


Is International Criminal Law Ready to Accommodate Online Harm?

Challenges and Opportunities

Sarah Zarnsky  *

PhD Candidate and Assistant Lecturer, University of Essex School of Law, Human Rights Centre

*Corresponding author. Email: s.zarnsky@essex.ac.uk

ABSTRACT

New technologies have the potential to both advance accountability for international crimes and to aid in their perpetration. Most of the existing literature to date focuses on the former, such as how digital evidence can be used in international criminal law (ICL) proceedings, or in the case of the latter, has taken a mainly rights-based approach (such as how technology can infringe upon rights to privacy or freedom of expression). This article answers the understudied question of how technology can serve as the vehicle by which certain international crimes are committed or lead to new offences, and how current ICL frameworks may be able to accommodate these ‘online harms’ to ensure that the law recognizes the full scope of harms caused to victims, who currently may not be able to access redress through the international criminal justice system. It identifies three examples of online harm that have a foreseeable nexus to the perpetration of international crimes, including hate speech and disinformation, sharing footage of crimes to the internet, and online sexual violence. The article evaluates the online harms alongside similar harms that have been encompassed by core ICL crimes, including genocide, crimes against humanity, and war crimes, to assess how online harms might fit into ICL frameworks (e.g. as an aggravating factor at sentencing, a new mode of commission of an existing crime, or a new crime). It concludes that some types of online harm may be more feasible to account for than others, and identifies where the existing ICL architecture falls short, which is important for providing a basis for future research as to how to best include novel online harms under ICL. Finally, the article emphasizes that as technology will only continue to develop and serve as a vehicle for an increasing array of harms, finding ways to account for online harm and bring redress to victims should be an issue at the forefront of ICL.

1. INTRODUCTION

The development of new technologies has led to a paradox for the fields of international law and human rights. On the one hand, the increased accessibility of smartphones and the internet has made it easier for civilians to collect evidence of human rights violations that may

then be used for accountability purposes.¹ On the other hand, new technologies can lead to new ways to infringe upon fundamental rights and to perpetrate international crimes. For instance, in the 2020 US Federal Elections, it was determined by the US National Intelligence Council that Russian intelligence used new technology to interfere with the election by amplifying negative content about President Biden and promoting disinformation about the coronavirus disease 2019 pandemic.² In addition, surveillance technologies, such as facial recognition tools, have been used by states to maintain power over their populations and are often programmed in ways that prejudice minority groups.³ Moreover, terrorist and other non-state armed groups have used social media to spread fear, propaganda, and to recruit new members.⁴

Much of the literature that focuses on the intersection between new technologies and human rights has concentrated on how technology can be used to enhance accountability (e.g. the use of open source information as evidence at international criminal trials),⁵ or, when focused on the negative impacts of technology, has taken a mainly rights-based approach (e.g. infringements upon privacy rights, racial or ethnic discrimination by algorithms, etc.).⁶ In contrast, there is little research on how technology enables novel forms of perpetration of existing international crimes or entirely new crimes, and how existing international criminal law (ICL) frameworks could be used for accountability purposes, despite the importance of being able to ensure that the law recognizes the full scope of harms caused to victims. Some scholars have written tangentially on related harms and international law, such as the international criminalization of online terrorist preparatory acts *via* UN Security Council resolutions or cybercrime conventions,⁷ the prohibition of hate propaganda under ICL,⁸ and who should be liable for potential international crimes committed by artificial intelligence.⁹ Yet, an analysis of how online harms might be addressed as crimes under the jurisdiction of international criminal courts and tribunals (ICCTs) remains missing. This article aims to bridge this gap in discussion around how harms perpetuated through online means ('online harms') may be accommodated under ICL frameworks, with a focus on crimes under the jurisdiction of the Rome Statute of the International Criminal Court (ICC).

Historically, ICL has been primarily concerned with physical violent crime.¹⁰ Progressively, ICL has recognized the importance of psychological forms of suffering (such

¹ See S. Dubberley, A. Koenig, and D. Murray (eds), *Digital Witness: Using Open Source Information for Human Rights Investigation, Documentation, and Accountability* (Oxford University Press (OUP), 2020).

² National Intelligence Council, *Foreign Threats to the 2020 US Federal Elections*, 10 March 2021, available online at <https://www.dni.gov/files/ODNI/documents/assessments/ICA-declass-16MAR21.pdf> (visited 12 April 2023).

³ M.K. Land and J.D. Aronson, 'Human Rights and Technology: New Challenges for Justice and Accountability', 16 *Annual Review of Law and Social Science* (2020) 223–240.

⁴ See A. Barr and A. Herfroy-Mischler, 'ISIL's Execution Videos: Audience Segmentation and Terrorist Communication in the Digital Age', 41 *Studies in Conflict & Terrorism* (2017) 946–967; E. Vermeersch et al., *The Role of Social Media in Mali and Its Relation to Violent Extremism: A Youth Perspective*, International Centre for Counter-Terrorism The Hague and United Nations Interregional Crime and Justice Research Institute, March 2020, available online at <https://www.icct.nl/sites/default/files/import/publication/Social-Media-in-Mali-and-Its-Relation-to-Violent-Extremism-A-Youth-Perspective.pdf> (visited 12 April 2023).

⁵ See L. Freeman, 'Digital Evidence and War Crimes Prosecutions: The Impact of Digital Technologies on International Criminal Investigations and Trials', 41 *Fordham International Law Journal* (Fordham Int'l L J) (2018) 283–336; A. Koenig, 'From 'Capture to Courtroom': Collaboration and the Digital Documentation of International Crimes in Ukraine', 20 *JICJ* (2022) 829–842.

⁶ See L. McGregor, D. Murray, and V. Ng, 'International Human Rights Law as a Framework for Algorithmic Accountability', 68 *International and Comparative Law Quarterly* (2019) 309–343; M. Risse, 'Human Rights and Artificial Intelligence: An Urgently Needed Agenda', 41 *Human Rights Quarterly* (2019) 1–16.

⁷ I. Couzigou, 'The Criminalization of Online Terrorism Preparatory Acts Under International Law', 45 *Studies in Conflict & Terrorism* (2019) 535–554.

⁸ W.K. Timmermann, 'The Relationship between Hate Propaganda and Incitement to Genocide: A New Trend in International Law Towards Criminalization of Hate Propaganda?' 18 *Leiden Journal of International Law* (2005) 257–282.

⁹ M. Swart, 'Constructing "Electronic Liability" for International Crimes: Transcending the Individual in International Criminal Law', 24 *German Law Journal* (2023) 589–602.

¹⁰ J. Werle and F. Jeßberger, *Principles of International Criminal Law* (4th edn., OUP, 2020), at 155.

as for the mental dimensions of torture and genocide), but this has always been an adjunct to cases that have focused on physical harm at their root. Until very recently, ICL did not prioritize any offences concerning other forms of harm, such as serious economic or environmental harm.¹¹ Yet, the proposed addition of new crimes such as ecocide under the Rome Statute of the ICC (hereinafter ‘Rome Statute’) demonstrates the capacity of ICL to evolve and accommodate novel (or at least previously unrepresented) types of harm.¹²

An emerging field of study is ‘online harm’,¹³ which encompasses a variety of acts or behaviours inflicted *via* online means, ranging from online bullying, to cyberstalking and harassment, to online sexual abuse and the sharing of violent or extremist content. The regulation and criminalization of online harm have been explored at the state level through proposed legislation such as the Online Safety Bill in the UK,¹⁴ and enacted in states such as Australia,¹⁵ Germany,¹⁶ and Indonesia.¹⁷ In the international law context, the need to counter online harms has been recognized through UN General Assembly Resolutions on disinformation,¹⁸ policy briefs by the Secretary-General concerning information integrity on digital platforms,¹⁹ and UN guidelines on countering and addressing online hate speech.²⁰ However, no known efforts have been made to utilize ICL to counter online harm, which could be a valuable addition to obtaining justice for victims of grave online harms who may not be able to receive redress through domestic criminal proceedings or regulatory processes.

In addition, existing literature on criminalizing online harm is situated within general criminal law theory and domestic criminal law spheres, such as in relation to image-based sexual abuse or pornographic ‘deepfakes’,²¹ cyberbullying,²² and false speech.²³ For the purposes of this paper, the term ‘online harm’ is used to refer to harmful acts perpetrated through online spaces. At the time of writing, the current literature does not link the concepts of ‘online harm’ and ICL. To fill this gap, this article brings together the themes of online harm and ICL through three examples of online harm that may have relevance or could

¹¹ *Ibid.* However, it is worth noting that scholars have argued that the doctrinal tools to enable focusing on such harms have been around for a while, see E. Schmid, *Taking Economic, Social and Cultural Rights Seriously in International Criminal Law* (Cambridge University Press (CUP), 2015).

¹² Stop Ecocide Foundation, *Independent Expert Panel for the Legal Definition of Ecocide: Commentary and Core Text*, June 2021, available online at <https://www.stopecocide.earth/expert-drafting-panel> (visited 12 April 2023). See also D. Bertram and G. Hill, ‘Visual Advocacy in the Ecocide Movement’ in this issue of the *Journal of International Criminal Justice*.

¹³ See J. Woodhouse, *Regulating Online Harms* (House of Commons Library, 2022); R.F. Jørgensen (ed.), *Human Rights in the Age of Platforms* (MIT Press, 2019); T. Keipi et al., *Online Hate and Harmful Content* (Routledge, 2017).

¹⁴ Online Safety Bill, HL Bill 164, available online at <https://bills.parliament.uk/publications/52368/documents/3841> (visited 26 March 2024).

¹⁵ Online Safety Bill 2021, available online at https://parlinfo.aph.gov.au/parlInfo/search/display/display.w3p;query=Id:22legislation/bills/r6680_aspassed/0000%22 (visited 26 March 2024).

¹⁶ Network Enforcement Act (*Netzwerkdurchsetzungsgesetz*) of 2017, available online at https://www.bmj.de/SharedDocs/Downloads/DE/Gesetzgebung/BGBl/BGBl_NetzDG.pdf?__blob=publicationFile&v=3 (visited 26 March 2024).

¹⁷ Electronic Information and Transactions, Law of the Republic of Indonesia No. 11 of 2008 and amended by No. 19 of 2016. Online harms such as child pornography and terrorist propaganda fall under many other domestic civil and criminal codes, but the examples mentioned are those with comprehensive laws specific to multiple online harms.

¹⁸ GA Res. 76/227, 24 December 2021.

¹⁹ United Nations, *Our Common Agenda Policy Brief 8: Information Integrity on Digital Platforms*, June 2023, available online at <https://www.un.org/sites/un2.un.org/files/our-common-agenda-policy-brief-information-integrity-en.pdf> (visited 26 March 2024).

²⁰ UN Office on Genocide Prevention in Collaboration with the ESRC Human Rights, Big Data and Technology Project, University of Essex, *Countering and Addressing Online Hate Speech: A Guide for policy makers and practitioners* (July 2023), available online at https://www.un.org/en/genocidprevention/documents/publications-and-resources/Countering_Online_Hate_Speech_Guide_policy_makers_practitioners_July_2023.pdf (visited 26 March 2024).

²¹ M. Goudsmit Samaritter, ‘What Makes a Sex Crime? A Fair Label for Image-based Sexual Abuse’, 2 *BJU Strafblad* (2021) 67; R.A. Delfino, ‘Pornographic Deepfakes: The Case for Federal Criminalization of Revenge Porn’s Next Tragic Act’, 88 *Fordham Law Review* (2019) 887–938.

²² L. Franco and K. Ghanayim, ‘The Criminalization of Cyberbullying Among Children and Youth’, 17 *Santa Clara Journal of International Law* (2019).

²³ L.W. Tompros et al., ‘The Constitutionality of Criminalizing False Speech Made on Social Networking Sites in a Post-Alvarez, Social Media-Obsessed World’, 31 *Harvard Journal of Law and Technology* (2017) 65.

potentially be encompassed under existing international crimes. These include the spread of hate speech and disinformation, the sharing of footage of crimes to social media, and online sexual violence. The article seeks to initiate a conversation about whether and to what extent online harms might fit into ICL frameworks (e.g. as an aggravating factor at sentencing, a new mode of commission of an existing crime, or as an entirely new crime).²⁴ Embarking on this exercise allows for similarities and differences between ‘old’ and ‘new’ forms of harm to come to light, which can help to more easily discern where online harm can be accommodated by existing law or otherwise where the gaps lie. The discussions that may be sparked by this article are relevant to the explorations of engaging with ICL ‘after critique’ presented in this Special Issue,²⁵ as it is important that ICL be able to recognize and adapt to new forms of harm to avoid the ‘privileging of existing criminal offences and harms’ that can reinforce traditional assumptions and stereotypes behind the law.²⁶ This is also in line with the concept of ‘kaleidoscopic justice’ noted by McGlynn in relation to survivors of sexual violence, which emphasizes need to embrace the ‘varied, nuanced, ever-changing experience and understandings of justice’.²⁷

Ultimately, this article concludes that some types of online harm may be easier to accommodate under ICL (as it stands) than others. For some online harms, there may be enough similarities to existing elements of crimes for them to be included within current ICL frameworks. In other cases, a completely new crime may need to be defined to fully account for the harm experienced by victims of the online act. Going forward, the conclusions reached in this article can provide a basis for future research as to how to best include newer online harms under ICL and in turn, to deliver more justice for victims.

2. ONLINE HARMS AND CORE INTERNATIONAL CRIMES

This section explores online harms which could conceivably be linked to core ICL crimes: the spread of hate speech and disinformation, sharing footage of crimes online, and sexualized abuse perpetrated in online spaces. These acts were chosen because of their close connection to core international crimes, as opposed to online harms such as cyberbullying or cyberstalking, which would likely be more relevant in a domestic criminal law sphere. The list of online harms examined in this article is not intended to be exhaustive, as other online harms could still have plausible links to ICL and will likely continue to arise in the future as technology evolves. Instead, this article surveys some key online harms that are currently being perpetuated with a nexus to core international crimes to demonstrate how a theory for how such novel harms may be accommodated by ICL frameworks can be developed.

A. Hate Speech and Disinformation

A first example of online harm that has relevance for international crimes is the spread of hate speech and disinformation. Hate speech has no universal definition under international law, though it may generally be thought of as speech that targets a group or an individual based on inherent characteristics (such as race, religion, or gender) and that may threaten peace.²⁸ Disinformation refers to false information that is intentionally spread, as opposed to

²⁴ This article limits its analysis to the crimes of genocide, crimes against humanity, and war crimes, and therefore does not discuss the crime of aggression that exists under the Rome Statute. It also does not discuss cyber warfare.

²⁵ M. Burgis-Kasthala and B. Sander, ‘Contemporary International Criminal Law after Critique: Towards Decolonial and Abolitionist (Dis-)Engagement in an Era of Anti-impunity’, in this issue of the *Journal*.

²⁶ C. McGlynn, ‘Challenging Anti-carceral Feminism: Criminalisation, Justice and Continuum Thinking’, 93 *Women’s Studies International Forum* (2022) 102614, 3.

²⁷ *Ibid.*; Burgis-Kasthala and Sander, *supra* note 25.

²⁸ United Nations, *Strategy and Plan of Action on Hate Speech*, 2019, available online at https://www.un.org/en/genocide-prevention/documents/advising-and-mobilizing/Action_plan_on_hate_speech_EN.pdf (visited 12 April 2023), at 2.

misinformation, which is false information spread unknowingly or by mistake.²⁹ Concerning ICL, disinformation has been linked to the commission of atrocity crimes at different stages of perpetration, including prompting such crimes, enabling their continuation as they occur, and concealing or denying them afterwards.³⁰ In the age of social media, hate speech and disinformation (which can include hateful false information or propaganda about a particular group) can be spread to a large audience with the click of a button.

Online hate speech and disinformation could constitute the crimes of persecution and (incitement to) genocide.³¹ First, in ICL jurisprudence, offline hate speech has been found to constitute an act of persecution.³² This occurred at the International Criminal Tribunal for Rwanda (ICTR), in which the Trial Chambers for both the *Ruggiu* and *Media* cases held that hate speech (regardless of whether it included calls for violence against the group) could constitute persecution.³³ In *Ruggiu*, the Trial Chamber noted that this was because the public radio broadcasts were ‘aimed at singling out and attacking’ the groups ‘on discriminatory grounds, by depriving them of the fundamental right to liberty and basic humanity enjoyed by members of wider society’.³⁴ Furthermore, the Chamber found that ‘[t]he deprivation of [those] rights can be said to have as its aim the death and removal of those persons from the society in which they live alongside the perpetrators, or eventually even from humanity itself’.³⁵ In the *Media* case, the Trial Chamber found that because the crime of persecution is defined in terms of impact as opposed to intent, ‘[i]t is not a provocation to cause harm’ and is ‘itself the harm’.³⁶ It is for this reason that, in the view of the Trial Chamber, ‘there need not be a call to action in communications that constitute persecution’.³⁷ Yet, in the same case, the Appeals Chamber stated that hate speech in combination with calls for violence could constitute persecution, but was uncertain as to whether ‘mere’ hate speech without calls for violence could amount to persecution (and ultimately decided that it was not relevant to answer this question).³⁸

On this note, an obstacle for prosecuting crimes before certain ICCTs — most notably the ICC — is the gravity requirement, which refers to the focus on only the ‘most serious crimes of concern to the international community’, as provided in the Rome Statute.³⁹ There seems to be uncertainty amongst ICCTs as to whether hate speech that does not explicitly call for physical violence meets the gravity requirement for the crime against humanity of persecution.⁴⁰ For example, while the ICTR in the *Media* case ruled that hate speech can meet the threshold to constitute persecution,⁴¹ the International Criminal Tribunal for the former Yugoslavia’s (ICTY) Trial Chamber in *Kordic and Čerkez* found that hate speech

²⁹ A.M. Guess and B.A. Lyons, ‘Misinformation, Disinformation, and Online Propaganda’, in N. Persily and J.A. Tucker (eds), *Social Media and Democracy: The State of the Field, Prospects for Reform* (CUP, 2020), at 11.

³⁰ M. Holvoet, ‘International Criminal Liability for Spreading Disinformation in the Context of Mass Atrocity’, 20 *IJCJ* (2022) 223–250, at 248.

³¹ Arts 7(1)(h) and 25(3)(e) ICCSt.

³² R. Kapoor and S. Aravindakshan, ‘Hate Speech as Persecution: Tackling the Gordian Knot’, EJIL: Talk!, 12 August 2020, available online at <https://www.ejiltalk.org/hate-speech-as-persecution-tackling-the-gordian-knot/> (visited 12 April 2023).

³³ Judgment, *Ruggiu* (ICTR-97-32-I), Trial Chamber I, 1 June 2000 (hereinafter referred to as ‘*Ruggiu* Trial Judgment’), § 22; Judgment, *Nahimana, Barayagwiza and Ngeze* (ICTR-99-52-T), Trial Chamber, 3 December 2003 (hereinafter referred to as ‘*Nahimana* Trial Judgment’), § 1073.

³⁴ *Ruggiu* Trial Judgment, *ibid.*, § 22.

³⁵ *Ibid.*

³⁶ *Nahimana* Trial Judgment, *supra* note 33, § 1073.

³⁷ *Ibid.*

³⁸ Judgment, *Nahimana, Barayagwiza and Ngeze* (ICTR-99-52-A), Appeals Chamber, 11 August 2008 (hereinafter referred to as ‘*Nahimana* Appeals Judgment’), § 987.

³⁹ Preamble, ICCSt. For a comprehensive overview of the gravity requirement, see M.M. deGuzman, ‘Gravity and the Legitimacy of the International Criminal Court’, 32 *Fordham Int’l L J* (2008) 1400–1465.

⁴⁰ G. S. Gordon, ‘Hate Speech and Persecution: A Contextual Approach’, 46 *Vanderbilt Journal of Transnational Law* (2013) 303–373.

⁴¹ *Nahimana* Trial Judgment, *supra* note 33, *Nahimana* Appeals Judgment, *supra* note 38.

without any direct call for violence does not meet the gravity requirement to be considered a crime against humanity.⁴² This uncertainty between ICL Chambers signals that in order for online hate speech to potentially constitute an act of persecution at a court such as the ICC, it may need to involve direct calls for violence.

Secondly, regarding incitement to genocide as its own offense under Article 25(3)(e) of the Rome Statute, this type of speech would need to be ‘direct’ and ‘public’ in its call for genocide against a protected group.⁴³ Under the Rome Statute, genocide is the only international crime in which incitement is expressly prohibited,⁴⁴ and it constitutes an inchoate offense, meaning the incitement is a crime regardless of whether a genocide subsequently occurs.⁴⁵ Historically, individuals have already been prosecuted whose hateful words directly incited acts of genocide to be committed against members of a targeted group. For example, at the Nuremberg Tribunal, Streicher and Fritzsche were tried, in part, for hate propaganda, as Streicher’s genocidal rants were published in a weekly anti-Semitic newspaper Fritzsche served as the head of the Radio Division of the Propaganda Ministry.⁴⁶ Additionally, at the ICTR, Akayesu was convicted of incitement to commit genocide for speeches urging a crowd to eliminate the Tutsis, and three men in the *Media* case were prosecuted for similar hate speech made in newspapers and on the radio.⁴⁷

A recent and ongoing example of the harm of online hate speech and disinformation is the suspected genocide in Myanmar against the Rohingya population. Social media websites, specifically Facebook (now under the parent company, Meta), have purportedly become tools used by the Myanmar military and its supporters to disseminate hateful and false messages about the Muslim Rohingya group.⁴⁸ For instance, Myanmar military personnel have allegedly posed as civilians on Facebook and posted various forms of hateful speech, such as that ‘Islam is a global threat to Buddhism’ or a false story about the rape of a Buddhist woman by a Muslim man.⁴⁹ As found by the Independent International Fact-Finding Mission on Myanmar, these posts are part of a systematic strategy used by the Myanmar military on Facebook to target the Rohingya population, allowing Facebook to serve as an ‘echo chamber of virulent anti-Rohingya content’ that ‘pour[s] fuel on the fire of long-standing discrimination and substantially increase[s] the risk of an outbreak of mass violence’.⁵⁰ The *New York Times* reported that as described by anonymous whistle-blowers, hundreds of military personnel created ‘troll accounts’ and fake news and celebrity pages on Facebook, then ‘flooded them with incendiary comments and posts timed for peak viewership’.⁵¹ This online

⁴² Judgement, *Kordic and Čerkez* (IT-95-14/2-T), Trial Chamber, 26 February 2001, § 209.

⁴³ Art. 25(3)(e) ICCSt.

⁴⁴ C. Stahn, *A Critical Introduction to International Criminal Law* (CUP, 2018), at 47.

⁴⁵ W.K. Timmermann, ‘Incitement in International Criminal Law’, 88 *International Review of the Red Cross* (2006) 823.

⁴⁶ Timmermann, *supra* note 8, 260–261; G.S. Gordon, ‘The Propaganda Prosecutions at Nuremberg: The Origin of Atrocity Speech Law and the Touchstone for Normative Evolution’, 39 *Loyola International & Comparative Law Review* (2017) 209–245, at 217.

⁴⁷ *Ibid.*, 266.

⁴⁸ OHCHR, *Report of the Independent International Fact-Finding Mission on Myanmar*, A/HRC/42/50, 16 September 2019; E. Irving, ‘“The Role of Facebook is Significant”: Facebook and the Fact Finding Mission on Myanmar’, *Opinio Juris*, 7 September 2018, available online at <http://opiniojuris.org/2018/09/07/the-role-of-social-media-is-significant-facebook-and-the-fact-finding-mission-on-myanmar/> (visited 12 April 2023); P. Mozur, ‘A Genocide Incited on Facebook, With Posts from Myanmar’s Military’, *The New York Times*, 15 October 2018, available online at <https://www.nytimes.com/2018/10/15/tech-nology/myanmar-facebook-genocide.html> (visited 12 April 2023). It is also worth noting that there is a growing body of literature focused on whether social media actors could be held responsible for the spread of hate speech under international law, see S.R. Singh, ‘Move Fast and Freak Societies: The Weaponization of Social Media and Options for Accountability under International Criminal Law’, 8 *Cambridge International Law Journal* (2019) 331–342; J. Domino, ‘Crime as Cognitive Constraint: Facebook’s Role in Myanmar’s Incitement Landscape and the Promise of International Tort Liability’, 52 *Case Western Reserve Journal of International Law* (2020) 143–198.

⁴⁹ *Ibid.*

⁵⁰ *Ibid.*; Amnesty International, *The Social Atrocity: Meta and the Right to Remedy for the Rohingya*, September 2022, available online at <https://www.amnesty.org/en/documents/ASA16/5933/2022/en/> (visited 12 April 2023), at 6.

⁵¹ Mozur, *supra* note 48.

hate speech has been argued to have incited murders, rapes, and ‘the largest forced migration in human history’,⁵² all of which have contributed to the alleged genocide in Myanmar (which is currently under investigation by the ICC). This practice of spreading hate speech has even been confirmed by Facebook’s parent company, Meta, whose Head of Security Policy stated that they had found ‘clear and deliberate attempts to covertly spread propaganda that were directly linked to the Myanmar military’.⁵³

What sets the harm of online hate speech and disinformation apart from offline speech of a similar nature is its ability to spread to much wider audiences in a very short timeframe. For example, while shouting something hateful about a minority group in a public town square might be heard by those in the immediate vicinity, posting the same comment on a website can allow for anyone around the world to see it and potentially act upon that message. In the context of the Rohingya situation, Facebook is many times equated with the internet itself, with an estimated 20 million users in Myanmar as of 2018 that is ‘growing exponentially’.⁵⁴ As reported by Amnesty International following interviews conducted in Myanmar, it is also apparent that the rapid transformation of the country’s telecommunications landscape and the rise of Facebook meant that digital media literacy is still fairly low, and the combination of Facebook being an American platform ‘populated by trusted friends and family’ and also used by the government for official announcements contributes to users’ perceptions of the social media website as a reliable source of information.⁵⁵ Amnesty International further observed that Facebook ‘was used not only as a means of communicating with friends, but for many people in Myanmar, it became their primary news source, business directory, online marketplace, and go-to search engine’.⁵⁶ Therefore, when anti-Rohingya groups and the government itself allegedly used Facebook to spread hateful messages and incite violence, they were able to influence an extremely wide audience in a short span of time, which would not have been possible without the online element.

In this sense, a comparison can be drawn with the Rwandan genocide, in which the hate speech spread on the Hutu radio stations about the Tutsi population were used to incite Hutu to ‘hate and kill the Tutsi’.⁵⁷ In the *Media* case at the ICTR, two defendants were found guilty of incitement to genocide for broadcasting hateful speech on their radio station, Radio Télévision Libre des Mille Collines (RTLM), and the third for speech published in a newspaper.⁵⁸ The two defendants charged in relation to the radio station RTLM, which was very popular in Rwanda at the time,⁵⁹ were found to have perpetuated ethnic stereotyping against the Tutsis through their broadcasts and deliberately called for violence against the group.⁶⁰ Just as hate speech and other propaganda spread through the radio could be classified as incitement to genocide in Rwanda, it is possible and even probable that the same type of speech directed against a group (such as the Rohingya) on social media could also meet this threshold. Online hate speech can be especially severe given that the internet can reach a much wider audience in quicker and more extended amounts of time (as posts remain visible on the internet after being uploaded, unlike speech given on live radio broadcasts).

⁵² *Ibid.*

⁵³ *Ibid.*

⁵⁴ Amnesty International, *supra* note 50, at 17.

⁵⁵ *Ibid.*

⁵⁶ *Ibid.*; OHCHR, *supra* note 48, § 1345.

⁵⁷ A. Hefti and L. Ausserladscheider Jonas, ‘From Hate Speech to Incitement to Genocide: The Role of the Media in the Rwandan Genocide’, 38 *Boston University International Law Journal* (2020) 15.

⁵⁸ *Nahimana Trial Judgment*, *supra* note 33; The third defendant, Barayagwiza, was the government minister who provided the political framework for the killings.

⁵⁹ *Ibid.*, § 342.

⁶⁰ *Ibid.*, § 371.

Though there has not yet been an international criminal trial in which *online* hate speech or disinformation has been considered, it is coming. In its request for authorization of an investigation into the situation of Bangladesh and Myanmar in 2019, the ICC Office of the Prosecutor (OTP) referred to numerous instances of hate speech attacking the Rohingya and other Muslims posted to Facebook.⁶¹ The case is still currently in the investigation phase, though it is foreseeable that should warrants of arrest be issued and a trial occur, that the ICC will consider the harms of online hate speech in the context of the alleged genocide depending on which crimes the Prosecutor decides to focus on. However, a key challenge which would likely emerge should hate speech and disinformation on social media websites be a main element of the OTP's case, would be the acquisition of evidence. As has already been an issue at proceedings for the same situation at the International Court of Justice (*The Gambia v. Myanmar*), Meta will likely refuse to disclose information to an international court without a legal obligation to do so, as shown by its limited cooperation to provide evidence that led The Gambia to bring a case in relation to the handover of information against it in a United States domestic court.⁶² Thus, should cases considering online hate speech or disinformation be heard at the ICC, which cannot compel a company to cooperate with it, there could be barriers to accessing the original posts containing the speech. In these instances, the ICC would need to take other measures, such as asking a state party to request the content from social media companies.

Additionally, another challenge which could emerge should online hate speech and disinformation become a crime under ICL is that of jurisdiction. In some cases with very specific contexts, such as government officials in Myanmar posting hateful speech on Facebook, it may be fairly straightforward to determine whether a court such as the ICC has jurisdiction, as some of the alleged crimes took place in Bangladesh (a state party to the ICC), and the government officials posting are in the area and involved with the conflict. However, in other cases where online hate speech is perpetrated from multiple states (some of which might not even be known, as users can be anonymous and hide their locations), it would be difficult to determine a jurisdictional or legal basis for prosecution.⁶³

B. Sharing Footage of Crimes Online

In recent years, non-state armed groups have accentuated their offline crimes either by posting footage of physical torture and killings online to further humiliate the victims and their families, or by forcing individuals (including children) to watch execution videos to instil fear.⁶⁴ This online harm encompasses three different scenarios, each with different effects.

First, the element of recording and/or sharing a video of someone being tortured causes additional humiliation and degradation to the 'direct' victim. This harm was alluded to by the OTP in its application for an arrest warrant in the case of *Al Hassan* at the ICC, which argued that the public nature of punishments by the non-state armed terrorist group Ansar Dine, with whom the accused was affiliated, who also posted videos of them on social media,

⁶¹ Request for authorization of an investigation pursuant to Art. 15, *Situation in the People's Republic of Bangladesh/Republic of the Union of Myanmar* (ICC-01/19), Office of the Prosecutor, 4 July 2019, §§ 59–60.

⁶² P. Pillai, 'The Republic of The Gambia v Facebook, Inc: Domestic Proceedings, International Implications' *Opinio Juris*, 8 August 2020, available online at <http://opiniojuris.org/2020/08/08/the-republic-of-the-gambia-v-facebook-inc-domestic-proceedings-international-implications/> (visited 12 April 2023); A. Koenig, 'Q&A on Court Ordering Facebook to Disclose Content on Myanmar Genocide', *Just Security*, 24 September 2021, available online at <https://www.justsecurity.org/78358/qa-on-court-ordering-facebook-to-disclose-content-on-myanmar-genocide/> (visited 12 April 2023).

⁶³ For a relevant discussion of jurisdictional challenges faced when prosecuting transnational cyber offences, see A. Perloff-Giles, 'Transnational Cyber Offenses: Overcoming Jurisdictional Challenges', 43 *Yale Journal of International Law* (2018) 191.

⁶⁴ See UN Human Rights Council, *Report of the Independent International Commission of Inquiry on the Syrian Arab Republic*, 5 February 2015, available online at <https://www.ohchr.org/en/hr-bodies/hrc/iici-syria/documentation> (visited 12 April 2023), which details multiple incidents of this nature committed by ISIS in Syria.

was particularly humiliating for the victims.⁶⁵ The OTP also submitted that the videos posted to YouTube by Ansar Dine depicted the victims ‘bent over in pain’ and with ‘private parts exposed for all to see’, which contributed to ‘a sense of public humiliation’.⁶⁶ A similar issue was also held by the Pre-Trial Chamber in the case of *Al-Werfalli*, where execution videos were posted on social media.⁶⁷ The Pre-Trial Chamber stated in its two arrest warrants that ‘the posting on social media of the videos depicting executions’ and ‘the manner in which the crime was committed and publicized was cruel, dehumanizing, and degrading’.⁶⁸ However, the Court never addressed this issue, as the case did not move forward past the arrest warrant stage due to *Al-Werfalli*’s death.

The posting of videos of torture and executions online further degrades victims before a large audience, and can intensify their suffering, as once something is shared online, it can be extremely difficult to remove and stop its circulation —⁶⁹ an act which is usually a singular instance fixed in time is different to one that is continuous and constantly available to taunt victims and their communities. As Emma Irving and I have argued, in the Chamber’s initial gravity assessment when deciding to issue an arrest warrant in *Al Hassan*, more focus should have been placed on the additional harm suffered by victims by having their suffering posted to public online platforms, which ‘extended the audience of this suffering far beyond the group of onlookers on the street that day’ and likely heightened victims’ sense of humiliation and degradation immensely.⁷⁰ This is because the posting of videos of crimes by Ansar Dine itself demonstrated a desire to further degrade victims before a widespread audience, which we argue should have been considered by the Pre-Trial Chamber as heightening the seriousness of the offense.⁷¹ This analysis is consistent with the ICTY Trial Chamber’s findings in *Kvočka* — that ‘[t]he presence of onlookers, particularly family members, also inflicts severe mental harm amounting to torture’.⁷²

The harms experienced by victims when footage of crimes is shared online are seemingly similar in their nature of severe degradation to the harms encompassed by torture,⁷³ ‘other inhumane acts’,⁷⁴ and the war crimes of ‘inhuman treatment’,⁷⁵ ‘wilfully causing great suffering’,⁷⁶ and ‘outrages upon personal dignity, in particular humiliating and degrading treatment’.⁷⁷ This is because these crimes have the potential to address instances in which either a victim suffers severe mental harm and a loss of dignity from the presence of onlookers, or witnesses to a violent act suffer mental distress. In relation to war crimes, though these crimes have not been prosecuted yet at ICCTs for online acts, domestic criminal courts have prosecuted them primarily in relation to open source photos and videos. For instance, in The Netherlands, a man was convicted for degrading treatment and an assault on personal dignity for posting a photo of himself with a crucified individual in Syria to Facebook.⁷⁸ In this case, the Hague Court found:

⁶⁵ Requête urgente du Bureau du Procureur, *Al Hassan* (ICC-01/12-01/18), La Chambre Préliminaire I, 20 mars 2018, § 136.

⁶⁶ *Ibid.*, § 304.

⁶⁷ Warrant of Arrest, *Al-Werfalli* (ICC-01/11-01/17), Pre-Trial Chamber, 15 August 2017, § 29; Second Warrant of Arrest, *Al-Werfalli* (ICC-01/11-01/17), Pre-Trial Chamber, 4 July 2018, § 31.

⁶⁸ *Ibid.*

⁶⁹ S. Zarnsky and E. Irving, ‘Rewrite of Decision on the Prosecutor’s Application for the Issuance of a Warrant of Arrest for *Al Hassan Ag Abdoul Aziz Ag Mohamed Ag Mahmoud*’, in K. McLoughlin et al. (eds), *Feminist Judgments: Reimagining the International Criminal Court* (CUP, forthcoming 2024), available on SSRN at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4765275

⁷⁰ *Ibid.*

⁷¹ *Ibid.*

⁷² *Ibid.*; Judgment, *Kvočka et al.* (IT-98-30/1-T), Trial Chamber, 2 November 2001, § 149.

⁷³ Art. 7(1)(f), Art. 8(2)(a)(ii), and Art. 8(2)(c)(i) ICCSt.

⁷⁴ Art. 7(1)(k) ICCSt.

⁷⁵ Art. 8(2)(a)(ii) ICCSt.

⁷⁶ Art. 8(2)(a)(iii) ICCSt.

⁷⁷ Art. 8(2)(b)(xxi) and Art. 8(2)(c)(ii) ICCSt.

⁷⁸ ECLI:NL:GHDHA:2021:103.

By posing next to the deceased person and having a photo taken, the suspect has contributed to further deepening the humiliation and/or dishonour of the deceased. The suspect has expressed that the body of the deceased should be seen as a trophy and that he is superior to the deceased. This humiliating and/or degrading behaviour is of such gravity that it is simply seen as an assault on the personal dignity of the deceased person. By subsequently posting the photo on his Facebook account, the suspect ensured that a large number of people were given the opportunity to take note of the photo.⁷⁹

In this ruling, the Hague Court considered what the ICC Pre-Trial Chamber arguably overlooked in *Al-Werfalli* and *Al Hassan* — that the posting of footage of victims in such a violent and vulnerable state onto social media was not just evidence of the commission of other crimes, but an act worthy of being a crime in itself. In doing so, the Hague Court was able to deliver more justice to the victim and his family, ensuring that the additional harm the defendant knowingly inflicted when choosing to share the image online was properly acknowledged.

At the same court, another man was convicted of the same crimes for posting a video of himself kicking and spitting on bodies of the dead in Syria to YouTube.⁸⁰ The Hague Court found in this case that uploading the video to YouTube ‘contin[ed] the humiliation and dishonour by allowing a large audience to learn about it’.⁸¹ In this case, the Hague Court referenced the *Kunarac* case at the ICTY, which defined an assault on personal dignity as being when ‘the accused intentionally committed or participated in an act or omission which would be generally considered to cause serious humiliation, degradation or otherwise be a serious attack on human dignity’ — a definition that was, in the Court’s view, met by posting a video desecrating dead bodies to YouTube.⁸² Additionally, the Hague Court recalled Article 13 of Geneva Convention III concerning the human treatment of prisoners, which states that prisoners must be protected ‘against insults and public curiosity’.⁸³ The Court noted the 2020 ICRC Commentary on Article 13, which states:

In modern conflicts, the prohibition also covers, subject to the considerations discussed below, the disclosure of photographic and video images, recordings of interrogations or private conversations or personal correspondence or any other private data, irrespective of which public communication channel is used, including the internet. Although this is seemingly different from being marched through a hostile crowd, such disclosure could still be humiliating and jeopardize the safety of the prisoners’ families and of the prisoners themselves once they are released.⁸⁴

The Hague Court applied the reasoning described in the ICRC Commentary to demonstrate that the acts which violate the prohibition of curiosity, first written in 1949, have changed to encompass modern technological advancements. It is thus possible that, in a similar vein, the acts which violate other international prohibitions and constitute crimes may evolve to accommodate other online harms.

In addition to the harm experienced by the immediate victims, once footage of torture is recorded and shared, those close to them and other internet users who stumble across the

⁷⁹ *Ibid*, translated from Dutch.

⁸⁰ ECLI:NL:RBDHA:2021:3998.

⁸¹ *Ibid*, translated from Dutch.

⁸² *Ibid*; Judgment, *Kunarac et al.* (IT-96-23-T & IT-96-23/1-T), Trial Chamber, 22 February 2001, § 514.

⁸³ *Ibid*; Art. 13 Geneva Convention III.

⁸⁴ ICRC, *Convention (III) relative to the Treatment of Prisoners of War. Geneva, 12 August 1949: Commentary of 2020*, Article 13: Humane Treatment of Prisoners, § 1624, available online at <https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/Comment.xsp?action=openDocument&documentId=3DEA78BSA19414AFC1258585004344BD> (visited 12 April 2023).

content can experience serious mental harm.⁸⁵ ICL jurisprudence has recognized the ‘serious mental harm’ that can come to third parties, particularly family or friends of the victim, when witnessing violent acts committed against others.⁸⁶ On this note, ICL jurisprudence has found that viewing torture of another person can be enough to constitute torture itself. For example, in the *Furundžija* case, the ICTY Trial Chamber found that being forced to watch sexual attacks on a female acquaintance met the severe mental suffering threshold needed to be considered as an act of torture.⁸⁷ The phrasing of the Trial Chamber (‘he was forced to watch sexual attacks on a woman, in particular, a woman whom he knew as a friend’) indicates that the harm stemmed from watching the assault (and not just because he was forced to do so).⁸⁸ Though this offense occurred in person and not through a computer screen, the harms can still be connected between the offline and online situations because viewing extremely violent content has been shown to cause the same types of effects as when individuals witness real-life violence, such as anxiety, depression, post-traumatic stress disorder (PTSD), or alcohol consumption.⁸⁹ These effects are very similar to the kinds of psychological issues faced by victims of in-person torture, which are known to also include frequent instances of depression and PTSD.⁹⁰ It should therefore not be assumed that the means of viewing the torture, whether it be in real life or through a screen, lessens the harm caused.

However, even if the harms of viewing violent acts against someone (and particularly a family member) in person versus online may be comparable, a major obstacle in considering this ‘online harm’ as torture constituting a crime against humanity is the requirement under Article 7(2)(e) of the Rome Statute that the victim be in the custody or under the control of the perpetrator. In this case, a victim’s family member simply viewing the footage from a remote location voluntarily (or even stumbling across it) without being under the immediate control of the perpetrator may not amount to torture, regardless of the level of harm caused. In these instances, it would be easier to envision such online harms as falling under other inhumane acts, wilfully causing great suffering, cruel or inhuman treatment, or outrages upon personal dignity, which do not possess custody requirements, unless the family members were physically forced to watch the footage on a screen. However, torture as a war crime under Article 8 (both for non-international and international armed conflicts) does not contain a custody requirement.⁹¹ It is thus possible that should the act take place in the context of an armed conflict (such as in a situation where a party to the conflict films footage of crimes and shares it either publicly on social media or by digitally sending it to those close to the victim), sharing this footage with the victim’s family members could potentially constitute a form of torture, even if the family members were not in the custody of the perpetrator.

⁸⁵ The psychological effects of viewing digital violent or graphic content have been studied in a variety of contexts, including (but not limited to) human rights, journalism, criminal justice, racial violence, and terrorism. See S. Dubberley, E. Griffin, and H.M. Bal, *Making Secondary Trauma a Primary Issue: A Study of Eyewitness Media and Vicarious Trauma on the Digital Frontline*, Eyewitness Media Hub, November 2015, available online at https://firstdraftnews.org/wp-content/uploads/2018/03/trauma_report.pdf; A. Birze, K. Regehr, and C. Regehr, ‘Workplace Trauma in a Digital Age: The Impact of Video Evidence of Violent Crime on Criminal Justice Professionals’, 38 *Journal of Interpersonal Violence* (2023) 1654–1689; S. Williams, ‘Stream of Sadness: Young Black Women’s Racial Trauma, Police Brutality and Social Media’, 21 *Feminist Media Studies* (2021) 1270–1284; M. Crayne and N. Shortland, *Vicarious Trauma via the Observation of Extremist Atrocities: A Rapid Evidence Assessment*, NCITE Reports, Projects, and Research, 2023, available online at <https://digitalcommons.unomaha.edu/cgi/viewcontent.cgi?article=1001&context=ncitereportsresearch> (websites visited 20 March 2024).

⁸⁶ Judgment, *Fofana and Kondewa* (SCSL-04-14-T), Trial Chamber I, 2 August 2007, § 153.

⁸⁷ Judgment, *Furundžija* (IT-95-17/1-T), Trial Chamber, 10 December 1998 (hereinafter referred to as the ‘*Furundžija* Trial Judgment’), § 267(ii).

⁸⁸ *Ibid.*

⁸⁹ A. Feinstein, B. Audet, and E. Waknine, ‘Witnessing Images of Extreme Violence: A Psychological Study of Journalists in the Newsroom’, 5 *JRSM Open* (2014).

⁹⁰ M. Basoglu et al., ‘Psychological Effects of Torture: A Comparison of Tortured With Nontortured Political Activists in Turkey’, 152 *American Journal of Psychiatry* (1994) 76–81.

⁹¹ Art. 8(2)(a)(ii-1) and Art. 8(2)(c)(i) Elements of Crimes.

A third potential effect of footage of torture and killings being recorded and shared relates to those who are targeted as recruits by terrorist organizations, particularly children. Generally, viewing violent and graphic content is known to cause emotional distress, which can result in vicarious psychological trauma to the viewer.⁹² More specifically, when that violent content is tailored to and strategically distributed to a vulnerable population, such as children, who may not be as resilient as adults,⁹³ the effects can be harmful.

Furthermore, terrorist groups have increasingly used social media platforms to target young audiences, using techniques such as ‘targeted advertising’ to those most vulnerable to propaganda and ‘grooming’, when the perpetrator learns about the child’s interests in order to tailor the approach and build a trusting relationship.⁹⁴ Thus, the harm of posting footage of a crime online has the potential to set off a cascade of more harm, as it could potentially be framed in a certain way or shared intentionally with a vulnerable audience in order to recruit more individuals and perpetuate more crimes (such as the recruitment of children to engage them in hostilities).

Overall, a likely challenge with prosecuting this type of online harm at an international criminal court or tribunal (or even in a domestic or regional court) would be the removal of evidence by social media platforms. This is because most companies (rightly) have content moderation policies or ‘community standards’ that prohibit users from posting violent or graphic content,⁹⁵ but this has also led to platforms removing content that could serve as important evidence for the prosecution of international crimes.⁹⁶ Once content is removed from platforms, there is a lack of transparency from companies about what then happens to the content and how it is preserved (if at all), as it is nearly impossible to find information about this in the public sphere. It is also likely that because of content removal policies that require violent or graphic footage to be taken down quickly, an investigative authority may not learn of its existence in the first place.⁹⁷ Thus, once footage of a crime is shared to social media and is removed, valuable evidence could be lost. The challenge that arises from this is that if online harm is to be prosecuted — including at the ICC — means of cooperation with social media companies will need to be explored and established in order to ensure that evidence is preserved for trial purposes and handed over to the investigators.

⁹² See C. Newton, ‘The Trauma Floor: The Secret Lives of Facebook Moderators in America’, *The Verge*, 25 February 2019, available online at <https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-inter-views-trauma-working-conditions-arizona> (visited 12 April 2023).

⁹³ M. Stoilova, S. Livingstone, and R. Khazbak, *Investigating Risks and Opportunities for Children in a Digital World: A Rapid Review of the Evidence on Children’s Internet Use and Outcomes*, UNICEF Innocenti Discussion Paper 2020-03, February 2021, available online at <https://www.end-violence.org/sites/default/files/paragraphs/download/Investigating-Risks-and-Opportunities-for-Children-in-a-Digital-World.pdf> (visited 12 April 2023).

⁹⁴ UNODC, *Handbook on Children Recruited and Exploited by Terrorist and Violent Extremist Groups: The Role of the Justice System*, 2017, available online at https://www.unodc.org/documents/justice-and-prison-reform/Child-Victims/Handbook_on_Children_Recruited_and_Exploited_by_Terrorist_and_Violent_Extremist_Groups_the_Role_of_the_Justice_System.E.pdf (visited 12 April 2023), at 13–14.

⁹⁵ See for example Meta’s policy on violent and graphic content, available online at <https://transparency.fb.com/en-gb/policies/community-standards/violent-graphic-content/>, or X’s (formerly Twitter’s) sensitive media policy, available online at <https://help.twitter.com/en/rules-and-policies/media-policy> (visited 16 August 2023).

⁹⁶ Human Rights Watch, ‘*Video Unavailable*’: *Social Media Platforms Remove Evidence of War Crimes*, 10 September 2020, available online at <https://www.hrw.org/report/2020/09/10/video-unavailable/social-media-platforms-remove-evidence-war-crimes> (visited 12 April 2023); Human Rights Center, UC Berkeley School of Law, *Digital Lockers: Archiving Social Media Evidence of Atrocity Crimes*, 2021, available online at https://humanrights.berkeley.edu/sites/default/files/digital_lockers_report5.pdf (visited 12 April 2023).

⁹⁷ However, it is worth noting that some NGOs, such as the Syrian Archive, work to quickly preserve online material before it is removed and lost. See M. Burgis-Kasthala, ‘Assembling Atrocity Archives for Syria: Assessing the Work of the CIJA and the IIMM’, 19 *JICJ* (2021) 1193–1220.

C. Online Sexual Violence

‘Sexual violence’ is an umbrella term which encompasses any unwanted sexual act or activity, from verbal harassment, to sexual abuse, to forced marriage or forced pregnancy.⁹⁸ In the context of online harms, sexual violence is closely linked to personal dignity, as the physical harms that can result from the sexual violence are not typically present given the nature of the digital space. Parallels can be drawn between online sexual violence and ICL cases dealing with public rapes and sexual assaults. For example, in *Furundžija*, the ICTY Trial Chamber found that the fact that victims were raped and sexually assaulted in front of soldiers who were watching and laughing caused ‘severe physical and mental pain, along with public humiliation’ and thus amounted to outrages upon personal dignity and sexual integrity.⁹⁹ Additionally, in the *Akayesu* case, the ICTR Trial Chamber found that forcing a female student to undress and do gymnastics naked in a public courtyard in front of a crowd also constituted the amount of seriousness to be considered sexual violence.¹⁰⁰ The idea of being naked or sexualized in front of a wider audience therefore seems to increase the severity of the act in the eyes of international criminal tribunals due to the additional harm it causes in terms of humiliation to the victim.

This same harm is present in many cases of sexual violence perpetrated through online environments, for instance in relation to image-based sexual abuse, or ‘revenge porn’.¹⁰¹ In some cases, intimate images or live-streams of individuals are posted online without permission from the person depicted, and are extremely difficult to remove from the internet once they have been uploaded, resulting in ‘immeasurable harm’ to victims.¹⁰² It is also possible that sexual imagery of a person is created without consent using AI technology (or ‘deepfakes’).¹⁰³ Both of these acts can also be perpetrated towards children, which can cause an irreparable amount of psychological trauma.¹⁰⁴ In addition, image-based sexual abuse subjects victims to the possibility of other types of harm, such as cyberbullying from other users who see the photos.¹⁰⁵ It is for these reasons that scholars have argued for the general or domestic criminalization of image based sexual abuse,¹⁰⁶ and new domestic criminal laws have been enacted by multiple states to accommodate this harm.¹⁰⁷

⁹⁸ OHCHR, *Sexual and gender-based violence in the context of transitional justice*, October 2014, available online at https://www.ohchr.org/sites/default/files/Documents/Issues/Women/WRGS/OnePagers/Sexual_and_gender-based_violence.pdf (visited 31 August 2023).

⁹⁹ Eurojust, *Prosecuting war crimes of outrage upon personal dignity based on evidence from open sources—Legal framework and recent developments in the Member States of the European Union* (2018), available online at https://www.eurojust.europa.eu/sites/default/files/Partners/Genocide/2018-02_Prosecuting-war-crimes-based-on-evidence-from-open-sources_EN.pdf (visited 31 August 2023); *Furundžija* Trial Judgment, *supra* note 87, § 272.

¹⁰⁰ Judgment, *Akayesu* (ICTR-96-4-T), Trial Chamber, 2 September 1998, § 688.

¹⁰¹ C. Rigotti and C. McGlynn, ‘Towards an EU criminal law on violence against women: The ambitions and limitations of the Commission’s proposal to criminalise image-based sexual abuse’, 13 *New Journal of European Criminal Law* (2022) 452–477.

¹⁰² A. Dodge, ‘Nudes are Forever: Judicial Interpretations of Digital Technology’s Impact on “Revenge Porn”’, 34 *Canadian Journal of Law and Society* (2019) 121–143.

¹⁰³ Rigotti and McGlynn, *supra* note 101, at 455.

¹⁰⁴ See for instance M. Joleby et al., ‘Experiences and Psychological Health among Children Exposed to Online Sexual Abuse—a Mixed Methods Study of Court Verdicts’, 27 *Psychology, Crime & Law* (2021) 159–181. Online child sexual abuse (or ‘child pornography’) has been widely discussed in literature and well-developed in domestic law, and thus falls outside the scope of this article. However, for a comprehensive overview of transnational and international responses to online child sexual abuse, see S.K. Witting, ‘Transnational by Default: Online Child Sexual Abuse Respects No Borders’, 29 *International Journal of Children’s Rights* (2021) 731–764.

¹⁰⁵ A. Dodge, ‘Digitizing Rape Culture: Online Sexual Violence and the Power of the Digital Photograph’, 12 *Crime, Media, Culture* (2016) 65–82.

¹⁰⁶ Goudsmit Samaritser, *supra* note 21, R.T. Toparlak, ‘Criminalising Deep Fake Pornography: A Gender-Specific Analysis of Image-Based Sexual Abuse’, 1 *cognitio—studentisches Forum für Recht und Gesellschaft* (2023); S. Bloom, ‘No Vengeance for “Revenge Porn” Victims: Unraveling Why This Latest Female-Centric, Intimate-Partner Offense Is Still Legal, and Why We Should Criminalize It’, 42 *Fordham Urban Law Journal* (2015) 233–290.

¹⁰⁷ For example, as of 2017, laws had been enacted in England and Wales, Scotland, Israel, Japan, Canada, Australia, New Zealand, and the United States; C. McGlynn and E. Rackley, ‘Image-Based Sexual Abuse’, 37 *Oxford Journal of Legal Studies* (2017) 534–561, at 554, fn 100.

Aside from the harm to direct victims of sexual violence, importantly, ICL has also recognized the harm to those who are forced to watch sexual violence as also being sexual in nature. For example, at the Special Court for Sierra Leone, the Trial Chamber recognized the harm caused to a man being forced to watch the rape and death of his wife as ‘fomenting terror by sexual means’.¹⁰⁸ This is also a relevant consideration for accommodating online sexual violence under ICL, since family members or friends of victims could also experience a form of sexual violence should they be forced to watch a video of a sexual assault or come across one that goes ‘viral’ online.

Overall, distributing footage of sexual assaults or sharing nude images without consent online or being made to watch such footage is an example of an online harm that existing ICL frameworks on sexual violence should be able to address. This is because even though they differ from ‘traditional’ offences where international courts have recognized being sexually assaulted or nude in public or physically witnessing a sexual offence as a crime, the fact that the harm occurs through a screen does not detract from its severity. Moreover, as argued previously in relation to sharing footage of crimes, the wide reach and instantaneousness of the online environment can exacerbate the sense of humiliation experienced by victims. As exemplified by the Rome Statute, provisions criminalizing forms of sexual violence are often vague — the lists of acts included under Articles 7(g) (crimes against humanity) and 8(2)(b)(xxii) and 8(2)(e)(vi) (war crimes) are, indeed, non-exhaustive and allow for room for interpretation.¹⁰⁹ It is thus plausible, given the similar nature of the harms previously recognized as grave enough to constitute sexual violence at ICCTs and the sharing of footage of sexual crimes or other nude imagery, that online sexual violence could also be considered a crime under ICL.

A final category of online sexual violence consists of when traditional international crimes, such as sexual slavery, are perpetrated directly through the online space. For example, in the context of the genocide against the Yazidi population in northern Iraq, ISIS used the internet to auction women and children, using posts with their registration information, photos, and minimum purchase prices.¹¹⁰ The crime of sexual slavery already exists under the Rome Statute,¹¹¹ and in this instance, using the internet as a forum to purchase or sell persons could constitute the offence as a crime against humanity or a war crime should the other elements of the crimes be met.¹¹² However, as noted by Koenig and Egan, using such online auctions also implied the sharing of a wealth of data about the victims to a global internet audience.¹¹³ In this instance, in addition to the crime of sexual slavery being committed, victims are further humiliated through their images and data being circulated online and their right to privacy is infringed upon when their personal information is shared. For this example of an online harm, though an offense is already available under ICL, the online component that exacerbates the harm can and should be considered as aggravating the offence, such as in the gravity assessment or at sentencing.

3. CONCLUSION

This article presented different forms of online harm that may be relevant in the context of ICL, including hate speech and disinformation, sharing footage of crimes online, and online sexual

¹⁰⁸ V. Oosterveld, ‘Sexual Violence Directed against Men and Boys in Armed Conflict or Mass Atrocity: Addressing a Gendered Harm in International Criminal Tribunals’, 10 *Journal of International Law & International Relations* (2014) 114.

¹⁰⁹ For instance, Art. 7(g) (crimes against humanity) includes ‘any other form of sexual violence of comparable gravity’ to the listed acts, and both provisions under Art. 8(2) (war crimes) include ‘any other form of sexual violence also constituting a grave breach of the Geneva Conventions’. Art. 7(g), 8(2)(b)(xxii), and 8(2)(e)(vi) ICCSt.

¹¹⁰ A. Koenig and U. Egan, ‘Power and Privilege: Investigating Sexual Violence with Digital Open Source Information’, 19 *JICJ* (2021) 55–84, at 66; P.V. Sellers and J.G. Kestenbaum, ‘Missing in Action: The International Crime of the Slave Trade’, 18 *JICJ* (2020) 517–542, at 524–525.

¹¹¹ Art. 7(1)(g), Art. 8(2)(b)(xii), and Art. 8(2)(e)(vi) ICCSt.

¹¹² Art. 7(1)(g)-2, Art. 8(2)(b)(xii)-2, and Art. 8(2)(e)(vi)-2 Elements of Crimes.

¹¹³ *Ibid.*

abuse. Though each is perpetrated using different technologies and by different actors, there are some commonalities. First, the nature of online harms of sharing footage of people in their most vulnerable circumstances (i.e. posting videos of executions, or online sexual violence such as sharing footage of rapes or images of nude victims), or of watching harm to a loved one, is one that severely degrades human dignity. The protection of human dignity, as argued by deGuzman, is a central reason for why acts are criminalized under ICL.¹¹⁴ Furthermore, as observed by Keenan, ICL chambers have historically relied on a harms-based approach to allow for new interpretations of international crimes, in which novel harms are compared to harms traditionally encompassed by ICL to justify their inclusion.¹¹⁵ If these two theories are applied to online harm, it would follow that because the online harms discussed cause similar types of harms to human dignity as those central to other ICL crimes, that they should also be able to be accommodated using existing ICL frameworks.

Nonetheless, a limitation of this approach is that other online harms that are ‘too novel’ and too far removed from existing ICL crimes would not be able to be encompassed in the same way — for instance, the use of surveillance technologies to monitor and suppress a minority group may be best addressed by developing a new crime, such as ‘digital apartheid’, which has been used by scholars and NGOs to describe current situations in South Africa and in Palestine.¹¹⁶ This is because digital surveillance methods can be particularly invasive and ever-present in all areas of life, in a way that was not possible prior to the development of such technologies. Because the technology lies at the heart of the offense, in line with the principle of ‘fair labelling’ in criminal law,¹¹⁷ it would be valuable for the label of the crime to reflect this for both the specific harm to victims to be fully recognized and to better show a distinction between the technology-facilitated offence and the ‘traditional’ crime (i.e. apartheid). As new technologies continue to develop, online harms may become more and more different from those recognized by ICL, and the only solution might be to create an entirely new crime, which would be much more difficult in practice.

Another overarching theme amongst the online harms discussed is the need for balancing rights. For instance, in terms of posting hate speech or circulating footage of crimes or intimate images online, any criminalization of this would need to be carefully balanced with the right of freedom of expression of the poster, which may be more clear-cut in some cases than others (e.g. posting a direct call for violence would obviously be prohibited as opposed to a post that is potentially offensive but not necessarily reaching the threshold for hate speech). Thus, if online harms are to be encompassed within existing ICL or accommodated *via* new crimes, these would need to consider a balance of fundamental rights and interests.

Furthermore, the examples of online harm discussed may fit already under the definitions of existing international crimes. For example, in the case of the spread of hate speech, this online harm could likely fall under existing definitions of persecution or incitement to

¹¹⁴ M.M. deGuzman, *Shocking the Conscience of Humanity: Gravity and the Legitimacy of International Criminal Law* (OUP, 2020).

¹¹⁵ P.J. Keenan, ‘Doctrinal Innovation in International Criminal Law: Harms, Victims, and the Evolution of the Law’, 42 *University of Philadelphia Journal of International Law* (2020) 407–444.

¹¹⁶ K. Hao and H. Swart, ‘South Africa’s Private Surveillance Machine is Fuelling a Digital Apartheid’, *MIT Technology Review*, 19 April 2022, available online at <https://www.technologyreview.com/2022/04/19/1049996/south-africa-ai-surveillance-digital-apartheid/> (visited 31 January 2024); Amnesty International, *Automated Apartheid: How Facial Recognition Fragments, Segregates and Controls Palestinians in the OPT*, May 2023, available online at <https://www.amnesty.org/en/documents/mde15/6701/2023/en/> (visited 31 January 2024); O. Zahzah, ‘Digital Apartheid: Palestinians being Silenced on Social Media’, *Aljazeera*, 13 May 2021, available online at <https://www.aljazeera.com/opinions/2021/5/13/social-media-companies-are-trying-to-silence-palestinian-voices> (visited 31 January 2024).

¹¹⁷ The principle of fair labelling ‘seeks to ensure that the definition of an offence would provide society with an accurate moral grasp of the defendant’s wrongdoing’ and that offences are categorized in a way that reflects different levels of wrongdoing and punishment; H.M. Zawati and T.A. Doherty, ‘Fair Labelling as a Common Legal Principle in Criminal Law’, in H.M. Zawati (ed.), *Fair Labelling and the Dilemma of Prosecuting Gender-Based Crimes at the International Criminal Tribunals* (OUP, 2014) 27.

genocide, or when footage of crimes is shared online, it could likely amount to an outrage upon personal dignity. However, this does not necessarily mean that they should be treated exactly the same as traditional crimes. As argued throughout the article, the addition of the online element typically serves to exacerbate the harm caused, which could make posting a video of a crime potentially even more humiliating than committing the same crime in a public square, where the footage is not preserved, distributed, and virtually impossible to get rid of. It is important to recognize the additional harm caused to victims by new technologies, and it could be argued that equating online harms with traditional offences is not enough to achieve justice. If this approach is taken in future cases, as argued above, it would be beneficial for ICL chambers to recognize online harms through means such as aggravating factors during gravity assessments or sentencing decisions to ensure the full scale of the harm to victims is acknowledged.

Finally, the aim of this article was to outline how online harms may foreseeably be included in future ICL cases to ensure greater accountability for online harms and justice for victims, which will likely be more or less difficult depending on the online harm. The article provided a survey of online harms that could imaginably be connected to the perpetration of international crimes and mapped the future direction of play for where those harms might fit best, given the nature of already-recognized harms. The process of either incorporating online harms into existing international criminal justice frameworks or recognizing new crimes in the future will require a specified and complex analysis, and would need to take various challenges into account, such as the acquisition and use of digital evidence, whether online harms can meet the gravity threshold for international crimes, and potential issues with jurisdiction for crimes committed in 'cyberspace'. Despite such challenges, it is clear that online harms are not insignificant and should be an issue at the forefront of the project of international criminal justice — as technology will only continue to evolve and inevitably serve as a vehicle for inflicting more and newer forms of harm.

ACKNOWLEDGEMENTS

The article was partly written while the author was a Visiting Scholar at the Human Rights Center at the University of California, Berkeley School of Law (USA). The author would like to thank Dr Barrie Sander, Dr Michelle Burgis-Kasthala, Professor Carla Ferstman, Dr Daragh Murray, Dr Alexa Koenig, and the anonymous reviewers for their helpful comments on previous drafts of this article.