

Dissertation Exposé

A double-edged sword?

Assessing the human rights impact of the introduction of user-generated digital evidence in international criminal proceedings

Submitted by:

Konstantina Despoina Stavrou, BA, LL.M (12044870)

PhD in Interdisciplinary Legal Studies

Supervisor: Univ.-Prof Dr Michael Lysander Fremuth

Research Field: International Human Rights Law, International Criminal Law

Seminar in International Law for the presentation of PhD Theses proposals (SE380034)

24 May 2021, Vienna

Table of Contents

Table of Abbreviations	iii
I. Introduction	1
II. Current State of Research & Relevance of the Topic	2
III. Description of the Issue	4
1. Human Rights-related Risks due to the Digital Nature of User-generated Digital Evidence	4
2. Involvement of Third Actors in International Criminal Proceedings	5
3. Reliance on Artificial Intelligence	6
4. Evidentiary Standards in International Criminal Law and User-generated Evidence	7
IV. Research Question and Sub-questions	8
V. Methodology	9
VI. Preliminary Structure	11
VII. Preliminary Bibliography	13

Table of Abbreviations

AI	Artificial Intelligence
ICC	International Criminal Court
ICCPR	International Covenant on Civil and Political Rights
ICL	International Criminal Law
ICTR	International Criminal Tribunal for Rwanda
ICTY	International Criminal Tribunal for the former Yugoslavia
IHRL	International Human Rights Law
NGO	Non-Governmental Organisation
STL	Special Tribunal for Lebanon

I. Introduction

Seventy-five years after the initiation of the first international criminal trials for the atrocities committed amid the Second World War,¹ the fight against impunity continues.² Core international crimes, encompassing genocide, war crimes, crimes against humanity and the crime of aggression,³ have continued to be perpetrated across the globe.⁴ In many cases the respective governments appear unable or unwilling to investigate the alleged commission of such crimes within their territories,⁵ reinforcing the need for alternative solutions, either through international(ised) courts and tribunals or through third states' domestic investigations. Nevertheless, the lack of sufficient evidence poses impediments in the initiation and progress of international criminal proceedings. Consequently, the question of how sufficient and proper evidence can be acquired to facilitate future investigations and prosecutions of core international crimes remains.

Lately, a novel type of digital evidence called 'user-generated evidence' is gaining increasing attention due to its potential to assist in overcoming one of the main challenges in international criminal proceedings, namely the lack of sufficient and proper evidence.⁶ User-generated evidence refers to digital evidence recorded by individuals in their digital devices.⁷ This information can be shared on the internet, such as on YouTube or on social media platforms like Facebook, or it can be gathered through NGO applications and be collected later as part of the investigation.⁸ User-generated digital evidence (hereinafter: user-generated evidence) becomes of significance especially in cases in which investigators do not have alternative ways of accessing documentation of core international crimes.⁹

Even though the use of user-generated evidence could assist in reducing impunity for core international crimes, the question is to what extent human rights of parties to the proceedings, such as the right to privacy and fair trial rights,¹⁰ might be impacted and impaired in the strive for justice. Despite the previous use of digital evidence in international criminal proceedings, research indicates that novel forms of digital evidence require particular attention and assessment by academia as they entail several new risks. The expansion of digital tools, making them broadly available, has resulted in the engagement of various private actors in the process of documentation and evidence collection.¹¹

¹ Steven R Ratner, Jason S Abrams, James L Bischoff, *Accountability for Human Rights Atrocities Beyond the Nuremberg Legacy* (3rd edn, OUP 2009), 209-210.

² Theodor Meron, 'Closing the Accountability Gap: Concrete Steps Toward Ending Impunity for Atrocity Crimes' (2018) 112(3) AJIL 433, 434.

³ Carsten Stahn, *A Critical Introduction to International Criminal Law* (CUP 2019), 15.

⁴ GCR2P, 'Atrocity Alert No. 198: Syria, Myanmar (Burma) and Sudan' (8 April 2020) <www.globalr2p.org/publications/atrocity-alert-no-198-syria-myanmar-burma-and-sudan/> accessed 10 April 2021.

⁵ Human Rights Watch, 'World Report 2020' (hereinafter: World Report 2020) <www.hrw.org/world-report/2020> accessed 10 April 2021; Seema Kassab, 'Justice in Syria: Individual Criminal Liability for Highest Officials in the Assad Regime' (2019) 39 MichJInt'lL 2, 289.

⁶ Rebecca J Hamilton, 'Social Media Platforms in International Criminal Investigations' (2020) 52(1) CaseWResJInt'lL 213.

⁷ Alexa Koenig and others, 'Open Source Fact-Finding in Preliminary examinations' in Morten Bergsmo and Carsten Stahn (eds) *Quality Control in Preliminary Examination: Volume 2* (TOAEP 2018), 684-685.

⁸ Berkeley Protocol, 6-7.

⁹ *ibid.*, v.

¹⁰ Those rights are enumerated in international and regional human rights treaties, eg International Covenant on Civil and Political Rights (adopted 16 December 1966, entered into force 23 March 1976) 999 UNTS 171 (ICCPR) art 17; Convention for the Protection of Human Rights and Fundamental Freedoms (European Convention on Human Rights, as amended) (ECHR), (adopted 4 November 1950, entered into force 3 September 1953) ETS 5, art 8.

¹¹ Rebecca Hamilton, 'User Generated Evidence' (2018) 57(1) ColumJTransnat'l L 1, 23.

Such actors do not necessarily adhere to criminal investigations' standards and practices and, thus, the legitimacy and reliability of evidence cannot always be ensured.¹² Moreover, technological advancements enable individuals to easily manipulate digital information, raising questions about its authenticity and credibility, in contrast to previous trials when technology was limited and mainly available in analogue form.¹³ Consequently, the introduction of innovative evidentiary material in courts requires a prior examination of the human rights implications, in addition to an assessment of their added value for international criminal proceedings.

II. Current State of Research & Relevance of the Topic

The use of technologically-derived evidence for international criminal justice purposes is long standing. Examples can be found in the Nuremberg trials where films and photographs were presented as evidence in court.¹⁴ However, the evidence used in the post-Second World War trials was in analogue form,¹⁵ with the introduction of digital evidence starting approximately fifty years later as a result of technological progress.¹⁶

The creation of the International Criminal Tribunals for the former Yugoslavia (ICTY) and Rwanda (ICTR) marks the first step in the evolution and use of digital evidence in international criminal proceedings. As an illustration, the ICTY prosecutors relied on satellite imagery as evidence to identify mass graves and to prove the destruction of towns through aerial bombardments.¹⁷ Similarly, the ICTR prosecution used radio broadcasts, which could be saved for years after the conflict due to technological advancements, to prove the incitement of violence against the Tutsi.¹⁸

The utilisation of user-generated evidence for international criminal justice purposes started receiving particular attention following the escalation of violence in Syria.¹⁹ More recently, in the cases of Sudan and Mali, the International Criminal Court (ICC) relied, among others, on leads from Facebook and YouTube.²⁰ Lastly, the key cases of *Al Mahdi*²¹ and *Al Werfalli*,²² where the court and the prosecution respectively relied on Facebook videos, suggest a new era in the use of user-generated evidence in international criminal proceedings.

¹² Alexander Heinze, 'Private International Criminal Investigations and Integrity' in Morten Bergsmo and Viviane E Ditttrich (eds), *Integrity in International Justice* (TOAEP 2020), 634-636.

¹³ UNOHCHR, HRC UC Berkeley School of Law, Berkeley Protocol on Digital Open Source Investigations- A Practical Guide on the Effective Use of Digital Open Source Information in Investigating Violations of International Criminal, Human Rights and Humanitarian Law (2020) (hereinafter: Berkeley Protocol), HR/PUB/20/2, 4.

¹⁴ Eric Stover, Victor Peskin and Alexa Koenig, *Hiding in Plain Sight: The Pursuit of War Criminals from Nuremberg to the War on Terror* (UCP 2016), 44.

¹⁵ *ibid.*

¹⁶ Lindsay Freeman, 'Digital Evidence and War Crimes Prosecutions: The Impact of Digital Technologies on International Criminal Investigations and Trials' (2018) 41(2) *Fordham Int'l LJ* 283, 286.

¹⁷ International Bar Association (IBA) International Criminal Court Programme, *Evidence Matters in ICC Trials* (2016), 24-25.

¹⁸ *Prosecutor v Nahimana* (Judgment and Sentence) ICTR-99-52-T (3 December 2003).

¹⁹ Rafael Braga da Silva, 'Sherlock at the ICC? Regulating Third-Party Investigations of International Crimes in the Rome Statute Legal Framework' (2020) 18 *JICJ* 59, 61-62.

²⁰ Freeman (n 16), 289.

²¹ *Prosecutor v Al Mahdi* (Judgment and Sentence) ICC-01/12-01/15-171 (27 September 2016), para 29; *Prosecutor v Al Mahdi* (Transcript) ICC-01/12-01/15-T-4-Red-ENG (22 August 2016), 28-29.

²² *Prosecutor v Mahmoud Mustafa Busayf Al-Werfalli* (Warrant of Arrest) ICC-01/11-01/17.

The Rohingya case in Myanmar further exemplifies the potential of user-generated evidence. The government of Myanmar has repeatedly refused access to the ground for independent, outside observers,²³ making user-generated content, complemented by interviews with refugees, a key source of documentation of the crimes committed against the religious minority.²⁴ According to an investigation conducted by Reuters, more than 1.000 examples of posts, comments and pornographic images attacking the Rohingya and other Muslims were found on Facebook,²⁵ while it is estimated that as many as 700 people were involved in the attacks on Facebook.²⁶ The documentation provided the basis for the establishment of the Independent Investigative Mechanism for Myanmar to promote accountability.²⁷ While the evidence has not yet been challenged before criminal courts, it shows its potential to decrease evidentiary gaps and to allow the initiation of criminal proceedings even in cases where domestic and international investigations are impeded.²⁸

The significance of user-generated evidence is growing in domestic proceedings for core international crimes. Cases have been opened in Germany, Finland and Sweden, in which the prosecution relied on images and videos shared through social media as evidence.²⁹ Several of those cases have led to convictions of the suspects,³⁰ proving the added value of novel types of digital information as a key source of evidence available (almost) immediately after the commission of a crime, which might otherwise be lost or destroyed over time.³¹

The topic of digital evidence in international criminal law is also gaining attention in the academic sphere.³² Legal scholars have assessed the added value of digital evidence as a way to reinforce accountability for core international crimes.³³ Due to the lack of sufficient and appropriate evidence in many cases, research discusses the potential of this evidence to advance criminal proceedings for core international crimes.³⁴ The academic field analyses the introduction of digital evidence, including user-generated evidence, as a positive development.³⁵

²³ ‘UN Rights Expert “Disappointed” by Myanmar’s Decision to Refuse Visit’ *UN News* (20 December 2017) <<https://news.un.org/en/story/2017/12/639982-un-rights-expert-disappointed-myanmars-decision-refuse-visit>> accessed 11 April 2021.

²⁴ UNHRC, Report of the Independent International Fact-Finding Mission on Myanmar, A/HRC/39/64(2018).

²⁵ Steve Stecklow, ‘Why Facebook is losing the war on hate speech in Myanmar’ *Reuters* (15 August 2018) <www.reuters.com/investigates/special-report/myanmar-facebook-hate/> accessed 11 April 2021.

²⁶ Paul Mozur, ‘A Genocide Incited on Facebook, With Posts From Myanmar’s Military’ *The New York Times* (15 October 2018) <www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html> accessed 11 April 2021.

²⁷ UNHRC, Resolution adopted by the Human Rights Council on 27 September 2018- Situation of human rights of Rohingya Muslims and other minorities in Myanmar, A/HRC/RES/39/2 (2018), paras 22-23.

²⁸ Konstantina Stavrou, ‘Civil Society and the IIMM in the Investigation and Prosecution of the Crimes Committed Against the Rohingya’ (2021) 36(1) *UJIEL* 95, 100-101.

²⁹ *UJAR* 2020, 27, 52, 74.

³⁰ *ibid.*

³¹ Sean Bain, ‘A legal path to justice emerges for Myanmar’ (*Justice Hub*, 7 October 2018) <www.icj.org/a-legal-path-to-justice-emerges-for-myanmar/> accessed 12 April 2021.

³² The forthcoming special issue of the JICJ focuses on ‘New Technologies and the Investigation of International Crimes’.

³³ Jay D Aronson, ‘The Utility of User-Generated Content in Human Rights Investigations’, in Molly K Land and Jay D Aronson (eds) *New Technologies for Human Rights Law and Practice* (CUP 2018), 130. Mark Kersten, ‘Challenges and Opportunities: Audio-Visual Evidence in International Criminal Proceedings’ (*Justice in Conflict*, 4 March 2020) <<https://justiceinconflict.org/2020/03/04/challenges-and-opportunities-audio-visual-evidence-in-international-criminal-proceedings/>> accessed 10 April 2021.

³⁴ Hamilton (n 11), 6-7.

³⁵ Keith Hiatt, ‘Open Source Evidence on Trial’ (2016) 125 *YaleLJF* 323.

Existing research and practice concerning user-generated evidence focuses on the evidence collection and analysis stages. Scholarly conversation on risks tackles the admissibility,³⁶ reliability³⁷ and biases of digital evidence.³⁸ Manipulation and misinformation through deepfakes have also been touched upon.³⁹ Recognising that such concerns could result in the exclusion of crucial digital evidence from criminal proceedings, studies deal with the ways risks could be mitigated and its admissibility could be ensured.⁴⁰ Lastly, due to the growing importance of digital evidence, research has also been launched on its possible impact on the rules of procedure and evidence at the international level.⁴¹

Thus far, there is little consideration of the human rights aspects of the introduction and utilisation of user-generated evidence, with publications including only brief mentions to fair trial rights.⁴² However, using user-generated evidence for criminal accountability purposes is a puzzling process on multiple levels, entailing several unforeseen legal challenges. Among the main issues are the manipulation of evidence,⁴³ authenticity,⁴⁴ biases, as well as the protection of privacy and fair trial rights.⁴⁵ Hence, its use should not remain unreflective, without a thorough analysis of its legal underpinnings. What is clearly missing from existing scientific work is an in-depth analysis of the human rights impact of the introduction of user-generated evidence in international criminal proceedings, with a long-term perspective on the ways to mitigate such possible implications. Before resorting to the use of such evidence it is, therefore, crucial to ensure that the material will not result in more harm than good. The need for this research is pressing to understand all human rights repercussions and legal consequences of the reliance on user-generated evidence, as well as to ensure the integrity of international accountability proceedings and the protection of the rights of the victims, witnesses and suspects.

III. Description of the Issue

1. Human Rights-related Risks due to the Digital Nature of User-generated Digital Evidence

An inquiry into literature regarding user-generated evidence reveals several new challenges due to its digital nature. One key issue related to user-generated evidence, raised from the initial stages of evidence collection, is its volatile nature due to the possibility of manipulation.⁴⁶ The technological developments, in combination with the proliferation of technology, have made digital technologies

³⁶ Freeman (n 16).

³⁷ IBA (n 17), 20.

³⁸ Jay D Aronson, 'Mobile Phones, Social Media and Big Data in Human Rights Fact-Finding', in Philip Alston and Sarah Knuckey (eds) *The Transformation of Human Rights Fact-Finding* (OUP 2018), 447.

³⁹ Lindsay Freeman, 'Law in Conflict- The Technological Transformation of War and Its Consequences for the International Criminal Court' (2019) 9 NYUJILP 16, 859.

⁴⁰ Freeman (n 16); Hamilton (n 11).

⁴¹ eg <www.nurembergacademy.org/projects/detail/45ed2d129b0e19459764c4684e317a95/digital-evidence-23/> accessed 6 April 2021.

⁴² Rebecca Hamilton, 'The Hidden Danger of User-Generated Evidence for International Criminal Justice' (*Just Security*, 23 January 2019) <www.justsecurity.org/62339/hidden-danger-user-generated-evidence-international-criminal-justice/> accessed 10 May 2021.

⁴³ Berkeley Protocol, 4.

⁴⁴ Hamilton (n 6), 218.

⁴⁵ Aronson (n 38), 447.

⁴⁶ Janosch Delcker, 'Welcome to the age of uncertainty. If everything can be faked, how can we know anything is real?' *Politico* (8 March 2021) <www.politico.eu/article/deepfake-videos-the-future-uncertainty/> accessed 15 April 2021.

broadly available. This allows individuals to easily access and falsify information. An illustration is the case of ‘deepfakes’, which are videos created by algorithms making it look like a person said or did something they did not.⁴⁷ Deepfake technologies are becoming increasingly available at a lower cost or for free, making them accessible for a variety of actors. By nature, deepfakes are difficult to detect, while concerns have been expressed regarding the possibility that in the future a detection of deepfakes might not be possible.⁴⁸ The utilisation of evidence manipulated by deepfake technologies could impact the rights of suspects.

Furthermore, collecting user-generated evidence is exclusively dependent on the availability of technologies. This can result in biases in evidence collection and distortion of the events. One main issue is the limited access to digital devices or the internet, which could affect the representativeness of the documentation as crimes might only be recorded from a selective perspective.⁴⁹ Moreover, because of the real-time documentation of the commission of crimes, the evidence presented might be distorted, based on the narrative of the person recording.

Finally, the overwhelming amount of documentation can pose challenges. In Syria, the broad use of digital devices resulted in what is estimated to more than four million videos uploaded to YouTube⁵⁰ with the hours documented exceeding the hours of the real war.⁵¹ The abundance of documentation makes it, on the one hand, tougher to find the significant pieces of information in all the noise, potentially affecting international criminal proceedings. On the other hand, the equality of arms could be impacted because of the over-collection of inculpatory compared to exculpatory evidence.

2. Involvement of Third Actors in International Criminal Proceedings

The emergence of user-generated evidence in international criminal law has resulted in the active involvement of private third actors in international criminal proceedings in the collection and analysis of evidence.⁵² Among the actors involved are NGOs, witnesses, as well as technologists.⁵³ The involvement of NGOs and witnesses is not new. Material gathered by NGOs on the ground had already been used in international criminal investigations by the ICTY⁵⁴ and the ICTR.⁵⁵ Nonetheless, the active involvement of private third actors in evidence collection intended specifically to be admitted in court is a new development.

Even though the work of those actors is central in the case of user-generated evidence, it comes with complications. A first challenge relates to the lack of a professional code of conduct and standards in

⁴⁷ Alexa Koenig, “‘Half the Truth is Often a Great Lie’: Deep Fakes, Open Source Information, and International Criminal Law” (2019) 213 *AJIL Unbound* 250, 251.

⁴⁸ Marie Helen Maras, Alex Alexandrou, ‘Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos’ (2019) 23(3) *IJE&P* 255, 256.

⁴⁹ Aronson (n 38), 447.

⁵⁰ Stephanie van der Berg, ‘Mass atrocities? There’s an app for that’ (*Justice Info*, 5 February 2019) <www.justiceinfo.net/en/40176-mass-atrocities-there-s-an-app-for-that.html> accessed 14 April 2021.

⁵¹ Armin Rosen, ‘Erasing History: YouTube’s Deletion of Syria War Videos Concerns Human Rights Groups’ (*Fast Company*, 7 March 2018) <www.fastcompany.com/40540411/erasing-history-youtubes-deletion-of-syria-war-videos-concerns-human-rights-groups> accessed 15 April 2021.

⁵² Hamilton (n 11), 22-23.

⁵³ *ibid.*, 27.

⁵⁴ *Prosecutor v Tolimir*, (Judgment) IT-05-8/2-T (12 December 2012), para 50.

⁵⁵ *Prosecutor v Setako* (Judgment and Sentence) ICTR-04-081-T (25 February 2010), para 164.

investigations,⁵⁶ which might raise concerns about rights' violations at the evidence collection stage. An illustration is the question whether the lack of consent of victims to be photographed or filmed would result in a violation of privacy rights of those individuals and how this could impact the admissibility of evidence. In such cases, the question that also arises is whether those actors are bound by human rights and what could be the reaction to counterbalance the lack of obligations in that respect.

Additionally, due to the lack of a professional code of conduct, the partiality of information could come under scrutiny. The diverse interests of actors involved in evidence collection, including due to donors supporting private actors, could substantially impact the type of information collected and the way it is presented. The Commission for International Justice and Accountability has already received such criticism due to its exclusive focus on the Syrian government and Daesh.⁵⁷ The collection of evidence representing only one side of the story regarding a crime would not only be distorting the truth, but, in case it is admitted as evidence in trial, it could also affect the rights of suspects.

Lastly, the involvement of third actors has resulted in abundance of information on the commission of core international crimes for some cases. As third actors involved in evidence collection and analysis are usually aiming to promote accountability, it is more likely that inculpatory evidence, rather than exculpatory, will be gathered.⁵⁸ This could affect the equality of arms, as defence teams might not have the same opportunities to counter the evidence collected and to obtain sufficient evidence to support their counter-claims.

3. Reliance on Artificial Intelligence

Following the proliferation of digital technologies, the evidentiary challenge in international criminal law has shifted from ensuring access to evidentiary material to finding the relevant material within the abundance of information. This development has led to the increased reliance on Artificial Intelligence (AI). AI is being used to support the discovery of user-generated information that could serve as evidence,⁵⁹ as well as to allow the timely processing and analysis of information to identify the relevant pieces of evidence.⁶⁰ Despite its potential contribution to international criminal proceedings, the use of AI comes with challenges.

Firstly, the reliance on AI, as means to identify evidence among the abundance of user-generated content available on the internet, is subject to technical biases.⁶¹ As algorithms tend to prioritise accounts and information which receive bigger attention or have more followers over smaller accounts, certain populations might be further marginalised, while there can be difficulties in the discovery of some crimes and over-documentation of some others.⁶² This could affect investigators'

⁵⁶ Morten Bergsmo and William H Wiley, 'Human Rights Professionals and the Criminal Investigation and Prosecution of Core International Crimes' in Skåre and others (eds), *Manual on Human Rights Monitoring. An Introduction for Human Rights Field Officers* (Norwegian Centre for Human Rights 2008), 5.

⁵⁷ Melinda Rankin, 'The Future of International Criminal Evidence in New Wars? The Evolution of the Commission for International Justice and Accountability' (2018) 20(3) JGR 392, 404.

⁵⁸ Hamilton (n 11), 39.

⁵⁹ Koenig (n 47), 251.

⁶⁰ eg the use of VFRAME by the Syrian Archive <<https://syrianarchive.org/en/about/methods-tools>> accessed 27 April 2021.

⁶¹ Yvonne McDermott, Alexa Koenig and Daragh Murray, 'Open Source Information's Blind Spot: Human and Machine Bias in International Criminal Investigations' (2021) JICJ 1, 7.

⁶² *ibid.*

perception of events and result in selection biases, where the identified information does not fully represent the situation on the ground. Representativeness of the results could impact the integrity of international criminal proceedings as the ICC, for instance, is obligated to ‘investigate incriminating and exonerating circumstances equally.’⁶³

Another problem is discrimination due to algorithmic biases, directly affecting the analysis of recovered evidence. Algorithms are informed by training data and can have inherent biases which lead to discrimination, even without the knowledge of the investigator.⁶⁴ In the case of investigations for the commission of core international crimes, this could raise serious concerns regarding the impartiality of evidence.

Lastly, AI is being used to identify deepfakes and, hence, to verify digital information.⁶⁵ An example is the extraction of data from the faces in individuals frames of videos, which is then tracked through sets of concurrent frames, as means to identify inconsistencies in deepfakes.⁶⁶ While such technologies can prove useful, they are not perfect. Detection software would need to keep up with innovations in deepfake technologies, while it cannot be ensured that the dangers would be eliminated.⁶⁷ A result could be false convictions based on unidentified deepfakes.

4. Evidentiary Standards in International Criminal Law and User-generated Evidence

User-generated evidence requires a thorough vetting to ensure the integrity of international criminal proceedings and the protection of human rights of victims, witnesses, and suspects. As different criminal accountability fora employ different evidentiary standards, it is necessary to examine whether there are guidelines regarding the assessment of the admissibility and use of (user-generated) digital evidence.

Taking as an illustration the ICC, the Chambers have a discretion to decide when to make a determination of the admissibility of evidence either during trial as evidence is presented or in the final judgement,⁶⁸ with the majority opting for the latter. Nonetheless, there have been exceptions in cases involving digital forms of evidence, especially when it is central to the Prosecution’s case, for instance *Bemba et. al.*⁶⁹ Despite the wide discretion afforded to the Chambers, the admissibility of evidence is guided by rules regarding the fair evaluation of evidence and fair trial⁷⁰ and evidence can be excluded under certain circumstances, including the violation of internationally recognised human rights.⁷¹

⁶³ ICCSt, art 54(1)(a).

⁶⁴ McDermott, Koenig, Murray (n 61), 13.

⁶⁵ Koenig (n 47), 254.

⁶⁶ John Sohrawardi, Matthew Wright, ‘In a Battle of AI versus AI, Researchers Are Preparing for the Coming Wave of Deepfake Propaganda’, (*The Conversation*, 9 October 2020) <<https://theconversation.com/in-a-battle-of-ai-versus-ai-researchers-are-preparing-for-the-coming-wave-of-deepfake-propaganda-146536>> accessed 27 April 2021.

⁶⁷ Bobby Chesney, Danielle Citron, ‘Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security’ (2019) 107 CalLRev 1753, 1787.

⁶⁸ *Prosecutor v Bemba* (Judgment on the Appeals of Mr Jean-Pierre Bemba Gombo and the Prosecutor against the decision of Trial Chamber III entitled “Decision on the admission into evidence of materials contained in the prosecution’s list of evidence”) ICC-01/05-01/08-1386 OA 5 OA 6 (3 May 2011), para 37.

⁶⁹ *Prosecutor v Bemba* (Decision on Requests to Exclude Dutch Intercepts and Call Data Records) ICC-01/05-01/13-1855 (29 April 2016).

⁷⁰ ICCSt, art 69(4).

⁷¹ ICCSt, art 69(7).

Due to the challenges entailed in the use of user-generated evidence, the Berkeley Human Rights Centre launched, in cooperation with the United Nations Office of the High Commissioner for Human Rights, the Berkeley Protocol on Digital Open Source Investigations.⁷² The Protocol includes the preparation and planning of investigations and delves into the perplexing procedures of evidence collection, verification, and analysis, setting universal standards for handling diverse types of digital evidence. The aim of the Protocol, among others, is to assist lawyers and judges to comprehend the opportunities and challenges of novel forms of digital evidence, such as user-generated evidence.

Nevertheless, the Protocol can only provide guidance to judges, lawyers, and investigators without introducing binding standards. Therefore, an in-depth analysis of the case law of international(ised) courts and tribunals is necessary to map the instances in which digital evidence has been excluded and the grounds for exclusion. Additionally, emphasising on the ICC, a case law examination is necessary to determine the international human rights recognised by the court, whose violation would amount to the exclusion of evidentiary material, and how this has been applied in practice.

Considering the human rights risks entailed in the use of user-generated evidence, a lack of specific guidelines for its thorough assessment could impact the integrity of international criminal proceedings. Moreover, the protection of relevant human rights of victims, witnesses, and suspects of core international crimes, could be affected. The predominant result could be the admission of potentially harmful evidence, in the terms of respect for the human rights of the parties to the proceedings. An example is the disclosure of documentary material violating the privacy or dignity of victims, thus, bringing into question whether and how the evidence should be used in trial.

IV. Research Question and Sub-questions

As the previous analysis showcases, the assumption of the proposed doctoral project is that, due to risks related to user-generated evidence, a human rights impact is expected. The dissertation will analyse the under-researched area of the human rights implications of the use of user-generated evidence in international criminal proceedings in light of the aforementioned risks. The research will additionally explore the possible legal responses to its findings.

The main research question the doctoral thesis will answer is:

To what extent could the introduction and use of user-generated digital evidence in international criminal proceedings impact the human rights of victims, witnesses, and suspects of core international crimes?

To guide the research and answer the main research question, several sub-questions will be addressed:

- (a) What are the human rights challenges posed by user-generated evidence due to its digital nature?
- (b) What are the risks entailed in third actors' involvement in international criminal proceedings?
- (c) To what extent could the reliance on AI in the case of user-generated digital evidence impact human rights guarantees for victims, witnesses, and suspects?
- (d) What are the standards and rules of evidence in the case of (user-generated) digital evidence?

⁷² Berkeley Protocol.

- (e) How could the possible human rights implications of the reliance on user-generated digital evidence be mitigated?

The thesis will be divided in four parts. The *first* part will set the legal framework. After defining its international criminal law framework, the research will identify the relevant human rights for each actor, i.e. victims, witnesses, and suspects. Some of the identified human rights will cover more than one actor, such as the right to privacy referring to victims and witnesses, while others, such as the fair trial guarantees, mainly referring to suspects, will focus exclusively on one actor. The research will cluster these rights according to the different actors and address potential conflicts among them. The *second* part will present the emergence of user-generated evidence and its significance, including through the presentation of recent examples involving it. The *third* part will be devoted to the analysis of the human rights impact of user generated evidence for victims, witnesses, and suspects of core international crimes. The part will be structured in four chapters reflecting the aforementioned issues, in light of which the different human rights implications for each actor, during the investigative and trial stage, will be presented. The *final* part will present the conclusions, solutions, and recommendations of the research, also considering the conflicting nature of some the rights affected.

V. Methodology

The primary method of the doctoral project will be doctrinal legal research through a review and analysis of identified relevant sources. To set its international criminal law framework and define its terms, the research will draw on primary sources, namely the statutes and selected case law from the international(ised) courts and tribunals, as well as domestic courts. The pool of cases has been preliminary identified through secondary literature review, while research on the respective databases of international courts and tribunals will allow the identification of additional case law. The decision to rely on international(ised) courts and tribunals, and domestic courts is partly justified by the principle of complementarity in international criminal law, guiding the work of the ICC and dictating that the Court is entitled to take action only under specific circumstances.⁷³ Moreover, the decision is justified due to the primacy of some of those courts over national courts, as in the case of the ICTY, ICTR and STL.⁷⁴ The inclusion of domestic courts is due to the primacy they (usually) enjoy over international courts' jurisdiction, as in the case of ICC.

To set the dissertation's human rights framework, international documents both from the international criminal law and the human rights law sphere will be relied on to encapsulate fair trials guarantees and due process rights, as well as general human rights that would be applicable. The elements of each identified human rights obligation will be determined, namely the scope of application, the category of obligation and the limitations allowed for each right. From the international criminal law side, the provisions of the statutes of international(ised) courts and tribunals (especially the ICC) will be considered in mapping fair trial guarantees and due process rights. From the human rights law side, the analysis will draw on international human rights standards, deriving from the ICCPR and the communications and views of the Human Rights Committee. Additionally, regional human rights

⁷³ ICCSt, preamble, arts 1, 17, 18, 19; Robert Cryer and others, *Introduction to International Criminal Law and Procedure* (OUP 2007), 127.

⁷⁴ *ibid* 104, 113.

conventions and the case law of regional human rights courts, such as the European Convention on Human Rights and the European Court of Human Rights, will be considered.

To present the emergence of user-generated evidence, the opportunities and the challenges entailed in its use, the research will engage in a critical analysis of secondary sources. The literature will include books, contributions to edited books, as well as articles in legal journals by leading experts on the topic of user-generated evidence. Complementarily, reports from NGOs involved in evidence collection and analysis will be studied when presenting the new forms of digital evidence and the actors involved in the process.

To assess the human rights impact of user-generated evidence on the different sets of rights, the research will rely on its defined human rights framework. In examining the rules of admissibility and exclusion of user-generated evidence, due to the limited cases involving it and the anticipated lack of specific provisions regarding its use in the rules of procedure and evidence of international(ised) courts and tribunals, the research will rely on case law of the aforementioned judicial bodies as well as soft law.

As a secondary method, limited amount of semi-structured interviews with professionals working in the international criminal justice field will be utilised.⁷⁵ Due to the novelty of the research subject and the potential limited sources on it (including case law), the empirical research will be necessary to facilitate and support the legal research. The use of semi-structured interviews as a research method will allow to ask additional questions to better interpret the obtained answers. The form of in-depth and semi-structured interviews will permit to draw attention to practitioners' experience with user-generated evidence and to hear their recommendations.⁷⁶ The interviews will be analysed by qualitative content analysis.

⁷⁵ Urszula Jaremba, Elaine Dr Mak, 'Interviewing Judges in the Transnational Context' (2014) LaM.

⁷⁶ P Ishwara Bhat, *Idea and Methods of Legal Research* (OUP 2019), 375.

VI. Preliminary Structure

This section presents the preliminary structure of the proposed doctoral thesis. Nonetheless, due to the novelty of the topic and potential developments in the field, minor alterations might be implemented.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND ACRONYMS

1. Introduction
 - 1.1. Background
 - 1.2. Research Question and Sub-questions
 - 1.3. Methodology
 - 1.4. Arguments and Structure of Dissertation
 - 1.5. Limitations

PART I- LEGAL FRAMEWORK

2. Introduction
3. International Criminal Law
 - 3.1. Definition of International Criminal Law & International Criminal Proceedings
 - 3.2. Sources of International Criminal Law
 - 3.3. Evidence in International Criminal Proceedings
4. Human Rights Guarantees in Court Proceedings: Where ICL and IHRL meet
 - 4.1. Human Rights Guarantees under IHRL
 - 4.1.1. Definition, Scope and Content of Guarantees
 - 4.1.2. Sources of Human Rights Guarantees
 - 4.2. Human Rights Guarantees under ICL
 - 4.2.1. Definition, Scope and Content of Guarantees
 - 4.2.2. Sources of Human Rights Guarantees
5. Conclusion

PART II- THE EMERGENCE OF USER-GENERATED DIGITAL EVIDENCE

6. Introduction
7. Evidence in ICL
 - 7.1. The Evolution of Evidence in ICL
 - 7.2. Evidentiary Problems in ICL
8. The Introduction of User-generated Digital Evidence as Part of the Solution
 - 8.1. Definition of User-generated Digital Evidence
 - 8.2. Actors Involved in Evidence Collection and Analysis
 - 8.2.1. Background
 - 8.2.2. The New Role for Third Actors in ICL
 - 8.2.3. Human Rights Obligations for Third Actors
 - 8.3. Recent Cases Involving User-generated Digital Evidence
 - 8.3.1. Syria
 - 8.3.2. Yemen

8.3.3. Myanmar

9. Conclusion

PART III- COMPLICATING THE PICTURE: USER-GENERATED EVIDENCE AS A DOUBLE-EDGED SWORD?

10. Introduction

11. Digital Nature of User-generated Digital Evidence

11.1. Human Rights Challenges at the Investigative Stage

11.2. Human Rights Challenges at the Trial Stage

12. Third Actors' Involvement

12.1. Human Rights Challenges at the Investigative Stage

12.2. Human Rights Challenges at the Trial Stage

13. Reliance on Artificial Intelligence

13.1. Definition of Artificial Intelligence

13.2. The Use of Artificial Intelligence in International Criminal Proceedings

13.3. Human Rights Challenges at the Investigative Stage

13.4. Human Rights Challenges at the Trial Stage

14. Standards and Rules of Evidence: The Case of User-generated Digital Evidence

14.1. Sources of Standards and Rules of User-generated Digital Evidence in ICL

14.2. Vetting of Evidence

14.3. Human Rights Challenges at the Investigative Stage

14.4. Human Rights Challenges at the Trial Stage

15. Conclusion

PART IV- CONCLUSIONS AND RECOMMENDATIONS

16. User-generated Evidence: A Challenge for Human Rights Guarantees?

17. Recommendations

BIBLIOGRAPHY

INDEX

VII. Preliminary Bibliography

- Literature

Books and contributions to edited books

Alston P and Knuckey S, *The Transformation of Human Rights Fact-Finding* (OUP 2018)

Aronson JD, 'The Utility of User-Generated Content in Human Rights Investigations' in Molly K Land and Jay D Aronson (eds) *New Technologies for Human Rights Law and Practice* (CUP 2018)

Bassiouni MC, *Introduction to International Criminal Law* (2nd rev edn, Martinus Nijhoff 2012)

Bergsmo M, Stahn C (eds) *Quality Control in Preliminary Examination: Volume 2* (Torkel Opsahl Academic EPublisher 2018)

Bergsmo M, Wiley WH, 'Human Rights Professionals and the Criminal Investigation and Prosecution of Core International Crimes' in Skåre and others (eds), *Manual on Human Rights Monitoring. An Introduction for Human Rights Field Officers* (Norwegian Centre for Human Rights 2008)

Black E, *IBM and the Holocaust: The Strategic Alliance Between Nazi Germany and America's Most Powerful Corporation* (Crown 2001)

Boas G and others, *International Criminal Law Practitioner Library: International Criminal Procedure*, vol 3 (CUP 2011)

Combs NA, *Fact-Finding without Facts: The Uncertain Evidentiary Foundations of International Criminal Convictions* (CUP 2010)

Cryer R and others, *An Introduction to International Criminal Law and Procedure* (CUP 2007)

Dubberley S, Koenig A and Murray D, *Digital Witness: Using Open Source Information for Human Rights Investigation, Documentation, and Accountability* (OUP 2019)

Dutelle AW, *An Introduction to Crime Scene Investigation* (3rd edn, Jones and Bartlett Learning 2016)

Hafetz J, *Punishing Atrocities through a Fair Trial- International Criminal Law from Nuremberg to the Age of Global Terrorism* (OUP 2018)

Heinze A, 'Private International Criminal Investigations and Integrity' in Morten Bergsmo and Viviane E Dittrich (eds), *Integrity in International Justice* (TOAEP 2020)

Heller K and others, *The Oxford Handbook of International Criminal Law* (OUP 2020)

Khan KAA, Buisman C and Gosnell C, *Principles of Evidence in International Criminal Justice* (OUP 2010)

Konstantina Despoina Stavrou

Klamberg M, *Evidence in International Criminal Trials: Confronting Legal Gaps and the Reconstruction of Disputed Events* (Martinus Nijhoff 2013)

Land MK, Aronson JD (eds) *New Technologies for Human Rights Law and Practice* (CUP 2018)

McBride J, *Human rights and Criminal Procedure- The Case Law of the European Court of Human Rights* (2nd edn, Council of Europe 2018)

McDermott Y, *Fairness in International Criminal Trials* (OUP 2016)

Ratner SR, Abrams JS, Bischoff JL, *Accountability for Human Rights Atrocities Beyond the Nuremberg Legacy* (3rd edn, OUP 2009)

Safferling C, *International Criminal Procedure* (OUP 2012)

Sluiter G and others, *International Criminal Procedure: Principles and Rules* (OUP 2013)

Stahn C, *A Critical Introduction to International Criminal Law* (CUP 2019)

Stover E, Peskin V and Koenig A, *Hiding in Plain Sight: The Pursuit of War Criminals from Nuremberg to the War on Terror* (UCP 2016)

Trechsel S, *Human Rights in Criminal Proceedings* (OUP 2005)

Legal Journals

Braga da Silva R, 'Sherlock at the ICC? Regulating Third-Party Investigations of International Crimes in the Rome Statute Legal Framework' (2020) 18 JICJ 59

Chesney B and Citron D, 'Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security' (2019) 107 CalLRev 1753

Hamilton RJ, 'User Generated Evidence' (2018) 57(1) ColumJTransnat'l L 1

Hamilton RJ, 'Social Media Platforms in International Criminal Investigations' (2020) 52(1) CaseWResJInt'l L 213

Hiatt K, 'Open Source Evidence on Trial' (2016) 125 YaleLJF 323

Freeman L, 'Digital Evidence and War Crimes Prosecutions: The Impact of Digital Technologies on International Criminal Investigations and Trials' (2018) 41(2) FordhamInter'l LJ 283

Freeman L, 'Law in Conflict- The Technological Transformation of War and Its Consequences for the International Criminal Court' (2019) 9 NYUJILP 16

Kassab S, 'Justice in Syria: Individual Criminal Liability for Highest Officials in the Assad Regime' (2019) 39 MichJInt'l L 2

Koenig A, "'Half the Truth is Often a Great Lie": Deep Fakes, Open Source Information, and International Criminal Law' (2019) 213 AJIL Unbound 250

Konstantina Despoina Stavrou

Maras ME, Alexandrou A, 'Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos' (2019) 23(3) IJE&P 255

McDermott Y, Koenig A and Murray D, 'Open Source Information's Blind Spot: Human and Machine Bias in International Criminal Investigations' (2021) JICJ 1

Mehandru N, Koenig A, 'Open Source Evidence and the International Criminal Court', HarvHumRtsJ (15 April 2019) <https://harvardhrj.com/2019/04/open-source-evidence-and-the-international-criminal-court/#_ftn10> accessed 17 May 2021

Meron T, 'Closing the Accountability Gap: Concrete Steps Toward Ending Impunity for Atrocity Crimes' (2018) 112 AJIL 3

Quattrococo S and others, 'Technical Solutions for Legal Challenges: Equality of Arms in Criminal Proceedings' (2020) 20(1) Global Jurist 1

Legal blogs

Hamilton RJ, 'The Hidden Danger of User-Generated Evidence for International Criminal Justice' (*Just Security*, 23 January 2019) <www.justsecurity.org/62339/hidden-danger-user-generated-evidence-international-criminal-justice/> accessed 17 May 2021

Kersten, M, 'Challenges and Opportunities: Audio-Visual Evidence in International Criminal Proceedings' (*Justice in Conflict*, 4 March 2020) <<https://justiceinconflict.org/2020/03/04/challenges-and-opportunities-audio-visual-evidence-in-international-criminal-proceedings/>> accessed 17 May 2021

van der Berg S, 'Mass atrocities? There's an app for that' (*Justice Info* 5 February 2019) <www.justiceinfo.net/en/other/40176-mass-atrocities-there-s-an-app-for-that.html> accessed 17 May 2021

Guides and studies

Eurojust Genocide Network, *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) <www.eurojust.europa.eu/sites/default/files/Partners/Genocide/2018-02_Prosecuting-war-crimes-based-on-evidence-from-open-sources_EN.pdf> accessed 17 May 2021

European Court of Human Rights, Guide on Article 6 of the European Convention on Human Rights: Right to a fair trial (criminal limb) (2020) <echr.coe.int/documents/guide_art_6_criminal_eng.pdf> accessed 17 May 2021

International Bar Association, *Evidence Matters in ICC Trials* (2016)

Matheson K, *Video as Evidence Field Guide* (WITNESS 2016) <<https://vae.witness.org/video-as-evidence-field-guide/>> accessed 17 May 2021

STL, *Primer on Telecommunications Evidence: Guide to understanding the testimony in Ayyash et al* (STL 2017) <www.stl-tsl.org/sites/default/files/bulletin/Primer.pdf> accessed 17 May 2021

NGO reports

Trial International, Universal Jurisdiction Annual Review 2021 (UJAR 2021)- *A year like no other? The impact of coronavirus on universal jurisdiction* (2021) < <https://trialinternational.org/latest-post/ujar-2021/>> accessed 17 May 2021

Trial International, Universal Jurisdiction Annual Review 2020 (UJAR 2020)- *Terrorism and international crimes: prosecuting atrocities for what they are* (2020) <https://trialinternational.org/wp-content/uploads/2020/03/TRIAL-International_UJAR-2020_DIGITAL.pdf> accessed 17 May 2021

Trial International, Universal Jurisdiction Annual Review 2019 (UJAR 2019)- *Evidentiary challenges in universal jurisdiction cases* (2019) <https://trialinternational.org/wp-content/uploads/2019/03/Universal_Jurisdiction_Annual_Review2019.pdf> accessed 17 May 2021

- **Sources**

Statutes and conventions

European Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by Protocols Nos 11 and 14, (adopted 4 November 1950, entered into force 3 September 1953) ETS 5

International Covenant on Civil and Political Rights (signed 16 December 1966, entered into force 23 March 1976) 999 UNTS 171

Statute of the International Tribunal for the Prosecution of Persons Responsible for Serious Violations of International Humanitarian Law Committed in the Territory of the Former Yugoslavia since 1991 (as amended on 17 May 2002), S/RES/808 (1993)

Statute of the International Criminal Tribunal for Rwanda (as last amended on 13 October 2006), S/RES/94 (1994)

Rome Statute of the International Criminal Court (adopted 17 July 1998, entered into force 1 July 2002) 2187 UNTS 3

Rules

Rules of Procedure and Evidence, ICC-ASP/1/3and Corr.1 (2013)

Rules of Procedure and Evidence, MICT/1/Rev.7 (2020)

Case law

Prosecutor v Nahimana (Judgment and Sentence) ICTR-99-52-T (3 December 2003)

Prosecutor v Milan Milutinović Nikola Šainović Dragoljub Ojdanić Nebojša Pavković Vladimir Lazarević Sreten Lukić, IT-05-87-T (10 October 2006)

Konstantina Despoina Stavrou

Prosecutor v Édouard Karemera Mathieu Ngirumpatse Joseph Nzirorera (Decision On The Prosecutor's Motion For Admission Of Certain Exhibits Into Evidence Rule 89(C) of the Rules of Procedure and Evidence) 98-44-T (25 January 2008)

Prosecutor v Setako (Judgment and Sentence) ICTR-04-081-T (25 February 2010)

Prosecutor v Bemba (Judgment on the Appeals of Mr Jean-Pierre Bemba Gombo and the Prosecutor against the decision of Trial Chamber III entitled "Decision on the admission into evidence of materials contained in the prosecution's list of evidence") ICC-01/05-01/08-1386 OA 5 OA 6 (3 May 2011)

Prosecutor v Lubanga (Judgment pursuant to article 74 of the Statute) ICC-01/04-01/06-2842 (14 March 2012)

Prosecutor v Tolimir (Judgment) IT-05-8/2-T (12 December 2012)

Prosecutor v Mathieu Ngudjolo (Judgment) ICC-01/04-02/12, (18 December 2012)

Prosecutor v Jean-Pierre Bemba Gombo, Aimé Kilolo Musamba, Jean-Jacques Mangenda Kabongo, Fidèle Babala Wandu and Narcisse Arido, ICC-01/05-01/13

Prosecutor v Bemba (Decision on Requests to Exclude Dutch Intercepts and Call Data Records) ICC-01/05-01/13-1855 (29 April 2016)

Prosecutor v Al Mahdi (Transcript) ICC-01/12-01/15-T-4-Red-ENG (22 August 2016)

Prosecutor v Al Mahdi (Judgment and Sentence), ICC-01/12-01/15-171 (27 September 2016)

Prosecutor v Ayyash et. al (Decision on the Admission of Call Sequence Tables Related to the Movements of Mr Rafik Hariri and Related Events, and Four Witness Statements) STL-11-01/T/TC (31 October 2016)

Prosecutor v Al-Werfalli (Warrant of Arrest) ICC-01/11-01/17 (15 August 2017)

Prosecutor v Ayyash et al (Judgment) STL-11-01/T/TC (18 August 2020)

UN Documents

UNOHCHR, HRC UC Berkeley School of Law, Berkeley Protocol on Digital Open Source Investigations- A Practical Guide on the Effective Use of Digital Open Source Information in Investigating Violations of International Criminal, Human Rights and Humanitarian Law (2020), HR/PUB/20/