[175]

It is the "physical power and control" that complicates matters for the extraterritorial application of IHRL in cyberspace. The Tallinn Manual 2.0 recognises that IHRL is applicable in cyberspace within a State's territory, but contends the extraterritorial applicability. The argument is that, based on the *Al-Skeini* judgement, a State should have physical control over the territory or the individual. Exercising power or effective control by virtual means is considered contrary to state practice and *opinio juris*,[176] when applying customary IHRL. Another complication is that these doctrines apply to State jurisdiction and responsibility when exerting military or administrative control on another State's territory, it rarely extends to control over extraterritorial human rights violations by TNCs.[177]

3.4. State Responsibility over TNCs and the UNGPs

The HRC asserts that States have the responsibility to ensure the protection of human rights against acts committed by private entities,[178] such as transnational corporations (TNCs). If a State fails to ensure the protection of these rights, by not taking appropriate measures or not exercising due diligence to prevent, punish, investigate or redress the harm caused by such acts, this may give rise to a violation of the ICCPR by that State.[179] The ECtHR has also established a State's duty to take reasonable measures to prevent human rights violations by private entities.[180] For the due diligence obligation to arise, the State should have the capacity to effectively influence the source of harm.[181] This means that the State should be able to exercise control over the TNC causing harm, effective control over the individual right-holder is not necessary.[182]

TNCs pose challenges through their structure, as they operate in multiple jurisdictions. If the host State, being the State where the TNC is conducting its business, has weak governance with respect to human rights, the TNC will not be held accountable for human rights abuses locally. In that case, the home State, being the State where the TNC is domiciled,[183] may have obligations under international law to exercise extraterritorial jurisdiction.[184] This may particularly be the case when violations of human rights that are serious in nature are at stake.[185] However, unless it concerns serious violations of peremptory norms, TNCs do not have duties under IHRL.[186]

---

[175] ECtHR 7 July 2011, ECLI:CE:ECHR:2011:0707JUD005572107 (*Al-Skeini/United Kingdom*), para 136.
[176] Tallinn Manual 2.0 rule 34.10, p. 185, with reference to ECtHR 7 July 2011, ECLI:CE:ECHR:2011:0707JUD005572107 (*Al-Skeini/United Kingdom*), para 136.
[177] McCorquodale & Simons, *The Modern Law Review* 2007, p. 609. The authors refer to the "effective control" doctrine as derived from the ICJ's *Nicaragua* judgment.
[178] UN Human Rights Committee, General Comment No. 31 (2004, CCPR/C/21/Rev.1/Add.13), para 8.
[179] UN Human Rights Committee, General Comment No. 31 (2004, CCPR/C/21/Rev.1/Add.13), para 8.
[180] Joseph & Dipnall 2018, p. 115, See also: ECtHR 9 December 1994, ECLI:CE:ECHR:1994:1209JUD001679890 (*López Ostra/Spain*), para 55-58 and ECtHR 28 October 1998, ECLI:CE:ECHR:1998:1028JUD002345294 (*Osman/United Kingdom*), para 115.
[181] ICJ 26 February 2007, *Case Concerning Application of The Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia and Herzegovina/Serbia and Montenegro)*, para 430.
[182] Besson, *ESIL Reflections* 28 April 2020, p. 5.
[183] Domicile is the State in which a business has been incorporated or is registered as headquarter, confirmed in ICJ 5 February 1970, *Case Concerning the Barcelona Traction, Light and Power Company, Limited (Belgium/Spain)* (Second Phase), para 70. See for definitions of host State and home State, Ruggie 2013, p. 89.
[184] Bernaz, *Journal of Business Ethics* 2013, p. 494.
[185] HRC 7 December 2017, Communication No. 2285/2013, UN Doc. CCPR/C/120/D/2285/2013 (*Issa Yassin/Canada*), para 6.5.
[186] Ruggie 2013, p. 66.

States recognise their duty under IHRL to prevent violations by private entities, however, actual State practice shows that measures are not sufficiently implemented.[187] Another challenge is that businesses resist international regulations with respect to corporate responsibility for human rights, as they consider these responsibilities to lie with the State,[188] although TNCs often adopt corporate social responsibility guidelines internally.[189] The UNGPs were developed to establish a global normative platform and policy guidance on the conduct of TNCs.[190] The UNGPs were endorsed by the HRC in 2011, and are now considered the most authoritative international standard in business and human rights.[191] The UNGPs are based on three pillars;[192] a State's duty to protect human rights,[193] corporate responsibility to respect human rights,[194] and access to effective remedy.[195]

It is important to note that responsibility in the context of "corporate responsibility" differs from the international law concept of State responsibility. The UNGPs require the TNC to respect human rights, but indicate that this is distinct from issues of legal liability and enforcement, which remain defined largely by national law provisions in relevant jurisdictions.[196] The UNGPs are non-binding principles that lack an enforceability mechanism in this respect.[197]

Under the UNGPs, there is no formalised process to assess if corporate accountability standards are being met. Accountability can be seen as both corporate liability, which implies a legal obligation or duty, as well as the non-legal implications of being answerable for the consequence of one's actions such as loss of reputation, denial of access to foreign markets and shareholder dissent.[198] States can enhance the TNC's accountability regarding IHRL by taking measures. State measures could be domestic with extraterritorial implications, such as a TNC headquarter reporting on the company's human rights policies and impacts.[199] However, home States tend to be more concerned about the competitive position of the TNC, and TNCs remain strongly opposed to extraterritorial jurisdiction. Also, host States may resist interference by other States in their domestic affairs.[200]

The UNGPs call on home States to ensure that TNCs operating in conflict-affected areas are not involved with gross human rights abuses, as the host State may be unable to protect human rights adequately.[201] As will be discussed in chapter 4, a host State can even be the cause of such abuses, using an SMC platform in the dissemination of atrocity speech, leading to atrocity crimes within its territory. To further lay the foundation for the case study of chapter 4, the next sections will focus on the (self-)regulation of SMCs incorporated under US law.

[187] Ruggie 2013, p. 102.
[188] Ruggie 2013, p. 66.
[189] Ruggie 2013 p. 103.
[190] Ruggie 2013, p. 104.
[191] Bernaz, *Human Rights Review* 2021, p. 49.
[192] UNGPs, p. iv.
[193] UNGPs, p. 3.
[194] UNGPs, p. 13.
[195] UNGPs, p. 27.
[196] UNGPs, p. 14.
[197] Bernaz, *Human Rights Review* 2021, p. 52.
[198] Bernaz, *Journal of Business Ethics* 2013, p. 494.
[199] UN Report of the Special Representative of the Secretary-General (John Ruggie) on the issue of human rights and transnational corporations and other business enterprises, UN Doc. A/HRC/14/27 (9 April 2010), para 49.
[200] Ruggie 2013, p. 89.
[201] UNGPs, p. 10, principle 7.

3.5. Applicable US Law to SMCs

Three very large social media platforms, Facebook, YouTube and Twitter, are owned by SMCs that are domiciled in the US and incorporated under US law. This is relevant for two reasons. The first reason is that the US has a different perspective from international law towards the freedom of expression.[202] The second reason is that the US considers the SMCs as internet intermediaries that fall under Section 230 of the U.S. Communications Decency Act (Section 230).[203]

The US made reservations to article 20(2) ICCPR and article 4 ICERD, and has not incorporated any anti-hate speech clauses into domestic law.[204] The US perceives these provisions to be contrary to the First Amendment to the US Constitution, which states

> Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances.

The US approaches freedom of expression as the "marketplace of ideas." This theory holds that the "truth" will most likely emerge when there is an open and unregulated market where all opinions may be freely expressed.[205] Proponents of this philosophy believe that if government is permitted to suppress speech, it will do so in vague and overbroad terms and will obtain unfettered censorial power.[206] This is perceived as a cure worse than the disease.[207] Potential hate speech laws could be used to suppress pro-civil rights speech as this would be considered speech that may trigger violence.[208] Such laws may then undermine universal principles of liberty, equality, and democracy.[209] Instead of restricting or prohibiting hate speech, counterspeech should be used.[210]

As a matter of principle, the freedom of expression should not be restricted under the First Amendment. Even speech conveying hateful or discriminatory ideas [211] that cause discomfort or mental pain is permitted. There are no limits to hate speech, unless it, in context, poses a clear and present danger. Such speech prohibitions under the First Amendment are only possible under very strict conditions and have developed in US case law. In *Schenck*,[212] Justice Holmes developed the "emergency test:" it must be clear that the speech will directly, demonstrably, and imminently cause certain specific, objectively ascertainable serious harms that cannot be averted by non-censorial measures, such as counterspeech.[213] Examples of punishable speech are, among others, "true threats," where the speaker communicates a serious intent to commit an act of unlawful violence to an individual or group of individuals; and "punishable incitement" which refers to a speaker intentionally inciting imminent violent or otherwise illegal conduct that is likely to occur immediately.[214]

Social media platforms provide the purest application of the "marketplace of ideas," as they provide an almost limitless forum for freedom of expression.[215] However, the First Amendment constrains governmental actors and protects private actors.[216] SMCs are considered private actors

[202] Gordon 2017, p. 76.
[203] 47 U.S.C. § 230. See Klonick, *Harvard Law Review* 2018, p. 1602.
[204] Gordon 2017, p. 76.
[205] Maggiore, *Michigan Telecommunications and Technology Law Review* 2012, p. 632.
[206] Strossen 2018, p. 6.
[207] Strossen 2018, p. 14.
[208] Strossen 2018, p. 17.
[209] Strossen 2018, p. 23.
[210] Strossen 2018, p. xxii.
[211] Strossen 2018, p. 7-8.
[212] *Schenck v. United States*, 249 U.S. 47 (1919).
[213] Strossen 2018, p. xx.
[214] Strossen 2018, p. 60-62.
[215] Maggiore, *Michigan Telecommunications and Technology Law Review* 2012, p. 642.
[216] *Manhattan Cmty. Access Corp. v. Halleck*, 139 S. Ct. 1921, 1926 (2019).

that are entitled to implement restrictions on their platforms,[217] for instance with respect to hate speech. Some states, such as Texas and Florida, are attempting to pass laws in which they regard SMC platforms as common carriers, arguing that the platforms have become as important for conveying public opinion as public utilities are for supporting modern society.[218] These laws would prohibit SMCs to restrict any speech, effectively imposing constitutional constraints on the SMCs. For now these laws are stopped by US courts.[219]

Section 230 protects internet intermediaries, or "interactive computer services,"[220] from liability from user-generated content.[221] This law was passed following a US court case suggesting internet intermediaries could be held liable for user content.[222] The rationale was explained in *Zeran*:[223] 1) to encourage the internet intermediary and its users to self-police for obscenity and other offensive material and 2) and to encourage the unfettered and unregulated development of free speech online.[224] In the context of the First Amendment, this approach is understandable.[225] In the years following *Zeran*, US court decisions have established Section 230 as a broad shield for internet intermediaries.[226] However, Section 230 pertains to the liability of internet intermediaries as if they were the publisher of defamatory content.[227] It does not shield the intermediary from legislative obligations directly applicable to the platform itself, such as user disclosure, transparency regarding hate speech or the content moderation algorithms.[228]

From a domestic point of view, the US government will not impose restrictions on hate speech on social media platforms, as this is contrary to the First Amendment. This also means that the US will not exercise extraterritorial jurisdiction with respect to hate speech on the platforms. An important case evidencing this approach is the *Yahoo!* case.[229] This case concerned the sale of Nazi memorabilia on a US Yahoo! auction webpage, which was also accessible to users in France. Displaying Nazi items for sale is considered a violation of the French Penal Code. The French court established jurisdiction based on the negative effect the act produced on French territory and the fact that the users could access the US auction site through French-language banners placed on Yahoo.fr.[230] Yahoo! Inc claimed that this request was in violation of the First Amendment and Yahoo France denied liability as the auction was not hosted on Yahoo.fr. Société Yahoo! France was ordered to warn users that they may breach French law if they followed links to sites operated by Yahoo! Inc. Yahoo! Inc. was ordered to take all appropriate measures to deter and prevent access to auctions of Nazi memorabilia on its site by users in France. Yahoo! Inc. also took the case to US court, with respect to its First Amendment rights. The US court considered the extraterritorial application of the First Amendment, whether it requires unrestricted access by users in France, even if it facilitates a violation of French criminal law.[231] The US court concluded that the extent or the

[217] Strossen 2018, p. 56.
[218] For instance: Florida SB 7072: Social Media Platforms, sections 1(5) and 1(6).
[219] Barnes & Zakrzewski, *The Washington Post* 31 May 2022.
[220] 47 U.S.C. § 230(f)(2): 'The term "interactive computer service" means any information service, system, or access software provider that provides or enables computer access by multiple users to a computer server, including specifically a service or system that provides access to the Internet and such systems operated or services offered by libraries or educational institutions.'
[221] 47 U.S.C. § 230(c)(2).
[222] *Stratton Oakmont, Inc. v. Prodigy Services Co.*, N.Y. Sup. Ct. 1794 (1995).
[223] *Zeran v. America Online, Inc.*, 129 F.3d 327 (1997).
[224] Klonick, *Harvard Law Review* 2018, p. 1607-1608.
[225] Oster 2017, p. 64.
[226] Hwang 2020, p. 261.
[227] Hwang 2020, p. 262.
[228] Hwang 2020, p. 268.
[229] *Ligue contre le racisme et l'antisémitisme et Union des étudiants juifs de France c. Yahoo! Inc. et Société Yahoo! France* (*LICRA c. Yahoo!*) (2000); *Yahoo! Inc. v. La Ligue Contre Le Racisme et L'Antisemitisme*, 433 F.3d 1199 (2006).
[230] Solmone 2020, p. 48.
[231] *Yahoo! Inc. v. La Ligue Contre Le Racisme et L'Antisemitisme*, 433 F.3d 1199 (2006).

very existence of such an extraterritorial right under the First Amendment is uncertain and that First Amendment issues arising out of international Internet use are 'new, important and difficult.'[232]

3.6. The Liability and Regulation of Social Media Platforms under International Law

SMCs are also considered internet intermediaries under international law.[233] Unlike Section 230, the SMCs are not completely immune from liability. A balance must be struck between the SMCs freedom of expression, including the right to impart information, and the rights of the general public and individuals to this information versus the interests of persons whose rights have been violated and the public order interest of States.[234] The ECtHR deems article 10 ECHR applicable to "everyone" and states that

> No distinction is made in it according to the nature of the aim pursued or the role played by natural or legal persons in the exercise of that freedom. It applies not only to the content of information but also to the means of dissemination, since any restriction imposed on the means necessarily interferes with the right to receive and impart information. Admittedly, publishers do not necessarily associate themselves with the opinions expressed in the works they publish. However, by providing authors with a medium they participate in the exercise of the freedom of expression, just as they are vicariously subject to the "duties and responsibilities" which authors take on when they disseminate their opinions to the public.[235]

However, as SMCs are not aware of the content disseminated through their platforms, they may lack accountability for that content.[236]

Establishing State jurisdiction on social media platforms can be equally challenging. It is often not possible to identify the perpetrator of an unlawful act and the territory in which the act originated.[237] This also poses a challenge with respect to extraterritorial jurisdiction. According to the Joint Declaration on Freedom of Expression and the Internet

> jurisdiction in legal cases relating to Internet content should be restricted to States to which those cases have a real and substantial connection, normally because the author is established there, the content is uploaded there and/or the content is specifically directed at that State.[238]

Nevertheless, even if State jurisdiction could be established, the measures imposed by a State can impact the freedom of expression of users that are not under the jurisdiction of that State. As such that State could violate the sovereignty of another State,[239] or even exercise universal jurisdiction over a matter that cannot be considered an international crime.[240]

---

[232] *Yahoo! Inc. v. La Ligue Contre Le Racisme et L'Antisemitisme*, 433 F.3d 1199 (2006).
[233] For instance: Recommendation CM/Rec(2018)2 of the Committee of Ministers of the Council of Europe (7 March 2018), *on the roles and responsibilities of internet intermediaries*, para 4.
[234] Oster 2017, p. 64.
[235] ECtHR 28 September 1999, ECLI:CE:ECHR:1999:0928JUD002247993 (*Özturk/Turkey*), para 49.
[236] Oster 2017, p. 66.
[237] Solmone 2020, p. 25.
[238] Joint Declaration on Freedom of Expression and the Internet The United Nations (UN) Special Rapporteur on Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe (OSCE) Representative on Freedom of the Media, the Organization of American States (OAS) Special Rapporteur on Freedom of Expression and the African Commission on Human and Peoples' Rights (ACHPR) Special Rapporteur on Freedom of Expression and Access to Information (1 June 2011).
[239] Solmone 2020, p. 87.
[240] Solmone 2020, p. 57.

The UNGPs are applicable to SMCs[241] and as such the SMCs have developed codes of conduct that refer to the UNGPs.[242] They have also developed hate speech policies[243] that are directly applicable to users. For instance, the Facebook Community Standard defines hate speech as a direct attack against people on the basis of protected characteristics.[244] Facebook declares to be committed to removing hate speech, and does so based on context of the culture and language as well as the intent of the user.[245]

Managing user expression on social media platform is called content moderation. As SMCs are active in multiple jurisdictions, this requires bespoke content moderation as the SMCs can be constrained by municipal law. States may wish to exercise control over the SMC with respect to content moderation, and influence which forms of expression can be removed or stay on the platform, or even consider the SMC itself as source of harm. Another question is whether the facilitation of atrocity speech could be consider an infringement on the human rights of others.

Establishing liability for and State jurisdiction on SMC platforms pose quite a challenge. As a consequence, SMCs are left to self-regulate their platforms in accordance with their codes of conduct. This results in a huge global impact of these codes of conduct on the exercise of the freedom of expression.

3.7. Preliminary Conclusion

The US rarely restricts hate speech in accordance with its First Amendment, and only prohibits speech that exhibits a serious intent to cause immediate violent conduct. The US deviates in this respect from IHRL. Additionally, as SMCs have constitutional rights and their liability as internet intermediaries is limited under US law, the US does not regulate expression on SMC platforms. As a result, the SMCs self-regulate their platforms under US law.

SMCs also enjoy limited liability under international law, and are in principle not responsible for content posted by users. State jurisdiction, especially extraterritorially, is difficult to establish in cyberspace, even if international law does apply. Due to lack of binding treaties or customary international law for the regulation of TNCs, the SMCs are also left to self-regulate their platforms under international law. Although the UNGPs serve as a basis for their codes of conduct, SMCs are not duty-bearers under IHRL. Yet, they restrict and even prohibit expression on their platforms.

The result is a lack of governmental oversight on SMC platforms, even though SMCs operate the platforms for commercial gain. Chapter 4 consists of a case study evidencing how this combination can result in atrocious consequences.

[241] UN Report of the Special Rapporteur (David Kaye) on the promotion and protection of the right to freedom of opinion and expression, UN Doc. A/HRC/38/35 (6 April 2018), para 10.

[242] Facebook: About.fb.com/wp-content/uploads/2021/03/Facebooks-Corporate-Human-Rights-Policy.pdf; Twitter: Help.twitter.com/en/rules-and-policies/defending-and-respecting-our-users-voice; YouTube: About.google/human-rights.

[243] YouTube: Support.google.com/youtube/answer/2801939?hl=en; Twitter: Help.twitter.com/en/rules-and-policies/hateful-conduct-policy; Facebook: Transparency.fb.com/en-gb/policies/community-standards/hate-speech

[244] The protected characteristics are race, ethnicity, national origin, disability, religious affiliation, caste, sexual orientation, sex, gender identity and serious disease. Attack is defined as violent or dehumanising speech, harmful stereotypes, statements of inferiority, expressions of contempt, disgust or dismissal, cursing and calls for exclusion or segregation. See: Transparency.fb.com/en-gb/policies/community-standards/hate-speech (last consulted on 29 may 2022)

[245] About.fb.com/news/2017/06/hard-questions-hate-speech

4. Case Study: Atrocity Speech on Facebook in Myanmar

4.1. Introduction

The SMC platforms are used globally by billions of users, which has given rise to a number of issues, such as disinformation campaigns, foreign interference in elections, privacy issues and hate speech.[246] Notable cases are Russian bots interfering in the 2016 US elections,[247] Covid-19 related disinformation,[248] but also the US Capitol attack on 6 January 2021.[249] These cases have in common that social media platforms were used to reach a large number of users to manipulate their opinion on important matters. The US Capitol attack resulted in Facebook and Twitter banning then President Trump from their platforms.[250]

A particularly notable case is the role of Meta in Myanmar. The Facebook platform played a pivotal role in the dissemination of hate speech and atrocity speech against Rohingya, an ethnic group that became victim of atrocity crimes. This chapter will determine whether the US, as home State of Meta, can be considered to have had an obligation under the CPPCG to prevent atrocity speech from occurring on Facebook. The chapter will also explore whether Meta can be considered complicit in the serious violations of peremptory norms that occurred in Myanmar.

4.2. The Case of Myanmar

Myanmar has a population of 54.80 million[251] and is inhabited by a large number of groups with various ethnic, cultural, linguistic and religious backgrounds.[252] The Rohingya are an Muslim ethnic minority group in Myanmar that have faced systemic oppression and discrimination for many decades.[253] Anti-Muslim rhetoric includes presenting Islam as a serious threat to Buddhism.[254] The Ministry of Labour, Immigration and Population's website features the motto 'The earth will not swallow a race to extinction but another race will.'[255] The military reportedly has received specific training regarding the 'fear of extinction of the race.'[256]

In August 2017, the "Tatmadaw," Myanmar's military, launched a disproportionate attack on the Rohingya, after a Rohingya rebel group attacked several military posts in Rakhine State. Several hundred Rohingya villages were destroyed and more than 700,000 Rohingya fled to Bangladesh.[257] The UN Human Rights Council established the Independent International Fact-Finding Mission on Myanmar (IIFFMM).[258] The IIFFMM submitted a report in September 2018, including a document[259] containing detailed factual and legal analysis of the serious human rights violations and abuses, as well as the serious violations of international humanitarian law that occurred in Myanmar.[260]

The IIFFMM concludes that gross human rights violations and abuses were committed that are shocking for their horrifying nature and ubiquity, and that undoubtedly amount to the gravest

---

[246] Persily & Tucker 2020, p. 1.
[247] Woolley 2020, p. 89-92, bots in this context are automated imposter accounts, most notably on Twitter, that generate content, mimic real users and manipulate public opinion.
[248] Persily & Tucker 2020, p. 2.
[249] Britannica.com/event/United-States-Capitol-attack-of-2021.
[250] Alba, Koeze & Silver, *The New York Times* 7 June 2021.
[251] Data.un.org/en/iso/mm.html.
[252] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 84.
[253] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 102. See also UN Human Rights Council (2018, A/HRC/39/64), para 20-23.
[254] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 697.
[255] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 698.
[256] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 716.
[257] Exhibitions.ushmm.org/burmas-path-to-genocide/timeline.
[258] UN Human Rights Council (2017, A/HRC/RES/34/22).
[259] *Report Independent International Fact-Finding Mission on Myanmar* 2018para 2
[260] UN Human Rights Council (2018, A/HRC/39/64), p. 1.

crimes under international law.[261] The crimes listed by the IIFFMM are genocide, crimes against humanity, and war crimes.[262] The IIFFMM indicates that the Tatmadaw was the main perpetrator of serious human rights violations and crimes under international law. Other perpetrators were the police, local authorities, militias, militant "civilian" groups, politicians and monks, who participated or assisted in the violations.[263]

Virulent hate speech against the Rohingya is widely prevalent in Myanmar.[264] As it concerns racial slurs and dehumanising language, this hate speech is deemed prohibited by the IIFFMM in accordance with article 20(2) ICCPR.[265] The hate campaigns have been led by radical Buddhist monk organisations such as the MaBaTha, officials and influential figures and Senior-General Min Aung Hlaing, the commander-in-chief of the Tatmadaw.[266] Social media, and in particular Facebook have played an important role in the dissemination of hate speech.[267] Due to a democratic reform process and the liberation of the telecommunications industry, the Myanmar population gained access to mobile telephony and the Internet as of 2011.[268] Facebook enjoys a more than 90% share among social media platforms in Myanmar, especially after Facebook launched a specific Myanmar version in 2015.[269] Facebook also launched an illustrated Myanmar-language version of its Community Standards.[270] In 2016, Facebook rolled out special products together with Myanmar Post and Telecommunications, through which subscribers had access to a basic version of Facebook without additional data charges.[271] Due to the relative unfamiliarity of the population with the internet and the easy and cheap access to Facebook, Facebook became the main platform for online news and internet use in Myanmar, with as many as 20 million users.[272] As such, Facebook also became widely used to spread hate speech and disinformation by the MaBaTha, government officials and the Tatmadaw.[273] Soldiers of the Tatmadaw created troll accounts, news and celebrity pages, and used these to post incendiary comments and posts.[274]

According to Facebook's Community Standards, Facebook will remove content that violates its policies, which includes hate speech. Content moderation is effectuated through moderators and technology. It became apparent that Facebook did not have enough moderators that were able to understand the Myanmar language and its nuances, as well as the context in which comments were made, nor was Facebook's technology compatible with Myanmar language fonts.[275] The IIFFMM put this to the test by flagging a post that contained hate speech and that was commented upon with racial slurs and dehumanising language. The post was shared and re-posted over a thousand times. The IIFFMM notes that it had to report the post to Facebook four times, and in each instance the post was not deemed to go against Facebook's Community Standards.[276]

In August 2018, Facebook started removing Myanmar military officials from Facebook for engaging in "coordinated inauthentic behaviour,"[277] which is when groups or pages work together to mislead others.[278] Facebook states that it has removed specific accounts, such as Senior-General Min

[261] UN Human Rights Council (2018, A/HRC/39/64), para 100.
[262] UN Human Rights Council (2018, A/HRC/39/64), para 83-89.
[263] UN Human Rights Council (2018, A/HRC/39/64), para 90.
[264] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1303.
[265] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1310.
[266] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1302-1303.
[267] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1342.
[268] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1343.
[269] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1344.
[270] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1349.
[271] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1344.
[272] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1344-1345.
[273] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1346.
[274] Mozur, *The New York Times* 15 October 2018.
[275] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1349.
[276] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1351.
[277] About.fb.com/news/2018/08/removing-myanmar-officials.
[278] About.fb.com/news/2018/12/inside-feed-coordinated-inauthentic-behavior.

Aung Hlaing,[279] in response to IIFFMM's evidence that these individuals and organisations committed or enabled serious human rights violations. Facebook wishes to prevent further misuse to inflame ethnic and religious tensions.[280]

In its recommendations, the IIFFMM advises the international community, through the UN, to assist Myanmar in meeting its responsibility to protect its people from genocide, crimes against humanity and war crimes, and that it should take collective action in accordance with the UN Charter as necessary.[281] The IIFFMM also has specific recommendations for the SMCs.[282] The IIFFMM recommends SMCs to apply IHRL as basis for content moderation, and to adopt the UNGPs approach. The SMCs should support existing research of the UN Special Rapporteur[283] on content moderation.[284] With respect to hate speech, the IIFFMM recommends that the UN Special Rapporteur explores the responsibilities of SMCs as carriers of hate speech. The SMCs should allow for independent and thorough examination of the use of their platforms to spread hate speech in Myanmar, and they should open themselves up to public accountability and transparency. The SMCs should hire content moderators that are familiar with the context, background and nuances of Myanmar language, and the issue of hate speech in the country. The SMCs should establish early warning systems for emergency escalation, and they should remove all death threats and threats of harm immediately when detected. This warning system should be developed and operated transparently and should be supported by a formal stakeholder group to provide advice and to monitor performance. The SMCs should review their advertising models to ensure that Myanmar users have access to credible sources of information and alternative views. Finally the IIFFMM recommends that the SMCs should conduct in-depth human right impact assessments before they enter any new market, but especially those with volatile ethnic, religious or other social tensions.[285]

4.3. Duty of Due Diligence of the United States as Home State of Meta

The primary responsibility with respect to the application of IHRL lies with the State.[286] Myanmar is not a party to the ICCPR.[287] However, Myanmar is bound by the UN Charter,[288] and voted in favour of the UDHR in 1948. The IFFMM considers the UDHR reflective of customary international law and applicable to Myanmar.[289] Article 19 UDHR protects the freedom of expression, and article 7 UDHR outlines the right to equality which includes the protection against incitement to discrimination. The right to equality is also enshrined in article 1(2) of the UN Charter. Myanmar is also party to the CPPCG.[290] Based on the report, one can conclude that Myanmar has systematically breached IHRL and *jus cogens*.

Article 1 CPPCG requires States to prevent and to punish the crime of genocide. Under article 48 ARSIWA, any State is entitled to invoke the responsibility of another State if the obligation breached is owed to the international community as a whole. On 11 November 2019, The Gambia instituted proceedings against Myanmar for violating the CPPCG, and requested the ICJ for

[279] In the 2021 military coup, Min Aung Hlaing overturned the elected government and established himself as Myanmar's prime minster. See also: Britannica.com/biography/Min-Aung-Hlaing.
[280] About.fb.com/news/2018/08/removing-myanmar-officials.
[281] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1699.
[282] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1718-1726.
[283] Being the rapporteur on the promotion and protection of the right to freedom of opinion and expression. *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1719.
[284] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 1720-1721.
[285] A broad summary of *Report Independent International Fact-Finding Mission on Myanmar* 2018, paras 1718-1726.
[286] Chinkin 2018, p. 66.
[287] Kittichaisaree 2022, p. 43.
[288] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 34.
[289] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 41.
[290] *Report Independent International Fact-Finding Mission on Myanmar* 2018, para 40.

provisional measures.[291] The ICJ has complied by issuing an order on 23 January 2020. In this order the ICJ reaffirms that State parties to the CPPCG have a common interest to ensure that acts of genocide are prevented.[292]

From the IIFFMM report it has become clear that Myanmar's government officials, religious authorities and military forces conducted much of the dissemination of both hate speech and atrocity speech through social media platforms, most notably on Facebook. The IIFFMM finds that there was a use of derogatory language, expressions that included instigation to rape, murder and "clearance operations" and concludes that there were factors present to a finding of genocidal intent.[293] In Chapter two, this type of hate speech was defined as atrocity speech.

The US has ratified the CPPCG.[294] One could argue that to prevent genocide, the jamming of social media in Myanmar could be deemed proportional.[295] This task could fall upon the US as home State of Meta,[296] as Meta is incorporated and domiciled in the US.[297] A counterargument would be that this would breach Myanmar's sovereignty under article 2(1) and 2(7) UN Charter. Such a breach is possible in relation to the protection of human rights,[298] but restricting social media would also incur a violation on Myanmar's freedom of expression.[299]

The Tallinn Manual 2.0 opines that extraterritorial jamming is lawful to prevent a violation of jus cogens norms.[300] The US has historically opposed such measures,[301] as it regards jamming as violating the US concept of the freedom of expression, and out of fear for precedent for interference with the Internet.[302] As became apparent in Chapter three, the freedom of expression should not be restricted under the First Amendment of the US Constitution.[303] This constraint is applicable to the US government. The US refers to the First Amendment as a universal human right.[304] Also, according to US case law, hate speech can only be limited if it poses a clear and present danger. This could be the case when it concerns punishable incitement, where the speaker should be intentionally inciting imminent violent conduct that is likely to occur immediately.[305] Regarding Myanmar, this is very difficult to establish. It would require active monitoring of the US government on Myanmar social media, with active knowledge of the Myanmar language to distil the proper context of the messages posted. Then the US government would have to assert that the violence is likely to occur immediately, which is also quite a feat to accomplish. However, the US Supreme Court ruled that foreign citizens abroad do not have any rights under the US Constitution.[306] This means that the extraterritorial application of the First Amendment is contentious towards Myanmar citizens.

Regardless of US' reservations to such a measure, it is questionable whether jamming social media access within Myanmar borders would be considered prescribed by law, having a legitimate aim, and necessary in accordance with article 19(3) ICCPR. Even if the restriction would pass the first

---

[291] ICJ 11 November 2019, *Application instituting proceedings and request for provisional measures* (*The Gambia/Myanmar*).
[292] ICJ 23 January 2020, *Application of the Convention on the Prevention and Punishment of the Crime of Genocide* (*The Gambia/Myanmar*) (Order), para 41. See also ICJ 26 February 2007, *Case Concerning Application of The Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia and Herzegovina/Serbia and Montenegro)*, para 427.
[293] *Report Independent International Fact-Finding Mission on Myanmar* 2018, paras 1419-1424.
[294] Treaties.un.org/pages/ViewDetails.aspx?src=IND&mtdsg_no=IV-1&chapter=4&clang=_en.
[295] Rapp, *Journal of Human Rights* 2021, p. 495.
[296] Rapp, *Journal of Human Rights* 2021, p. 494.
[297] Rapp, *Journal of Human Rights* 2021, p. 491.
[298] Kittichaisaree 2022, p. 243.
[299] Rapp, *Journal of Human Rights* 2021, p. 494.
[300] Tallinn Manual 2.0, rule 63 no. 11, p. 297.
[301] See in general: Metzl, *The American Journal of International Law* 1997.
[302] Metzl, *The American Journal of International Law* 1997, p. 646.
[303] Strossen 2018, p. 7-8.
[304] Zick, *Notre Dame Law Review* 2010, p. 1549.
[305] Strossen 2018, p. 60-62.
[306] Editorial, *Harvard Law Review* 2020, p. 490.

two criteria, it would arguably not be reasonable and proportionate to the goal sought, as it would affect all Myanmar's citizens' rights. It is not a single radio station that would be jammed, but general access to an essential utility, widely used by Myanmar citizens for all kinds of purposes.

Another issue is the fact that, under US law, SMCs have distinct legal rights and obligations, including First Amendment rights.[307] An SMC can restrict speech, but the US government cannot impose restrictions. Attempts to allow for more government involvement on SMC platforms have been stopped by US courts. It is unlikely that the US government would violate an SMCs First Amendment rights outside US borders, as it is considered a US citizen.

The US has not prevented the further dissemination of atrocity speech on Facebook in Myanmar due to its constitutional constraints. The US has taken other countermeasures under international law by imposing sanctions on Myanmar, including on military leaders, for human rights abuses against the Rohingya.[308]

4.4. Meta's Corporate Responsibility for Breaches of IHRL and Jus Cogens on Facebook

International law does not impose duties directly to Meta to refrain from abusing human rights, unless it concerns "egregious" conduct, or serious violations of peremptory norms. The question is whether Meta, as internet intermediary, may be considered complicit for allowing atrocity speech on Facebook. The legal enforcement mechanisms for either egregious conduct or complicity therein by businesses lie with domestic courts, which may determine such a breach based on customary international law, human rights treaties or even the Rome Statute.[309]

The IIFFMM report includes detailed examples of hate speech and atrocity speech disseminated on Facebook. Some of these examples could be considered direct and public incitement to genocide in accordance with article III(c) CPPCG, or persecution as a CAH.[310] This means the platform was used for violations of peremptory norms. In December 2021, a $150 billion lawsuit was filed in California on behalf of Rohingya refugees against Meta, for strict product liability and negligence, arguing that Meta's failure to police content and its platform's design contributed to violence against the Rohingya.[311] In the lawsuit, the plaintiff refers to a declaration from a Facebook whistle blower claiming that Facebook was fully aware of the use of its platform for atrocity speech for years.[312] A spokesperson for Meta states that this allegation is wrong,[313] however the company has acknowledged that it failed to prevent its platform from being used to "incite offline violence" in Myanmar.[314] From an international law perspective, negligence is not sufficient to determine liability in this respect.

As established in Chapter three, Meta is not liable for user content under US Section 230. Under international law, if Meta is not aware of the content that is disseminated on Facebook, it may lack accountability.[315] In Chapter two, a number of cases were examined to determine the requirements for direct and public incitement to genocide, and secondarily, persecution as a CAH. A requirement for the element "public incitement" was dissemination by such means as mass

[307] Editorial, *Harvard Law Review* 2020, p. 493.
[308] Kittichaisaree 2022, p. 242. See also: State.gov/burma-sanctions.
[309] Ruggie 2013, p. 68-71. The ICC itself does not have jurisdiction over corporations: in article 25(1) of the Rome Statute, it is clearly stated that the ICC has jurisdiction over natural persons.
[310] Gordon 2017, p. 166.
[311] Chandran & Asher-Schapiro, *Reuters* 10 December 2021, with reference to *Doe v. Meta Platforms, Inc.*, Class Action Complaint filed with San Mateo Superior Court (2021).
[312] *Doe v. Meta Platforms, Inc.*, Class Action Complaint filed with San Mateo Superior Court (2021), para 2. Facebook has since changed its name to Meta.
[313] Timberg, *The Washington Post* 22 October 2021.
[314] Bbc.com/news/world-asia-46105934.
[315] Oster 2017, p. 66.

media.[316] Facebook is considered mass media.[317] The application of article III(c) to mass media is based on the purpose of the publisher or editor's message as well as the context thereof.[318] Notably, in mass media cases, the publisher or editor runs a television or radio station, or a newspaper. However, as an internet intermediary, Meta offers platforms for all kinds of mass media outlets. It cannot be considered the publisher or editor of their messages, as in principle it does not review messages before they are posted by users. Rather, Meta takes messages down after they have either been flagged by another user or flagged and reviewed by its (automatic) content moderators. Most importantly, as determined in Chapter two, both the crime of direct and public incitement to genocide and the crime of persecution require intent. Intent is also required for complicity.[319] Intent cannot be established with respect to Facebook's role in Myanmar. In conclusion, Meta cannot be regarded as complicit in the use of Facebook by Myanmar officials to propagate atrocity speech resulting in atrocity crimes.

However, this conclusion is without prejudice to the recommendations of the IIFFMM for SMCs. In addition, Meta has worked on its corporate responsibilities under the UNGPs to address adverse human rights impacts by updating its code of conduct and blocking several Myanmar Facebook accounts.

4.5. Preliminary Conclusion

The case study of Myanmar shows how precarious the self-regulation of SMC platforms is, lacking governmental oversight. Facebook has the means to prohibit atrocity speech on its platforms, but did not intervene in a timely manner in Myanmar, despite the ensuing atrocities against the Rohingya. The US as home State did not intervene on the platform due to constitutional constraints. This leads to the important question of whether SMCs should still be left to self-regulate from an international law perspective, especially when in conflict-affected areas it is almost impossible for the SMC platforms to not become involved with gross human rights abuses. In Chapter 5 two possible solutions for the regulation of SMCs will be explored.

[316] ICTR 2 September 1998, Case No. ICTR-96-4-T, (*Prosecutor/Akayesu*), Judgement, para 556.
[317] Oster 2017, p. 1. Oster defines media 'as a means to disseminate information and ideas to a mass audience,' which are any communication intermediaries transmitting third-party information, such as internet service providers.
[318] ICTR 3 December 2003, Case No. ICTR-99-52-T (*Prosecutor/Nahimana et al.*), Judgement, para 1000-1004.
[319] Greenfield, *Journal of Criminal Law and Criminology* 2008, p. 921.

5. Possible Solutions: the Meta Oversight Board and the EU Digital Services Act

5.1. Introduction

The regulation of the freedom of expression on social media platforms is conducted mainly by the SMCs. As the SMCs operate in multiple jurisdictions, their liability as internet intermediaries varies per State.

As was discussed in Chapter three, Meta, Alphabet and Twitter have codes of conduct in place addressing hate speech. Additionally, Meta has installed the Oversight Board (OB) for its platforms which 'ensures respect for freedom of expression, through independent judgment.'[320] In the introduction of this thesis, the possible acquisition of Twitter by Elon Musk was mentioned. Mr. Musk is a proponent of the US First Amendment, and has announced that under his leadership, Twitter would allow all speech protected by the First Amendment.[321] As evidenced in Chapter three, the First Amendment allows fewer restrictions on hate speech than international law does. It is one thing for Twitter to apply First Amendment standards to US-based users, but the case of Myanmar showed that ungoverned online hate speech can have atrocious consequences.[322] The SMCs policies should not depend on the SMCs' ownership, nor should the policies be easily amended if ownership changes.

The EU reminded Mr. Musk that any company, regardless of their shareholding, needs to comply with EU rules.[323] This refers to a new EU regulation, the Digital Services Act (DSA), aimed to create oversight on the SMCs and to harmonise intermediary liability legislation within the territory of the EU Member States on the subject of fundamental human rights. The DSA will enter into force on 1 January 2024 at the latest.[324] Following the announcement of the DSA, several authors commented that the DSA could garner the so-called 'Brussels Effect,'[325] similar to the extraterritorial effect of the EU General Data Protection Regulation (GDPR).[326] This may be the case if the SMCs decide to implement the new EU rules into their global governance, rather than in one geography only.[327]

This chapter will look at two possible solutions for the regulation of social media platforms. First, the OB will be discussed as an example of SMC self-regulation. Thereafter the proposal of the European Commission for the DSA will be explored,[328] as well as its potential to have the Brussels Effect, similarly to the GDPR. The DSA could serve as a good example of government regulation of SMCs.

[320] Oversightboard.com.
[321] Rosen, *The Atlantic* 2 May 2022.
[322] UN Report of the Special Rapporteur (David Kaye) on the promotion and protection of the right to freedom of opinion and expression, UN Doc. A/74/486 (9 October 2019), para 41.
[323] Milmo & Sweney, *The Guardian* 27 April 2022. See also: Twitter.com/ThierryBreton/status/1518904143066320896.
[324] Digital-strategy.ec.europa.eu/en/policies/digital-services-act-package. The related Digital Markets Act falls outside of the scope of this thesis.
[325] Bradford, *Northwestern University Law Review* 2012 and Bradford 2020.
[326] For instance: Satariano, *The New York Times* 24 March 2022.
[327] Satariano, *The New York Times* 24 March 2022.
[328] COM(2020)825 final.

5.2. The Meta Oversight Board

The OB was constituted in 2020 to help Meta[329] weigh difficult questions regarding the freedom of expression,[330] following insight within Meta that it should not be making so many decisions in this respect on its own. The OB is empowered to select highly emblematic cases for review and to uphold or reverse Meta's content decisions in accordance with Meta's values and policies.[331]

An example of a highly emblematic case is Meta's decision on 7 January 2021 to restrict then-President Donald Trump's access to his Facebook and Instagram account, following the Capitol Riots that occurred on 6 January 2021.[332] In its decision, in which the OB upholds Meta's decision, it specifically refers to the ICCPR, and not the First Amendment. This choice can be explained by the fact that international courts weigh and balance individual rights against other rights with an emphasis on preserving the integrity of the broader system of rights. The American common law system tends to treat rights as property that individuals either have or do not have.[333] The decisions of the OB are binding on Meta according to article 4 of its charter.[334]

The number of posts that are reviewed by Meta are in the millions each day.[335] The OB only reviews a few cases per year, and only if a user appeals Meta's decision[336] to either remove or keep a post on its platforms. It seems quite unlikely that removed posts containing atrocity speech will be appealed, and that the OB would pick such a case for review.

The OB cannot be compared with an independent court simply because it reviews cases in which it references IHRL. Moreover, it operates through a Meta framework that addresses a narrow and predetermined set of concerns. Meta is the creator of the OB and the settlor of the board of Trustees.[337] Finally, the OB decisions are applicable to Facebook and Instagram only, and have no effect on other social media platforms.

In a recent decision, the OB advises Meta to clarify the Hate Speech Community Standard, as it found that content moderators interpret this policy to only apply to explicit content. Implicit violations, such as comparing certain ethnic groups to rats, were not removed from Meta's platforms.[338] This decision shows how the OB, in time, could enhance Meta's self-regulation, insofar as Meta upholds and implements the advice included in the decisions. This in turn could help content moderators act more swiftly in removing abusive content on Meta's platforms.

[329] At the time Meta was called Facebook.
[330] Klonick, *The New Yorker* 12 February 2021.
[331] Oversightboard.com.
[332] Meta Oversight Board, Case decision 2021-001-FB-FBR.
[333] Gopal, *American Bar Association Business Law Section* 12 October 2021.
[334] Oversightboard.com/governance.
[335] Facebook.com/notes/751449002072082/?hc_location=ufi.
[336] Oversightboard.com/appeals-process.
[337] Gopal, *American Bar Association Business Law Section* 12 October 2021. See also: Oversightboard.com/governance.
[338] Meta Oversight Board, Case decision 2022-001-FB-UA.

5.3. The EU and the Digital Services Act

Article 2 of the Treaty on European Union (TEU) declares that the EU is founded on the value of respect for human rights.[339] All EU Member States are also parties to the ICCPR and the ECHR.[340] The EU has exclusive competence on defined areas.[341] This means that the EU may adopt legally binding acts such as regulations and directives.[342]

The EU has determined that there is legal fragmentation in the EU with respect to the protection of fundamental rights online, and that this results in rule setting and enforcement by large SMCs.[343] The DSA is a regulation which aims to achieve harmonisation of legislation across the EU.[344] It will amend the current 2000 E-Commerce Directive,[345] which came into existence when the SMC platforms did not yet exist. The DSA provides a governance framework applicable to all providers of intermediary services.[346] It will establish additional due diligence obligations for large SMCs,[347] aiming to set a higher standard of transparency and accountability on how the SMCs moderate content.[348] Provided the SMC is compliant with the DSA, it will enjoy exemption from liability.[349]

The DSA is explicitly applicable to SMCs outside of the EU who operate on the EU's internal market,[350] which includes Meta, Alphabet and Twitter. The more restrictive measures targeted at these large SMCs are considered proportionate to their ability to comply. A specific obligation that will be imposed on the SMCs is the appointment of a legal representative in the EU to allow for effective oversight.[351] The SMC will be under the jurisdiction of the Member State where the legal representative is appointed.[352] This will have an effect on domestic court jurisdiction with respect to the SMCs liability cases within the EU. The DSA will also require Member States to appoint Digital Services Coordinators who shall be responsible for the application and enforcement of the DSA.[353] A

[339] Article 2 TEU: 'The Union is founded on the values of respect for human dignity, freedom, democracy, equality, the rule of law and respect for human rights, including the rights of persons belonging to minorities. These values are common to the Member States in a society in which pluralism, non-discrimination, tolerance, justice, solidarity and equality between women and men prevail.'

[340] The EU is not party to the ECHR yet, despite its article 6(2) TEU: 'The Union shall accede to the [ECHR].' The accession is currently being negotiated between the EU and the Council of Europe. See: coe.int/en/web/human-rights-intergovernmental-cooperation/accession-of-the-european-union-to-the-european-convention-on-human-rights. The EU has its own catalogue of human rights, the Charter of Fundamental Rights of the European Union, which forms part of the binding treaties of the EU, and refers explicitly to the ECHR and the ECtHR in its preamble and in its article 52(3).

[341] Article 3 TFEU

[342] Article 288 TFEU: 'To exercise the Union's competences, the institutions shall adopt regulations, directives […]. A regulation shall have general application. It shall be binding in its entirety and directly applicable in all Member States. A directive shall be binding, as to the result to be achieved, upon each Member State to which it is addressed, but shall leave to the national authorities the choice of form and methods.'

[343] COM(2020)825 final, p. 11.

[344] COM(2020)825 final, p. 3.

[345] Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce').

[346] COM(2020)825 final, p. 12.

[347] COM(2020)825 final, article 1(a) and 1(b). The threshold for SMCs in scope of the additional DSA obligations are those with a significant reach in the Union, being 10% of the EU's population, or 45 million recipients of the SMCs service: COM(2020)825, final p. 3.

[348] COM(2020)825 final, p. 2.

[349] Generally, COM(2020)825 final, as an example p. 22, para 23.

[350] COM(2020)825 final, p. 6.

[351] COM(2020)825 final, p. 26, para 37.

[352] COM(2020)825 final, p. 37, para 76.

[353] COM(2020)825 final, article 38.

European Board for Digital Services will be established, composed of Digital Services Coordinators, to contribute to the supervision of the large SMCs.[354]

The DSA also requires the implementation of Codes of Conduct that contribute to the application of the DSA,[355] with reference to the EU Code of Conduct against illegal hate speech (among other codes).[356] The DSA chooses not to define harmful content, as it considers it a delicate area with severe implications for the protection of the freedom of expression.[357] It does give a broad description of "illegal content," which should be understood to also refer to "illegal hate speech."[358] The SMCs will be required to put notice and action mechanisms in place to allow any individual or entity to notify the SMC of illegal content.[359] The DSA will also introduce "trusted flaggers."[360] Trusted flaggers are entities that have particular expertise and competence in tackling illegal content, such as national or European law enforcement authorities or (non-governmental) organisations 'committed to notifying illegal racist and xenophobic expressions online.'[361] The SMCs will be required to ensure that notices of trusted flaggers will be processed and decided upon with priority and without delay.[362] This mechanism could be especially effective for identifying and removing hate speech and atrocity speech, possibly ensuring the prevention of the rapid spread of such content. This in turn could prevent physical harm from occurring. It is imaginable that this could even be effective in the identification and removal of hate speech and atrocity speech outside of EU territory. It would not be an exercise of extraterritorial jurisdiction, as it would still be up to the SMC to decide to remove the flagged content.

Other specific obligations imposed on the SMCs will be conducting systemic risk assessments, including on hate speech;[363] the deployment of mitigating measures such as enhancing or adapting their content moderation;[364] and accountability through independent auditing.[365]

The DSA does emphasise the freedom of SMCs to determine terms and conditions and establish stricter measures in the case of manifestly illegal content related to serious crimes, provided that this is set out clearly and in sufficient details in the terms and conditions.[366] In line with a recent CJEU ruling,[367] the DSA will also prohibit a "general monitoring obligation," which consists of active monitoring of information that is transmitted or stored by the internet intermediary or active searching for facts or circumstances indicating illegal activity.[368] Such an obligation is perceived as disproportionately limiting user's freedom of expression and an excessive burden on service providers.[369]

---

[354] COM(2020)825 final, article 47.
[355] COM(2020)825 final, p. 34, para 67.
[356] COM(2020)825 final, p.35, para 69, the EU Code of Conduct was drafted together with Facebook, Twitter and YouTube.
[357] COM(2020)825 final, p. 9.
[358] COM(2020)825 final, p. 20, para 12.
[359] COM(2020)825 final, article 14(1).
[360] COM(2020)825 final, p. 13.
[361] COM(2020)825 final, p. 28, para 46.
[362] COM(2020)825 final, article 19(1).
[363] COM(2020)825 final, p. 31, para 57.
[364] COM(2020)825 final, p. 32, para 58.
[365] COM(2020)825 final, p. 32, para 60.
[366] COM(2020)825 final, p. 28, para 47.
[367] CJEU 3 October 2019, C-18/18, ECLI:EU:C:2019:821 (*Glawischnig-Piesczek/Facebook Ireland Limited*), para 34. The Court of Justice of the European Union (CJEU) ensures EU law is interpreted and applied alike in the Member States, and that the Member States and EU institutions abide by EU law. See also: European-union.europa.eu/institutions-law-budget/institutions-and-bodies/institutions-and-bodies-profiles/court-justice-european-union-cjeu_en.
[368] COM(2020)825 final, article 7.
[369] COM(2020)825 final, p. 13.

5.4. The Possible Brussels Effect of the DSA

The effects of EU law can impact third countries. This effect is not so much of an extraterritorial nature as being a territorial extension of EU law. Territorial extension is when an assessment of compliance with EU law requires an evaluation of foreign conduct and/or third country law.[370] As such, EU law's global reach is extended.[371] The territorial extension of EU law can also garner the Brussels Effect.

The theory underpinning the Brussels Effect is that a jurisdiction can externalise its laws and regulations outside its borders through market mechanisms, resulting in unilateral regulatory globalisation of standards.[372] The standards should be aimed at consumers. As Bradford states

> This presumes that the benefits of adopting a uniform global standard exceed the benefits of adhering to multiple, including laxer, regulatory standards. This is the case in particular when the firms' conduct or production is nondivisible, meaning that it is not legally or technically feasible, or economically viable, for the firm to maintain different standards in different markets.[373]

The Brussels Effect has two components. The *de facto* component occurs when third country companies that are required to comply with EU rules, standardise their production globally by adhering to the EU rule. The *de jure* component occurs when the third country decides to incorporate the substance of EU law into its domestic law, often lobbied to do so by TNCs.[374]

The GDPR is used as an example of an EU regulation that has arguably garnered the Brussels Effect, where "data subjects" are considered similar to "consumers."[375] The GDPR protects individuals with respect to the personal data that is collected and used by organisations engaged in professional or commercial activity. Organisations that do not comply with the GDPR will face heavy penalties. The GDPR also applies to non-EU organisations that offer goods or services to people in the EU, or if they monitor online behaviour.[376] The GDPR covers a broad spectrum of organisations, including SMCs.

The DSA indicates the EU's intention for governmental regulation of the SMCs. The implication is that the DSA may garner the Brussels Effect.[377] An important indication may be Meta's implementation of the GDPR.[378] Meta's users were either governed by the company's US based structure, or the Ireland based structure. Until the introduction of the GDPR, Meta's non-US users were governed by the Ireland structure. Meta then decided to change its terms of service, moving all non-EU users to the US structure.[379] These users were now removed effectively from the GDPR's reach. The GDPR cannot be completely disregarded with respect to these users, but Meta can still be selective in its application.[380] A more in-depth analysis showed that Meta has a different data policy in place for non-EU users than its EU users.[381] A comparative research between the GDPR and the California Consumer Privacy Act did show that some GDPR provisions garnered *de facto* Brussels Effect, as companies implemented the provisions in their global standards.[382] However, the GDPR's liability and judicial remedy provisions in particular are not applied globally.[383]

---

[370] Scott 2019, p. 22.
[371] Scott 2019, p. 31.
[372] Bradford, *Northwestern University Law Review* 2012, p. 5.
[373] Bradford, *Northwestern University Law Review* 2012, p. 5.
[374] Bradford, *Northwestern University Law Review* 2012, p. 6.
[375] Gunst & De Ville, *European Foreign Affairs Review* 2021, p. 439.
[376] Gdpr.eu/companies-outside-of-europe.
[377] Also posited by Bradford, the author of 'The Brussels Effect': Bradford, *Project Syndicate* 17 December 2020.
[378] Meta was called Facebook at the time.
[379] Bradford 2020, p. 57.
[380] Gunst & De Ville, *European Foreign Affairs Review* 2021, p. 457.
[381] Gunst & De Ville, *European Foreign Affairs Review* 2021, p. 444.
[382] Gunst & De Ville, *European Foreign Affairs Review* 2021, p. 437.
[383] Gunst & De Ville, *European Foreign Affairs Review* 2021, p. 457.

For now it remains to be seen whether specific provisions of the DSA will also acquire the Brussels Effect, whether *de facto* or *de jure*. It seems unlikely that the SMCs will lobby in the US or in third countries to amend domestic provisions to reflect those of the DSA. However, the SMCs may choose to implement the DSA code of conduct requirements globally. Third countries may use the DSA as an example to introduce or amend legislation on governmental regulation and intermediary liability.

5.5. Preliminary Conclusion

The Meta OB is an interesting instance of SMC self-regulation, but its effectiveness for the long term remains to be seen. The OB does provide recommendations on the improvement of Meta's user policies, and its decisions may serve as examples of what types of expression should be protected, restricted or prohibited on Meta's platforms in accordance with IHRL. If other SMCs introduce similar independent bodies for policy review, this could improve self-regulation by attuning their codes of conduct to IHRL. However, the OB does not add to governmental oversight on Meta's platforms.

The DSA aims to harmonise legislation within the EU to protect fundamental human rights online, which includes the freedom of expression. The DSA allows SMCs to regulate their platforms in accordance with their own codes of conduct, as long as these policies are in compliance with the DSA's minimum standards. These standards are aligned with IHRL as EU Member States are also parties to the ICCPR and EHRM. The "trusted flaggers" may prove to be effective in the restriction of hate speech and the prohibition of atrocity speech. These entities could potentially flag content outside of EU jurisdiction, and consequently work together with SMCs to prevent atrocity crimes in third countries. Whether the DSA's will garner the Brussels Effect is not yet certain. However, the DSA is an important step forward in the governmental regulation of SMC platforms.

6. Conclusion

This thesis aimed to clarify to what extent States have international human rights law obligations with respect to the regulation of social media platforms regarding hate speech, especially when atrocity speech leads to breaches of jus cogens. Under IHRL, States have obligations to respect, protect and fulfil human rights, including in cyberspace. The freedom of expression is a human right which should be protected. It may only be restricted or even prohibited under certain conditions. Hate speech is a form of expression that is not clearly defined in international law, but is generally considered to be a form of expression which is abusive towards individuals or groups, and breaches their right to equality and non-discrimination. Under IHRL, it may be restricted. Atrocity speech is a form of hate speech of such nature that it may lead to gross violations of *jus cogens*, such as direct and public incitement to genocide and persecution as a CAH. States have obligations towards the international community as a whole to protect against atrocity crimes, including the duties to prevent and to act. On this basis, one could conclude that international law provides for a robust IHRL regime with respect to the duties of the State.

The thesis focussed on SMCs incorporated under US law for two reasons. First, the US-based SMCs are the largest globally, both in number of users and revenue. These SMCs have a huge commercial interest to keep users engaged on their platforms, which are used for the dissemination of all types of expression. Secondly, the US rarely restricts hate speech in accordance with the First Amendment, and it offers SMCs a broad exemption from liability for user content. In both respects, the US deviates from international law, although international law also maintains a limited liability regime regarding SMCs.

SMCs do not have duties under IHRL. Based on the UNGPs, a soft law instrument, SMCs do have a corporate responsibility to respect human rights. This is reflected in their codes of conduct, which are mostly aligned with IHRL. However, this proves to be precarious as the codes of conduct are drafted by the SMCs, and the contents may depend on ownership. If SMC ownership changes, the policies may be amended as well. There is no international law instrument in place to enforce IHRL compliancy by SMCs, this is a domestic State affair. The US will not interfere as it regards SMCs as private entities with their own constitutional rights.

The SMCs are allowed a lot of room to self-regulate their platforms. That this can be very problematic was evidenced by the case study of the role of Facebook in Myanmar. Myanmar did not comply with its IHRL obligations as its own government was responsible for the spread of atrocity speech on Facebook, resulting in atrocity crimes against Rohingya. The US did not intervene due to constitutional constraints, it was up to Meta to restrict access to and regulate expression on Facebook within the borders of Myanmar. Meta failed to do this properly, contributing to gross violations of *jus cogens*. However, Meta's responsibility in this respect is contentious. This leads to the conclusion that there is a serious gap in IHRL where it concerns the governmental regulation, whether territorially or extraterritorially, of SMC platforms. Especially where it concerns conflict-affected areas and egregious human rights abuses, such responsibility should not be solely in the hands of an SMC.

Meta has enhanced its self-regulation by installing the OB, although the OB does not add to governmental oversight. The DSA imposes obligations on the largest SMCs regarding the regulation of expression. The DSA will also curb whims of SMC shareholders to change policies, as it sets a minimum standard for the codes of conduct in line with IHRL. Compliance with these requirements will exempt the SMCs from liability. The DSA could garner the Brussels Effect if SMCs choose to implement these standards outside of the EU, and if the DSA serves as an example for third countries to introduce similar legislation. In conclusion, the DSA has great potential to fill the void regarding government regulation of SMC platforms concerning all forms of expression.

Bibliography

Books

Bradford 2020
A. Bradford, *The Brussels Effect, How the European Union Rules the World*, New York: Oxford University Press 2020.

Gordon 2017
G.S. Gordon, *Atrocity Speech Law, Foundation, Fragmentation, Fruition*, New York: Oxford University Press 2017.

Hernández 2019
G. Hernández, *International Law*, Oxford: Oxford University Press 2019.

Kittichaisaree 2022
K. Kittichaisaree, *The Rohingya, Justice and International Law*, Oxon: Routledge 2022.

Milanovic 2011
M. Milanovic, *Extraterritorial Application of Human Rights Treaties. Law, Principles, and Policy*, Oxford: Oxford University Press 2011.

Orakhelashvili 2006
A. Orakhelashvili, *Peremptory Norms in International Law*, Oxford: Oxford University Press 2006.

Oster 2017
J. Oster, *European and International Media Law*, Cambridge: Cambridge University Press 2017 (kindle version).

Ruggie 2013
J.G. Ruggie, *Just Business, Multinational Corporations and Human Rights*, New York: W.W. Norton & Company 2013 (kindle version).

Schabas 2021
W.A. Schabas, *The Customary International Law of Human Rights*, Oxford: Oxford University Press 2021.

Strossen 2018
N. Strossen, *HATE Why We Should Resist It with Free Speech, Not Censorship*, New York, Oxford University Press 2018.

Tallinn Manual 2.0
M.N. Schmitt & L. Vihul (Eds.), *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations*, Cambridge: Cambridge University Press 2017.

Book Contributions

Byrnes & Renshaw 2018
Byrnes & Renshaw, 'Within the State', in: D. Moeckli et al. (Eds.), *International Human Rights Law*, Oxford: Oxford University Press 2018.

Chinkin 2018
C. Chinkin, 'Sources', in: D. Moeckli et al. (Eds.), *International Human Rights Law*, Oxford: Oxford University Press 2018.

Clapham 2018
A. Clapham, 'Non-State Actors', in: D. Moeckli et al. (Eds.), *International Human Rights Law*, Oxford: Oxford University Press 2018.

De Wet 2013
E. de Wet, 'Jus Cogens and Obligations Erga Omnes', in: D. Shelton (Ed.), *The Oxford Handbook of International Human Rights Law*, Oxford: Oxford University Press 2013.

Hwang 2020
T. Hwang, 'Dealing with Disinformation: Evaluating the Case for Amendment of Section 230 of the Communications Decency Act', in: N. Persily & J.A. Tucker (Eds.), *Social Media and Democracy, The State of the Field, Prospects for Reform*, Cambridge: Cambridge University Press 2020.

Joseph & Dipnall 2018
S. Joseph & S. Dipnall, 'Scope of Application', in: D. Moeckli et al. (Eds.), *International Human Rights Law*, Oxford: Oxford University Press 2018.

McGoldrick 2018
D. McGoldrick, 'Thought, Expression, Association, and Assembly', in: D. Moeckli et al. (Eds.), *International Human Rights Law*, Oxford: Oxford University Press 2018.

Mégret 2018
F. Mégret, 'Nature of Obligations', in: D. Moeckli et al. (Eds.), *International Human Rights Law*, Oxford: Oxford University Press 2018.

Persily & Tucker 2020
N. Persily & J.A. Tucker, 'Introduction', in: N. Persily & J.A. Tucker (Eds.), *Social Media and Democracy, The State of the Field, Prospects for Reform*, Cambridge: Cambridge University Press 2020.

Scott 2019
J. Scott, 'The Global Reach of EU Law', in: M. Cremona & J. Scott (Eds.), *EU Law Beyond EU Borders, The Extraterritorial Reach of EU Law*, Oxford: Oxford University Press 2019.

Van Aaken, Motoc & Vasel 2018
A. van Aaken, I. Motoc & J.J. Vasel, 'Introduction', in: Van Aaken & Motoc (Eds.), *The European Convention on Human Rights and General International Law*, Oxford: Oxford University Press 2018.

Van Boven 2018
T. Van Boven, 'Categories of Rights', in: D. Moeckli et al. (Eds.), *International Human Rights Law*, Oxford: Oxford University Press 2018.

Williams 2015
A. Williams, 'Human Rights in the EU', in: A. Arnull & D. Chalmers (Eds.), *The Oxford Handbook of European Union Law*, Oxford: Oxford University Press 2015.

Woolley 2020
S.C. Woolley, 'Bots and Computational Propaganda: Automation for Communication and Control', in: N. Persily & J.A. Tucker (Eds.), *Social Media and Democracy, The State of the Field, Prospects for Reform*, Cambridge: Cambridge University Press 2020.

Dissertations
Solmone 2020
S. Solmone, *The Respect for Freedom of Expression and the Exercise of State Jurisdiction Online in International Human Rights Law* (Doctor of Philosophy thesis, Royal Docks School of Business and Law), 2020.

Journal Articles
Bernaz, *Journal of Business Ethics* 2013
N. Bernaz, 'Enhancing Corporate Accountability for Human Rights Violations: Is Extraterritoriality the Magic Potion?', *Journal of Business Ethics* 2013 (Vol 117) p. 493-511.

Bernaz, *Journal of International Criminal Justice* 2015
N. Bernaz, 'Corporate Criminal Liability under International Law', *Journal of International Criminal Justice* 2015 (Vol. 13 no. 2), p. 313-330.

Bernaz, *Human Rights Review* 2021
N. Bernaz, 'Conceptualizing Corporate Accountability in International Law: Models for a Business and Human Rights Treaty', *Human Rights Review* 2021 (Vol. 22), p. 45-64.

Besson, *ESIL Reflections* 28 April 2020
S. Besson, 'Due Diligence and Extraterritorial Human Rights Obligations – Mind the Gap!', *ESIL Reflections* 28 April 2020 (Vol. 9, Issue 1), esil-sedi.eu.

Bradford, *Northwestern University Law Review* 2012
A. Bradford, 'The Brussels Effect', *Northwestern University Law Review* 2012 (Vol. 107, No. 1), p. 1-67.

Editorial, *Harvard Law Review* 2020
Editorial, 'Agency for International Development v. Alliance for Open Society International, Inc.', *Harvard Law Review* 2020 (Vol: 134:410), p. 490-499.

Farrior, *Berkeley Journal of International Law* 1996
S. Farrior, 'Molding The Matrix: The Historical and Theoretical Foundations of International Law Concerning Hate Speech', *Berkeley Journal of International Law* 1996 (Vol. 14:1), p. 1-98.

Greenfield, *Journal of Criminal Law and Criminology* 2008
D.M. Greenfield, 'The Crime of Complicity in Genocide: How the International Criminal Tribunals for Rwanda and Yugoslavia Got It Wrong, and Why It Matters', *Journal of Criminal Law and Criminology* 2008 (Vol. 98, No. 3), p. 921-952.

Gunst & De Ville, *European Foreign Affairs Review* 2021
S. Gunst & F. De Ville, 'The Brussels Effect: How the GDPR Conquered Silicon Valley', *European Foreign Affairs Review* 2021, p. 437-458.

Hummelbrunner, *Cleer Papers* 2019
S. Hummelbrunner, 'Beyond Extraterritoriality: Towards an EU Obligation to Ensure Human Rights Abroad? Reflections in Light of the *Front Polisario* Saga', *Cleer Papers* 2019/2.

Klonick, *Harvard Law Review* 2018
K. Klonick, 'The New Governors: the People, Rules, and Process Governing Online Speech', *Harvard Law Review* 2018 (Vol. 131), p. 1598-1670.

Maggiore, *Michigan Telecommunications and Technology Law Review* 2012
P. Maggiore, 'Viewer Discretion is Advised: Disconnects between the Marketplace of Ideas and Social Media Used to Communicate Information during Emergencies and Public Health Crises', *Michigan Telecommunications and Technology Law Review* 2012 (Vol. 18, Issue 2), p. 627-660.

McCorquodale & Simons, *The Modern Law Review* 2007
R. McCorquodale & P. Simons, 'Responsibility beyond Borders: State Responsibility for Extraterritorial Violations by Corporations of International Human Rights Law', *The Modern Law Review* (Vol. 70, No. 4), p. 598-625.

Metzl, *The American Journal of International Law* 1997
J.F. Metzl, 'Rwandan Genocide and the International Law of Radio Jamming', *The American Journal of International Law* 1997 (Vol. 91 No. 4), p. 628-651.

Milanovic, *EJIL* 2012
M. Milanovic, 'Al-Skeini and Al-Jedda in Strasbourg', EJIL 2012 (Vol. 23 No. 1), p. 121-139.

Rapp, *Journal of Human Rights* 2021
K. Rapp, 'Social media and genocide: The case for home state responsibility', Journal of Human Rights 2021 (Vol. 20, No. 4), p. 486-502.

Zick, *Notre Dame Law Review* 2010
T. Zick, 'Territoriality and the First Amendment: Free Speech At – And Beyond – Our Borders', *Notre Dame Law Review* 2010 (Vol. 85:4), p. 1543-1628.

News and Magazine Articles

Alba, Koeze & Silver, *The New York Times* 7 June 2021
D. Alba, E. Koeze & J. Silver, 'What Happened When Trump Was Banned on Social Media', *The New York Times* 7 June 2021, nytimes.com.

Barnes & Zakrzewski, *The Washington Post* 31 May 2022
R. Barnes & C. Zakrzewski, 'Supreme Court puts Texas social media law on hold while legal battle continues', *The Washington Post* 31 May 2022, washingtonpost.com.

Bradford, *Project Syndicate* 17 December 2020
A. Bradford, 'The Brussels Effect Comes for Big Tech', *Project Syndicate* 17 December 2020, project-syndicate.org.

Chandran & Asher-Schapiro, *Reuters* 10 December 2021
R. Chandran & A. Asher-Schapiro, 'Analysis: Rohingya lawsuit against Facebook a "wake-up call" for social media', *Reuters* 10 December 2021, reuters.com.

Fried, *Axios.com* 28 April 2022
I. Fried, 'What Musk's free-speech Twitter could unleash', *Axios.com*, 28 April 2022.

Gopal, *American Bar Association Business Law Section* 12 October 2021
L. Gopal, 'Facebook's Oversight Board & the Rule of Law: The Importance of Being Earnest', *American Bar Association Business Law Section* 12 October 2021, businesslawtoday.org.

Klonick, *The New Yorker* 12 February 2021
K. Klonick, Inside the Making of Facebook's Supreme Court, *The New Yorker* 12 February 2021, newyorker.com

Milmo & Sweney, *The Guardian* 27 April 2022
D. Milmo & M. Sweney, 'Twitter takeover: EU and UK warn Elon Musk must comply or face sanctions', *The Guardian* 27 April 2022, theguardian.com.

Mozur, *The New York Times* 15 October 2018
P. Mozur, 'A Genocide Incited on Facebook, With Posts From Myanmar's Military', *The New York Times* 15 October 2018, nytimes.com.

Oremus et al., *The Washington Post* 26 October 2021
W. Oremus et al., 'How Facebook shapes your feed', The Washington Post 26 October 2021, washingtonpost.com.

Rosen, *The Atlantic* 2 May 2022
J. Rosen, 'Elon Musk Is Right That Twitter Should Follow the First Amendment', *The Atlantic* 2 May 2022, theatlantic.com.

Satariano, *The New York Times* 24 March 2022
A. Satariano, 'E.U. Takes Aim at Big Tech's Power With Landmark Digital Act', *The New York Times* 24 March 2022, nytimes.com.

Smith & Van Alstyne, *Harvard Business Review* 12 August 2021
M.D. Smith & M. Van Alstyne, 'It's Time to Update Section 230', *Harvard Business Review* 12 August 2021, hbr.org.

Timberg, *The Washington Post* 22 October 2021
C. Timberg, 'New whistleblower claims Facebook allowed hate, illegal activity to go unchecked', *The Washington Post* 22 October 2021, washingtonpost.com.

Reports and other unpublished work

***Report Independent International Fact-Finding Mission on Myanmar* 2018**
Report of the detailed findings of the Independent International Fact-Finding Mission on Myanmar (2018, UN Doc. A/HRC/39/CRP.2).

Report of the International Law Commission 2001
Report of the International Law Commission on the work of its fifty-third session (2001, UN Doc. A/56/10).

Report of the International Law Commission 2006
Report of the International Law Commission fifty-eighth session (2006, UN Doc. A/61/10).

Šturma 2020
P. Šturma, 'State Responsibility and the European Convention on Human Rights', rm.coe.int.

Online Content

About.facebook.com/company-info.
About.fb.com/news/2017/06/hard-questions-hate-speech.
About.fb.com/news/2018/08/removing-myanmar-officials.
About.fb.com/news/2018/12/inside-feed-coordinated-inauthentic-behavior.
About.fb.com/wp-content/uploads/2021/03/Facebooks-Corporate-Human-Rights-Policy.pdf.
About.google/human-rights.
Abc.xyz/investor/#product-and-business-updates.
Ala.org/advocacy/intfreedom/hate.
Amnesty.org/en/latest/campaigns/2022/02/what-is-big-techs-surveillance-based-business-model.
Bbc.com/news/world-asia-46105934.
Britannica.com/biography/Min-Aung-Hlaing.
Britannica.com/event/United-States-Capitol-attack-of-2021.
Constitution.congress.gov/constitution/amendment-1.
Data.un.org/en/iso/mm.html.
Digital-strategy.ec.europa.eu/en/policies/digital-services-act-package.
Echr.coe.int/documents/questions_answers_eng.pdf .
Coe.int/en/web/human-rights-intergovernmental-cooperation/accession-of-the-european-union-to-the-european-convention-on-human-rights.
Europa.eu/citizens-initiative/how-it-works/faq/faq-eu-competences-and-commission-powers_en.
European-union.europa.eu/institutions-law-budget/institutions-and-bodies/institutions-and-bodies-
Exhibitions.ushmm.org/burmas-path-to-genocide/timeline.
Profiles/court-justice-european-union-cjeu_en.
Facebook.com/notes/751449002072082/?hc_location=ufi.
Gdpr.eu/companies-outside-of-europe.
Help.twitter.com/en/rules-and-policies/defending-and-respecting-our-users-voice.
Help.twitter.com/en/rules-and-policies/hateful-conduct-policy.
Investor.twitterinc.com/home/default.aspx.
Law.justia.com/cases/federal/appellate-courts/F3/433/1199/546158.
Ohchr.org/EN/HRBodies/CCPR/Pages/CCPRIntro.aspx.
Ohchr.org/en/professionalinterest/pages/internationallaw.aspx.
Oversightboard.com.
Oversightboard.com/appeals-process.
Oversightboard.com/governance.
Prnewswire.com/news-releases/elon-musk-to-acquire-twitter-301532245.html.
Statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users.
State.gov/burma-sanctions.
Support.google.com/youtube/answer/2801939?hl=en.
Transparency.fb.com/en-gb/policies/community-standards.
Transparency.fb.com/en-gb/policies/community-standards/hate-speech.
Transparencyreport.google.com/youtube-policy/featured-policies/hate-speech?hl=en.
Treaties.un.org/pages/ViewDetails.aspx?src=IND&mtdsg_no=IV-1&chapter=4&clang=_en.
Twitter.com/ThierryBreton/status/1518904143066320896.
Un.org/sg/en/content/sg/statement/2019-06-18/secretary-generals-remarks-the-launch-of-the-united-nations-strategy-and-plan-of-action-hate-speech-delivered .

Register

Customary International Law
Universal Declaration of Human Rights, General Assembly resolution 217, 10 December 1948.

International Treaties
United Nations Charter, 26 June 1945, entry into force 24 October 1945.
Statute of the International Court of Justice, 24 October 1945.
Convention on the Prevention and Punishment of the Crime of Genocide, 9 December 1948, entry into force 12 January 1951.
International Convention on the Elimination of All Forms of Racial Discrimination Preface, 21 December 1965, entry into force January 4, 1969.
International Covenant on Civil and Political Rights, 16 December 1966, entry into force 23 March 1976.
Vienna Convention on the Law of Treaties, 23 May 1969, entry into force 27 January 1980.
Rome Statute of the International Criminal Court, 17 July 1998, entry into force 1 July 2002.

United Nations Security Council resolutions
Statute of the International Criminal Tribunal for the Former Yugoslavia, Resolution 827, 25 May 1993.
Statute of the International Criminal Tribunal for Rwanda, Resolution 955, 8 November 1994.

International Soft Law
United Nations Guiding Principles on Business and Human Rights, UN Doc. HR/PUB/11/04 (2011) (UNGPs).

Europe
Convention for the Protection of Human Rights and Fundamental Freedoms, 4 November 1950, entry into force 3 September 1953.

European Union
Treaty on European Union, entry into force 1 December 2000.
Treaty on Function of the European Union, entry into force 1 December 2000.
Charter of the Fundamental Rights of the European Union, entry into force 7 December 2000.

EU Law
E-Commerce Directive
Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce').

COM(2020)825 final
Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC.

EU Soft Law
EU Human Rights Guidelines on Freedom of Expression Online and Offline (2014).

Domestic Law
United States of America
47 U.S.C. § 230.
Florida SB 7072: Social Media Platforms.

International Reports and Other Publications

United Nations

International Law Commission

Draft Code of Crimes Against the Peace and Security of Mankind; in: Report of the International Law Commission to the General Assembly, UN Doc A/51/10 (1996).

Articles on the Responsibility of States for Internationally Wrongful Acts (2001).

ILC, 'Second Report on Jus Cogens' (Special Rapporteur Dire Tladi) UN Doc A/CN.4/706 (16 March 2017).

ILC, 'Fourth Report on Jus Cogens' (Special Rapporteur Dire Tladi) UN Doc A/CN.4/727 (31 January 2019).

UN Human Rights Committee

UN Human Rights Committee, General Comment No. 31 (2004, CCPR/C/21/Rev.1/Add.13).

UN Human Rights Committee, General Comment No. 34 (2011, CCPR/C/GC/34).

UN Human Rights Council

UN Human Rights Council (2016, A/HRC/32/L.20).

UN Human Rights Council (2017, A/HRC/RES/34/22).

UN Human Rights Council (2018, A/HRC/39/64).

UN Reports

UN Report of the Special Representative of the Secretary-General (John Ruggie) on the issue of human rights and transnational corporations and other business enterprises, UN Doc. A/HRC/14/27 (9 April 2010).

UN Report of the Special Rapporteur (Frank La Rue) on the promotion and protection of the right to freedom of opinion and expression, UN Doc. A/67/357 (7 September 2012).

UN Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security (2015, A/70/174).

UN Report of the Special Rapporteur (David Kaye) on the promotion and protection of the right to freedom of opinion and expression, UN Doc. A/HRC/38/35 (6 April 2018).

UN Report of the Special Rapporteur (David Kaye) on the promotion and protection of the right to freedom of opinion and expression, UN Doc. A/74/486 (9 October 2019).

UN Report of the Group of Governmental Experts on Advancing Responsible State Behaviour in Cyberspace in the Context of International Security (2021, UN Doc. A/76/135).

Other publications

Joint Declaration on Freedom of Expression and the Internet The United Nations (UN) Special Rapporteur on Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe (OSCE) Representative on Freedom of the Media, the Organization of American States (OAS) Special Rapporteur on Freedom of Expression and the African Commission on Human and Peoples' Rights (ACHPR) Special Rapporteur on Freedom of Expression and Access to Information (1 June 2011).

Council of Europe

Recommendation CM/Rec(2018)2 of the Committee of Ministers of the Council of Europe (7 March 2018), *on the roles and responsibilities of internet intermediaries*.

International Case Law

Arbitration

Arbitral Award 4 April 1928, *Island of Palmas Case* (*Netherlands/United States of America*).

International Court of Justice

ICJ 5 February 1970, *Case Concerning the Barcelona Traction, Light and Power Company, Limited (Belgium/Spain)* (Second Phase).
ICJ 24 May 1980, *Case Concerning United States Diplomatic and Consular Staff in Tehran (United States/Iran).*
ICJ 27 June 1986*, Case Concerning Military and Paramilitary Activities In and Against Nicaragua (Nicaragua/United States of America)*.
ICJ 8 July 1996*, Legality of the Threat or Use of Nuclear Weapons* (Advisory Opinion).
ICJ 6 July 2004, *Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory* (Advisory Opinion).
ICJ 26 February 2007, *Case Concerning Application of The Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia and Herzegovina/Serbia and Montenegro)*.
ICJ 11 November 2019, *Application instituting proceedings and request for provisional measures* (*The Gambia/Myanmar*).
ICJ 23 January 2020, *Application of the Convention on the Prevention and Punishment of the Crime of Genocide* (*The Gambia/Myanmar*) (Order).

International Criminal Tribunal for Rwanda

ICTR 2 September 1998, Case No. ICTR-96-4-T, (*Prosecutor/Akayesu*), Judgement.
ICTR 4 September 1998, Case No. ICTR-97-23-S (*Prosecutor/Kambanda*), Judgement.
ICTR 1 June 2000, Case No. ICTR-97-32-I (*Prosecutor/Ruggiu*), Judgement.
ICTR 3 December 2003, Case No. ICTR-99-52-T (*Prosecutor/Nahimana et al.*), Judgement.

International Criminal Tribunal for the former Yugoslavia

ICTY 7 May 1997, Case No. IT-94-1-T (*Prosecutor/Tadic*), Judgement.

Human Rights Committee

HRC 29 July 1981, Communication No. R.12/52, UN Doc. Supp. No. 40 (*Lopez Burgos/Uruguay*).
HRC 7 December 2017, Communication No. 2285/2013, UN Doc. CCPR/C/120/D/2285/2013 (*Issa Yassin/Canada*).

European Court of Human Rights

ECtHR 7 December 1976, ECLI:CE:ECHR:1976:1207JUD000549372 (*Handyside/United Kingdom*).
ECtHR 26 April 1979, ECLI:CE:ECHR:1979:0426JUD000653874 (*Sunday Times/United Kingdom*).
ECtHR 23 September 1994, ECLI:CE:ECHR:1994:0923JUD001589089 (*Jersild/Denmark*).
ECtHR 9 December 1994, ECLI:CE:ECHR:1994:1209JUD001679890 (*López Ostra/Spain*).
ECtHR 18 December 1996, ECLI:CE:ECHR:1996:1218JUD001531889 (*Loizidou/Turkey*).
ECtHR 9 June 1998, ECLI:CE:ECHR:1998:0609JUD002267893 (*Incal/Turkey*).
ECtHR 28 October 1998, ECLI:CE:ECHR:1998:1028JUD002345294 (*Osman/United Kingdom*).
ECtHR 8 July 1999, ECLI:CE:ECHR:1999:0708JUD002668295 (*Sürek/Turkey (No. 1)*).
ECtHR 28 September 1999, ECLI:CE:ECHR:1999:0928JUD002247993 (*Özturk/Turkey*).
ECtHR 21 November 2001, ECLI:CE:ECHR:2001:1121JUD003576397 (*Al-Adsani/United Kingdom*).
ECtHR 12 December 2001, ECLI:CE:ECHR:2001:1212DEC005220799 (*Banković/Belgium*)
ECtHR 16 November 2004, ECLI:CE:ECHR:2004:1116JUD003182196 (*Issa/Turkey*).
ECtHR 6 July 2006, ECLI:CE:ECHR:2006:0706JUD005940500 (*Erbakan/Turkey*).
ECtHR 7 July 2011, ECLI:CE:ECHR:2011:0707JUD005572107 (*Al-Skeini/United Kingdom*).
ECtHR 18 December 2012, ECLI:CE:ECHR:2012:1218JUD000311110 (*Yildirim/Turkey*).

Court of Justice of the European Union
CJEU 3 October 2019, C-18/18, ECLI:EU:C:2019:821 (*Glawischnig-Piesczek/Facebook Ireland Limited*).

Domestic Case Law
United States of America
Schenck v. United States, 249 U.S. 47 (1919).
Stratton Oakmont, Inc. v. Prodigy Services Co., N.Y. Sup. Ct. 1794 (1995).
Zeran v. America Online, Inc., 129 F.3d 327 (1997).
Yahoo! Inc. v. La Ligue Contre Le Racisme et L'Antisemitisme, 433 F.3d 1199 (2006).
Manhattan Cmty. Access Corp. v. Halleck, 139 S. Ct. 1921, 1926 (2019).
Doe v. Meta Platforms, Inc., Class Action Complaint filed with San Mateo Superior Court (2021).

France
Ligue contre le racisme et l'antisémitisme et Union des étudiants juifs de France c. Yahoo! Inc. et Société Yahoo! France (*LICRA c. Yahoo!*) (2000).