

The Promise and Challenges of New Actors and New Technologies in International Justice

Federica D'Alessandra* and Kirsty Sutherland**

Abstract

This article addresses the role of new technologies in the international justice and accountability landscape, drawing from research we conducted into new United Nations (UN) accountability mechanisms that have the explicit mandate to collect, collate, analyse and preserve evidence of international crimes according to criminal justice standards. The article is divided in four parts. First, we contextualize our research by discussing some of our findings and situating them against what we define as a 'third wave' of institutional developments in international justice, prompted by an 'accountability-turn' affecting civil

* Federica D'Alessandra is Executive Director of the Oxford Programme on International Peace and Security, Institute for Ethics, Law and Armed Conflict, Blavatnik School of Government, University of Oxford, UK. She is also the immediate past Co-Chair of the International Bar Association War Crimes and Human Rights Law Committees. [Federica.dalessandra@bsg.ox.ac.uk]

** Kirsty Sutherland is Visiting Fellow of Practice, Oxford Programme on International Peace and Security, Institute for Ethics, Law and Armed Conflict, Blavatnik School of Government, University of Oxford, UK. She is also an officer of the International Bar Association War Crimes Committee, and a Barrister at 9 Bedford Row Chambers. [Kirsty.sutherland@bsg.ox.ac.uk]

The research for this article was conducted in connection with our project 'Anchoring Accountability for Mass Atrocities: Providing the Support Necessary to Fulfil International Investigative Mandates' at the University of Oxford. The project is the result of a partnership between the Oxford Programme on International Peace and Security, the International Bar Association War Crimes Committee, and the Simon Skjoldt Center for the Prevention of Genocide at the US Holocaust Memorial Museum, Washington, DC.

The authors are grateful to the editors of the journal and symposium for the opportunity to publish our research, and in particular to Alexa Koenig and Urmila Pullat for their input. We are also grateful to Lindsay Freeman for her insight; to our team members Ambassador Stephen J. Rapp and Sareta Ashraph for their contributions and data gathering over the summer; to all members of our project's Advisory Group for their guidance, and particularly to Cécile Aptel and David Akerson for their insight in framing some of the issues tackled in this piece; and to Ross Gildea for his editorial and research assistance.

society groups and UN mandates. Secondly, we discuss — using real-world examples — both the opportunities and challenges arising from the use of digital and new documentation technologies in the field. Thirdly, the article pays particular attention to the role of UN mandates affected by the ‘accountability-turn’; our research reveals such UN mandates now often sit at the heart of the ‘life cycle’ of information and evidence collected for justice and accountability purposes. In this section of the article, we also briefly discuss issues relating to third party control of information, in particular by social media companies. Finally, we discuss the need (and welcome initiative) to develop better international guidance and best practices for actors across the board in order to maximize the effective use of new technologies and digital evidence in international justice and accountability processes.

1. A Changing Landscape in International Justice

The past 10 years have been characterized by important shifts in the international justice and accountability landscape.¹ These include an increased exercise of universal or extraterritorial jurisdiction by domestic authorities; increased involvement of international courts and tribunals around issues of state responsibility for atrocities; and the emergence of new actors — including civil society groups and mandates of the United Nations (UN) — seeking to contribute directly to international justice and accountability processes, including criminal accountability.² We have termed this re-orientation of stakeholders an ‘accountability turn’.³ This trend has culminated, since 2016, in a ‘third wave’ of institutional developments within the international justice field,⁴ namely, the establishment of a ‘novel generation’ of UN accountability mechanisms, which have the explicit mandate to collect, collate, analyse and preserve evidence of international crimes — including in Syria, Myanmar and by the Islamic State of Iraq and Syria (ISIS) in Iraq — according to criminal justice standards, and to make this evidence available for domestic or international prosecutions.⁵ Though significant differences in mandate, jurisdiction,

1 By ‘international justice and accountability’ we do not refer exclusively to international criminal law and institutions, but rather to the broader field seeking to uphold accountability and fight impunity for gross violations of international human rights law, international crimes and other systemic abuse. This includes courts and tribunals, as well as non-judicial mechanisms. Where we intend international criminal justice, we will so specify.

2 F. D’Alessandra, ‘Anchoring Accountability for Mass Atrocities Through Stronger States’ Support of UN Investigative Mandates’, Oxford Programme on International Peace and Security (forthcoming 2021).

3 F. D’Alessandra, ‘The Accountability Turn in Third Wave Human Rights Fact-Finding’, 33 *Utrecht Journal of International and European Law* (2017) 59–76.

4 *Ibid.*

5 The International, Impartial and Independent Mechanism for Syria (IIIM) created by the UN General Assembly in December 2016 pursuant to UN GA Res. 71/248; the UN Investigative Team to Promote Accountability for Crimes Committed by Da’esh/ISIL (UNITAD), created by the UNSC in September 2017 pursuant to UN SC Res. 2379; and the International Independent Mechanism for Myanmar (IIMM) — the sister mechanism of the IIIM — created by the UN HRC in December 2018 pursuant to UN HRC Res. 39/2. These mechanisms however are not the only impacted by the ‘accountability-turn’. See D’Alessandra, *supra* note 3.

resources and operational realities distinguish these investigative mechanisms from one another, they may be conceptualized as 'quasi-offices-of-an-international-prosecutor-at-large', albeit not attached to a specific court.⁶

To understand the challenges and opportunities offered by these changing dynamics in the justice and accountability panorama, our team at the University of Oxford has been gathering significant data and research over the past few months.⁷ This includes interviews of 57 personnel working across the accountability ecosystem, from members of evidence-gathering organizations to national and international investigators and prosecutors,⁸ and an anonymous survey of 103 UN staff members working with mandates affected by the 'accountability-turn'.⁹ What emerges is a fascinating picture.¹⁰

Data from our interviews indicate that as competent legal systems gain willingness and confidence to pursue cases against individuals under their jurisdiction, and as other accountability mechanisms are triggered to uphold state responsibility, UN mechanisms can perform important support functions, such as providing contextual analyses with regard to broader political and case circumstances; gathering both crime-base and linkage evidence (i.e. evidence that connects individuals with crimes); helping to collect, verify and collate information; and closing evidential/information gaps.¹¹ Prosecutorial authorities also recognize civil society documentation as an essential to their ability to investigate and try cases, especially where they cannot access crime scenes themselves. The increased 'professionalisation' of documentation efforts from civil society actors is also seen favourably by those ultimately seeking to

6 D'Alessandra, *supra* note 3.

7 See Oxford Institute for Ethics, Law and Armed Conflict, 'Anchoring Accountability for Mass Atrocities', 2020, available online at <https://www.elac.ox.ac.uk/moving-fact-finding-case-build-ing> (visited 27 November 2020).

8 We interviewed investigating and prosecuting authorities of all 12 domestic jurisdictions most active in universal jurisdiction cases of core international crimes, as well as investigators and prosecutors at the ICC, and 34 of the leading civil society groups documenting international crimes in Africa, the Middle East, Asia and Europe.

9 S. Ashraph, F. D'Alessandra and S. Rapp, 'Structural Challenges Confronted by UN Accountability Mandates: Perspectives from Current and Former Staff', *Opinio Juris*, 14 October 2020, available online at <http://opiniojuris.org/2020/10/14/structural-challenges-confronted-by-un-accountability-mandates-perspectives-from-current-and-former-staff-part-i/>; <http://opiniojuris.org/2020/10/14/structural-challenges-confronted-by-un-accountability-mandates-perspectives-from-current-and-former-staff-part-ii/>; <http://opiniojuris.org/2020/10/14/structural-challenges-confronted-by-un-accountability-mandates-perspectives-from-current-and-former-staff-part-iii/> (visited 28 November 2020).

10 For a discussion of preliminary findings, see F. D'Alessandra et al., 'Anchoring Accountability for Mass Atrocities: Providing the Support Necessary to Fulfil International Investigative Mandates', *Opinio Juris*, 18 September 2020, available online at <http://opiniojuris.org/2020/09/18/anchoring-accountability-for-mass-atrocities-providing-the-support-necessary-to-fulfil-international-investigative-mandates/>; <http://opiniojuris.org/2020/09/18/anchoring-accountability-for-mass-atrocities-providing-the-permanent-support-necessary-to-fulfil-international-investigative-mandates-part-ii/>; <http://opiniojuris.org/2020/09/19/anchoring-accountability-for-mass-atrocities-providing-the-permanent-support-necessary-to-fulfil-international-investigative-mandates-part-iii/> (visited 28 November 2020).

11 Rapp, D'Alessandra and Sutherland, *Part iii, ibid.*

leverage digital information in judicial proceedings.¹² Indeed, our data show that many actors increasingly think of international justice as an ‘evolving landscape’, a complex system of information where streams of data originate with civil society groups, are processed in some form by UN mandates, and then arrive with judicial and non-judicial authorities.¹³ Interestingly, civil society actors consistently rated ‘criminal accountability’ as either their ‘top priority’ or a ‘very high priority’ for their documentation efforts, and often looked to UN mandates to provide qualitative guidance in terms of what information is most relevant for these purposes, although they also valued capacity-building initiatives aimed at supporting the long-haul fight against impunity.¹⁴

Of course, challenges persist. These range from discrepancies in operational and methodological approaches, to the quality and quantity of information collected. These challenges are relevant irrespective of which form of accountability is pursued. In fact, the true utility of information collected might not be apparent at the time of gathering. However, consensus is emerging that actors should seek to uphold the highest methodological and procedural standards to ensure that, if information might ever become relevant to criminal proceedings, its probative value will not be lost or compromised.¹⁵ Our data also showed clear consensus that, as more actors and mandates operate within or around the same area, the provision of better coordination and guidance becomes crucial. Indeed, the UN Office of the High Commissioner for Human Rights (OHCHR) has recognized the crucial role that UN mandates can play in supporting accountability purposes, and has already begun to build an investigations support unit.¹⁶

Most fascinatingly, these developments have coincided with a transformation of the information landscape that has the potential to revolutionize the field. The Internet has dramatically altered conditions of speech, becoming the latest

12 *Ibid.*

13 D'Alessandra et al., *Part i*, *supra* note 10.

14 D'Alessandra et al., *Part ii*, *supra* note 10.

15 F. D'Alessandra et al., *Handbook on Civil Society Documentation Serious Human Rights Violations*, Public International Law and Policy Group (2016), available online at https://static1.squarespace.com/static/5900b58e1b631bffa367167e/t/59dfab4480bd5ef9add73271/1507830600233/Handbook-on-Civil-Society-Documentation-of-Serious-Human-Rights-Violations_c.pdf (visited 28 November 2020); also see Group of Practitioners on Fact-Finding and Accountability, ‘Bridging The Hague – Geneva Divide: Recommendations to Maximize Benefit and Minimize Harm for Human Rights Inquiries and Criminal Investigations at the Same Scenes of Mass Violence’, 6 January 2017, available online at <https://www.elac.ox.ac.uk/files/bridgingthehague-genevadiide-finalrecommendations6jan2017revpdf> (visited 28 November 2020).

16 The establishment of an investigations support unit, with the support of the Dutch government, is consistent with recommendation 1 of the Bridging The Hague – Geneva Divide recommendations. See Group of Practitioners on Fact-Finding and Accountability, *ibid.*, 15. For more background on this, see <https://www.elac.ox.ac.uk/moving-fact-finding-case-building>; also see Ashraph, D'Alessandra and Rapp, *supra* note 9 and D'Alessandra et al., *supra* note 10.

medium for the recording, publication, and indeed commission, of crimes.¹⁷ Even when not purposively recording our actions, the digital footprint we leave behind by virtue of the technology we carry around, from our phones to our digital watches, is inescapable. In addition, intelligence sources and capabilities once the exclusive tool of states, are now publicly available to civil society groups and individuals on the frontlines, providing the opportunity to capture or prove the commission of crimes extemporaneously, even without stepping foot into a crime scene.

Based on our data, four ‘new technologies’ are chief candidates to propel the international justice field into its new frontiers¹⁸: (i) geospatial intelligence and remote sensing (GEOINT);¹⁹ (ii) online open source intelligence (OSINT);²⁰ (iii) financial intelligence (FININT);²¹ and (iv) documentation technologies, ranging from simple photo cameras to specialized software — such as the ground-breaking ‘eyeWitness to Atrocities’²² application — that allows for the recording of photos and videos and packages those items with the relevant metadata needed to demonstrate their authenticity in court. Of course, the more tools become available to those seeking to document violence, the more guidance on how these tools can be leveraged ethically and lawfully becomes necessary. Efforts such as the Berkeley Protocol on Open-Source Investigations, developed by the Berkeley Human Rights Center and the UN Human Rights Office, are a critical contribution to this rapidly developing field.²³ The conversation,

17 E. Irving, ‘Suppressing Atrocity Speech on Social Media’, 113 *American Society of International Law* (2019) 256–261.

18 F. D’Alessandra, S. Raj Singh and S. Rapp, ‘Atrocity Prevention in a Transatlantic Setting: A Paper on the Need to Foster Knowledge in Governmental and Non-governmental Experts’, Oxford Programme on International Peace and Security, June 2020, available online at <https://www.elac.ox.ac.uk/files/atrocitypreventioninatransatlanticsetting-finalpdf> (visited 28 November 2020).

19 Such as satellite imagery through Google Earth, USGS EarthExplorer, LandViewer, Copernicus Open Access Hub and NASA’s Earthdata Search, to name a few.

20 Open source intelligence (OSINT) ‘... refers to a subcategory of open source information that is collected and used for the specific purpose of aiding policymaking and decision-making, most often in a military or political context’. See UC Berkeley Human Rights Center and OHCHR, *Berkeley Protocol on Digital Open-Source Investigations* (2020), available online at https://www.ohchr.org/Documents/Publications/OHCHR_BerkeleyProtocol.pdf (visited 28 November 2020), at 7–8. By OSINT, we refer to all data that is publicly available and can be collected and shared without breaking laws or policies, needing a warrant, or participating in unethical practices. This includes: videos of events streamed live on YouTube, Facebook and other social media platforms; audio clips found online; and communications, instructions and commands posted on social media.

21 For example, Chainalysis’s KYT and Reactor, and Elliptic’s Navigator and Forensic software, which are being used by government agencies to monitor and investigate tor network and cryptocurrency financial flows. See Chainalysis, ‘Chainalysis Professional Services’, available online at <https://www.chainalysis.com/professional-services/> (visited 26 March 2021); Elliptic, ‘Bringing Compliance to Cryptoassets’, available online at <https://www.elliptic.co> (visited 26 March 2021).

22 See <https://www.eyewitness.global/welcome>. EyeWitness to Atrocities’ ground-breaking work is followed by new projects aimed at capturing, storing and transmitting court-standard secure, signed digital material, such as ProofMode and Tella. See at <https://guardianproject.info/apps/org.witness.proofmode/> and <https://tella-app.org/> (visited 30 April 2021).

23 UC Berkeley Human Rights Center and OHCHR, *supra* note 20.

however, is only just starting. From the collection of explosive information challenging official narratives, to the verification and corroboration of information and evidence, to the co-opting of new technologies for preventive purposes, the opportunities presented by closed and open-source digital material are worth exploring in detail.

2. Understanding the Promise and Potential of New Technologies in the International Justice and Accountability Ecosystem

The advantages of new technologies, digital documentary and investigative methods, can be significant.²⁴ Collecting and transmitting evidence digitally have been lauded as having the potential to ‘democratise the process of human rights fact-finding’ by opening it up to ‘ordinary people.’²⁵ This is because such methods are open to anyone with the right equipment — a smart-phone, a camera, a Facebook account or similar — and access to the Internet. Open-source investigation has the power to overcome or at least mitigate biases identified in ‘traditional’ investigative methods, for example, selection bias in witness interviews, or confirmation bias in investigative decisions. Authentic photographic evidence contains more visual details more accurately than eye-witness testimony ever could. In addition, because digital technologies can theoretically be used anywhere, they permit meaningful monitoring, documentation and investigation of possible atrocities by investigators remotely even for so-called ‘black hole’ environments, where information is deliberately hidden by local authorities or otherwise scarce.

This information cannot, of course, entirely substitute more traditional forms of evidence. However, as discussed below, it can act as a ‘force multiplier’ for other evidence. For this reason, it can and has infinitely expanded its potential uses, including but not limited to judicial and non-judicial accountability, including criminal and civil proceedings, other transitional justice strategies, truth and reconciliation efforts, memorialization and restorative justice processes.

GEOINT documentary technologies — including satellite imagery, radio, radar and other forms of remote sensing capacity, such as commercial unmanned aerial vehicles — can be used alongside OSINT methodologies to

²⁴ See, for example, J. Deutch and N. Para, ‘Targeted Mass Archiving of Open Source Information: A Case Study’, in S. Dubberley, A. Koenig and D. Murray (eds), *Digital Witness: Using Open Source Information for Human Rights Investigation, Documentation, and Accountability* (Oxford University Press, 2020) 163–184; A. Koenig, ‘Open Source Evidence and Human Rights Cases: A Modern Social History’, in Dubberley, Koenig and Murray (eds), *ibid.*, 32–47. Also 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* (2018) 283–335.

²⁵ M. Land, ‘Democratizing Human Rights Fact-Finding’, in P. Alston and S. Knuckey (eds), *The Transformation of Human Rights Fact-Finding* (Oxford University Press, 2016) 399–417, at 402.

corroborate human testimony. This includes, for example, by helping to identify the construction and use of military installations and other buildings, track the movement of convoys and document incidents. Satellite imagery has shown the trail of death and destruction left by militias and military forces in Sudan, South Sudan, the Central African Republic, the Democratic Republic of the Congo,²⁶ Niger and Myanmar.²⁷ Contradicting Chinese government claims, GEOINT has indicated the rapid recent development of detention camps holding an estimated 1–1.5 million Uyghurs and other Turkic minorities in Xinjiang.²⁸ Similarly, satellite imagery, corroborated by witness affidavits, has not only unveiled the operation and reporting structures of special political prisoners' camps (*kwanli'so*), north of Pyongyang, where 80,000–130,000 people are currently being detained, but also established the exact location of camps 14, 15, 16, 18, 22 and 25, the very existence of which the North Korean government has vehemently denied.²⁹ While measures can be taken to 'fool' satellites, and simple 'birds eye' views may not adequately show destruction or erosion of building complexes, the prevalence, independence and consistency of satellite oversight — no longer the preserve of powerful state authorities — renders it more difficult to sustain denials of wrongdoing.

New technologies can also be used for early-warning and preventive purposes.³⁰ FININT, GEOINT and OSINT (including social network and big data analysis) can assist with tracking the movement of individuals or groups (including militias and refugees), or even the 'mood' of specific groups³¹ — sometimes being able to predict with amazing precision the outbreak and location of identity-based protests or other atrocity risk factors.³² FININT can also assist with criminal deterrence (through the freezing of assets and other financial sanctions, the issuing of travel bans and the exposure of money laundering³³), thus diminishing perpetrators' agency and structural

26 See, for example, The Sentry, 'Fingerprints and Money Trails', available online at <https://the-sentry.org> (visited 26 March 2021).

27 B. Strick, *How to Identify Burnt Villages by Satellite Imagery — Case-Studies from California, Nigeria and Myanmar*, Bellingcat, 4 September 2018, available online at <https://www.bellingcat.com/resources/how-tos/2018/09/04/identify-burnt-villages-satellite-imagery%E2%80%8A-case-studies-california-nigeria-myanmar/> (visited 28 November 2020).

28 C. Buckley and A. Ramzy, 'Night Images Reveal Many New Detention Sites in China's Xinjiang Region', *New York Times*, 24 September 2020, available online at <https://www.nytimes.com/2020/09/24/world/asia/china-muslims-xinjiang-detention.html> (visited 20 November 2020).

29 International Bar Association, *Report: Inquiry on Crimes Against Humanity in North Korean Political Prisons*, December 2017, available online at <https://www.ibanet.org/IBA-War-Crimes-Committee--Inquiry-on-Crimes-Against-Humanity-in.aspx> (visited 28 November 2020) at 21–26.

30 Irving, *supra* note 17.

31 R. Rotberg, 'Deterring Mass Atrocity Crimes: The Cause of Our Era', in R. Rotberg (ed.) *Mass Atrocity Crimes: Preventing Future Outrages* (Brookings International Press, 2010) 1–24.

32 C. Mahony, E. Albrecht and M. Sensoy, *The Relationship Between Influential Actors' Language and Violence: A Kenyan Case Study Using Artificial Intelligence*, Commission on State Fragility, Growth and Development (2019), available online at https://www.theigc.org/wp-content/uploads/2019/02/Language-and-violence-in-Kenya_Final.pdf (visited 28 November 2020).

33 The Sentry, *supra* note 26.

opportunities. Finally, by bringing hidden and underrepresented voices to the fore, the ‘democratising’ effect of technology³⁴ can overturn state narratives, allow greater oversight and transparency, and significantly contribute to shaping public policy. Within international justice, these technologies hold enormous potential to ensure public, political and criminal accountability for atrocities. In the following sub-sections, we discuss some illustrative and practical examples.

A. Ensuring Public Transparency and Accountability

One example of the use of new technologies and digital documentation to promote public transparency and accountability is the work of the London-based watchdog ‘Airwars’, which monitors and documents allegations of civilian harm caused by international military airpower in Syria, Iraq, Libya, Yemen and Somalia. Airwars monitors open source material including relevant media and social media, uploaded footage of international strikes, local casualty-monitoring reports and militant propaganda; by doing so, it has been able to highlight stark divergences from official casualty counts reported by state militaries, including the anti-ISIS US-led Coalition.³⁵ Challenges to official reports posed by Airwars and others seem to have encouraged greater military transparency, at least with coalition-powers that engage with such findings. The Pentagon, for example, reported that in ‘continu[ing] to refine its practices and procedures for reviewing reports of civilian casualties, ... it considers reports available from any source, including after-action reporting of military units, and information provided by external sources, such as NGOs, the news media, social media and individuals who were present during the operation, including military personnel and local civilians’.³⁶ In addition to promoting improvements in standards, the publication of these reports also vastly

34 A. Powell, ‘Democratizing Production through Open Source Knowledge: from Open Software to Open Hardware’, 34 *Media, Culture & Society* (2012) 691–708; F. Gilardi, ‘Digital Democracy: How Digital Technology is Changing Democracy and Its Study’, Working Paper, 18 August 2016, available online at <https://www.fabriziogilardi.org/resources/papers/Digital-Democracy.pdf> (visited 28 November 2020).

35 S. Oakford, *Credibility Gap: United Kingdom Civilian Harm Assessments for the Battles of Mosul and Raqqa*, Airwars, September 2018, available online at <https://airwars.org/wp-content/uploads/2018/09/UK-Inquiry-into-Mosul-and-Raqqa-2018.pdf> (visited 28 November 2020); Airwars, *US Military Assessments of Civilian Harm: Lessons Learned from the International Fight Against ISIS*, March 2019, available online at <https://airwars.org/wp-content/uploads/2019/03/Airwars-2019-Interim-Better-Practice-Recommendations-for-DoD.pdf> (visited 28 November 2020). For example, while the US-led Coalition conceded 1,410 deaths in Iraq and Syria, Airwars found at least 8,310 ‘confirmed or fair’ such deaths — with estimates possibly going as high as 13,187, with locally reported allegations numbers rising to 29,639. See Airwars, *US-Led Coalition in Iraq and Syria*, available online at <https://airwars.org/conflict/coalition-in-iraq-and-syria/> (visited 28 November 2020).

36 US Department of Defense, *Annual Report on Civilian Casualties in Connection with United States Military Operations in 2019*, 29 April 2019, available online at <https://media.defense.gov/2019/May/02/2002126767/-1/-1/1/ANNUAL-REPORT-CIVILIAN-CASUALTIES-IN-CONNECTION-WITH-US-MILITARY-OPERATIONS.PDF> (visited 28 November 2020) at 15; Airwars, *ibid.*

improves public (and even military) understanding of the real impact of military actions, and affords those affected by the conflict better information regarding the extent of the harm suffered.³⁷ This information can, eventually, lead to efforts to repair or redress that harm.³⁸

B. Challenging False Narratives and Upholding State Responsibility

In addition to promoting public transparency and accountability, new investigative and documentation technologies can assist with attributing wrongful international acts and upholding state responsibility. One such example is the work of Bellingcat and open-source investigative journalists in eastern Ukraine. By determining the locations of videos and photographs posted online, investigators have been able to determine and confirm that it was a Russian military Buk missile — launched from Donetsk on 17 July 2014 — that hit Malaysia Airlines Flight MH17 as it flew over Ukraine in 2014, killing all 283 passengers (most whom were Dutch nationals) and 15 crew members.³⁹ The official Dutch-led Joint Investigation Team (JIT) investigation into the downing of MH17 relied on digital open-source information, alongside forensic assessments; witness and expert testimony; radar data; satellite imagery; and telecommunications records and data analysis. On the basis of these findings, the Dutch government challenged Russia before the European Court of Human Rights for its alleged role and its failure to conduct an appropriate investigation into the actions of the paramilitary groups it supported.⁴⁰ Bellingcat was also able to discredit alternative scenarios presented by the Russian Ministry of Defence,⁴¹

37 Airwars, *The Credibles*, available online at <https://airwars.org/conflict-data/the-credibles/> (visited 28 November 2020).

38 For example, open-source intelligence proved vital in the investigation into the downing of Ukraine Airlines passenger flight PS752, shot down by Iranian forces outside Tehran in January 2020. Iran eventually created a compensation fund to pay the families of the 176 victims \$150,000 for each victim. See C. Stoker-Walker, 'How Digital Sleuths Unravelling the Mystery of Iran's Plane Crash', *Wired*, 13 January 2020, available online at <https://www.wired.co.uk/article/iran-plane-crash-news> (visited 4 March, 2021).

39 Bellingcat, *MH17: The Open-Source Evidence*, available online at <https://www.bellingcat.com/app/uploads/2015/10/MH17-The-Open-Source-Evidence-EN.pdf> (visited 28 November 2020), at 23.

40 Government of the Netherlands, *The Netherlands Brings MH17 Case Against Russia before European Court of Human Rights*, Ministry of Foreign Affairs, 10 July 2020, available online at <https://www.government.nl/latest/news/2020/07/10/the-netherlands-brings-mh17-case-against-russia-before-european-court-of-human-rights> (visited 28 November 2020); European Court of Human Rights, 'New Inter-State application brought by the Netherlands against Russia concerning downing of Malaysia Airlines flight MH17', 213 ECHR (2020), 15 July.

41 Bellingcat, *A Post Mortem of Russia's Claim that Crucial MH 17 Video Evidence was Falsified*, 10 March 2020, available online at <https://www.bellingcat.com/news/2020/03/10/a-post-mortem-of-russias-claim-that-crucial-mh17-video-evidence-was-falsified/> (visited 28 November 2020); Again relying on open source imagery, Bellingcat investigators were able to debunk a Russian assertion that a video posted on YouTube on 17 July 2014 showing the Russian BUK Telar near the missile launch site was not authentic.

and later to identify those likely responsible for the downing of the flight.⁴² When the Ukrainian Security Service and the JIT publicly released telephone intercepts, independent investigators were able to assess (and confirm) their authenticity against open source material, leading to the identification of further potential suspects.⁴³

In a similar manner, while the Syrian government denies ever having used chemical weapons, analyses of digital open-source information and declassified intelligence allowed the French government to determine ‘with a high degree of confidence’ that lethal chemical attacks occurred in Douma, Syria, at the hands of the Syrian regime.⁴⁴ In reaching this conclusion, French authorities relied on reports issued by civil society organizations as well as media, authenticating and verifying these through forensic examination and against witness testimony, finding that ‘the spontaneous circulation of these images across all social networks confirms that they were not video montages or recycled images’.⁴⁵ While assertions by states seeking to attribute international wrongful conduct to other states must always be handled with care, the references to civil society and open source material in both the French and Dutch/JIT reports demonstrate very clearly the value of publishing on-the-ground information in places otherwise inaccessible to international investigators. Further, the enforced methodological transparency meant that the French and JIT/Ukrainian claims could at least themselves be subject to rigorous scrutiny.

In another example, when pursuing its case against Myanmar for violations of the Convention on the Prevention and Punishment of the Crime of Genocide at the International Court of Justice (ICJ), The Gambia relied on satellite

42 *Ibid.* Also see N. Beauman, ‘How to Conduct an Open-Source Investigation, according to the Founder of Bellingcat’, *The New Yorker*, 30 August 2018, available online at <https://www.newyorker.com/culture/culture-desk/how-to-conduct-an-open-source-investigation-according-to-the-founder-of-bellingcat> (visited 28 November 2020).

43 Bellingcat, ‘“A Birdie is Flying Towards You”: Identifying the Separatists Linked to the Downing of MH17’, 19 June 2019, available online at <https://www.bellingcat.com/app/uploads/2019/06/a-birdie-is-flying-towards-you.pdf> (visited 28 November 2020).

44 French Ministry for Europe and Foreign Affairs, ‘Chemical Attack of 7 April 2018 (Douma, Eastern Ghouta, Syria): Syria’s Clandestine Chemical Weapons Programme’, 14 April 2018, available online at <https://www.diplomatie.gouv.fr/en/country-files/syria/news/article/national-assessment-document-on-chemical-attack-of-7-april-2018-douma-eastern> (visited 28 November 2020) at 2–3; B. Hubbard, ‘Dozens Suffocate in Syria as Government is Accused of Chemical Attack’, *The New York Times*, 8 April 2018, available online at <https://www.nytimes.com/2018/04/08/world/middleeast/syria-chemical-attack-ghouta.html> (visited 28 November 2020).

45 French Ministry for Europe and Foreign Affairs, *supra* note 44. The same conclusions reached by the French government with regard to the use of chemical weapons in Syria, was also independently reached by a fact-finding mission deployed by the Organization for the Prohibition of Chemical Weapons (OPCW). The Syrian government denied OPCW access to inspection sites and Russia opposed the renewal of the UN–OPCW Joint Investigative Mechanism’s mandate for attributing responsibility to the Syrian regime. See OPCW, *Report of the Fact-Finding Mission Regarding the Incident of Alleged Use of Toxic Chemicals as a Weapon in Douma, Syrian Arab Republic*, 7 April 2018, available online at <https://www.opcw.org/sites/default/files/documents/2019/03/s-1731-2019%28e%29.pdf>, (visited 28 November 2020); Security Council Report, *In Hindsight: The Demise of the JIM*, 28 December 2017, available online at https://www.securitycouncilreport.org/monthly-forecast/2018-01/in_hindsight_the_demise_of_the_jim.php (visited 28 November 2020).

imagery analysed by the UN Fact-Finding Mission on Myanmar showing ‘irrefutable documentation of the scale of destruction perpetrated’ against the Rohingya by the Tatmadaw. Digital opensource material was also leveraged by The Gambia to support claims of ‘genocidal intent’.⁴⁶ Likewise, the Dutch and Canadian governments recently announced their intentions to hold Syria accountable for violations of the UN Convention Against Torture.⁴⁷ The decision relies heavily on documentary evidence provided by Syrian military police defector, ‘Caesar’, who smuggled 55,000 photographs of emaciated and visibly battered bodies out of Syria on a USB stick. The photographs were not ‘evidentiary’ photographs in the style of those taken in homicide investigations. They lacked, for example, images of the backs of bodies, scales, close-ups or any information about internal injuries or disease. Despite the limitations that this placed on the forensic analysis, experts in forensic digital imagery, anthropology and medicine were able to conclude that the images had not been digitally altered, and that the prevalence of emaciation and injuries consistent with torture would support findings of systematic torture and killing of detained persons amounting to war crimes and crimes against humanity.⁴⁸

At a minimum, as the above examples demonstrate, open-source investigations by civil groups can play an important role in helping to counter false state narratives. Importantly, they can also play a role in assisting judicial proceedings seeking to uphold state responsibility.

C. Pursuing Individual Responsibility

Finally, as new technologies and open-source investigations assist inter-state proceedings, they may assist the pursuit of individual responsibility. The attack on the USA capitol on 6 January 2021, during which five people — including

46 *Application Instituting Proceedings and Request for Provisional Measures (The Gambia v. Myanmar)*, 11 November 2019, available at <https://www.icj-cij.org/public/files/case-related/178/178-20191111-APP-01-00-EN.pdf> (visited 28 November 2020). The FFM Myanmar scrutinised ‘the vast extent of Myanmar’s hate campaign against the Rohingya group’ by examining, inter alia, Facebook posts and audio–visual materials. The Gambia relied on such findings to argue that Myanmar mounted a ‘pervasive campaign of dehumanisation’ and noted that ‘[t]he head of cybersecurity policy at Facebook said the company had found “clear and deliberate attempts to covertly spread propaganda that were directly linked to the Myanmar military”’.

47 Government of The Netherlands, ‘The Netherlands holds Syria Responsible for Gross Human Rights Violations’, 18 September 2020, available online at <https://www.government.nl/latest/news/2020/09/18/the-netherlands-holds-syria-responsible-for-gross-human-rights-violations> (visited 28 November 2020); Government of Canada, ‘Minister of Foreign Affairs Takes Action on Syria’s Human Rights Violations’, 4 March 2021, available online at <https://www.canada.ca/en/global-affairs/news/2021/03/minister-of-foreign-affairs-takes-action-on-syrias-human-rights-violations.html> (visited 5 March 2021).

48 *A report into the credibility of certain evidence with regard to torture and execution of persons incarcerated by the current Syrian regime*, UN Doc. S/2014/244, 4 April 2014. The pictures were independently scrutinized by digital imagery, forensic anthropology, and medical experts who were unaware of the source, but aware that the images were alleged to have been produced during an armed conflict and thus injuries could have resulted from lawful military action.

one police officer — lost their lives, is perhaps one of the most recent and emblematic examples. As a violent mob of rioters, supporters of then-President Trump, stormed the US Congress with the intent to intimidate lawmakers and interrupt the certification of the election of President Joe Biden, they filmed and posted videos of themselves, leaving a digital record of their identities and involvement in the ensuing violence.⁴⁹ This information was collated by open-source investigators, who cooperated with the Federal Bureau of Investigations — which had posted digital billboards soliciting potential leads from the public in the aftermath of the attack⁵⁰ — leading to over 100,000 pieces of digital evidence now assisting with serving hundreds of indictments.⁵¹

At the international level, open source investigations, including social network analysis, allowed the opening of a civil lawsuit in the USA against the former Sri Lanka Secretary of Defence for his alleged role in the killing and persecution of journalists.⁵² Alongside the above-mentioned ‘Caesar’ pictures, open source materials are also supporting criminal proceedings for international crimes against Syrian government officials in six European jurisdictions.⁵³ Similarly, the above-cited work of open source investigators in eastern Ukraine, combined with the JIT’s own findings, have led to criminal proceedings against Ukrainian and Russian commanders of the ‘Donetsk People’s Republic’, and investigations into individuals within the Russian Federation’s chain of command.⁵⁴ Digital open-source materials have also formed the basis for the International Criminal Court (ICC) arrest warrants issued for Mohamad Al Werfalli,⁵⁵ supported the guilty plea conviction of Ahmad al-Faqi al-Mahdi,⁵⁶ and led to the trial of four soldiers in Cameroon for the brutal killing of two women and two young children in Cameroon.⁵⁷ In addition, documentary evidence collected using the ‘eyeWitness’ technology — which is not, however,

49 R. Goodman and J. Hendrix, “‘Fight for Trump’: Video Evidence of Incitement at the Capitol”, *Just Security*, 25 January 2021, available online at <https://www.justsecurity.org/74335/fight-for-trump-video-evidence-of-incitement-at-the-capitol/> (visited 28 November 2020).

50 Federal Bureau of Investigations, ‘U.S. Capitol Violence: FBI Seeking Information Related to Violent Activity at the U.S Capitol Building’, available online at <https://www.fbi.gov/uscapitol1> (visited 28 November 2020).

51 The United States’ Attorney’s Office, District of Columbia, ‘Capitol Breach Cases’, available online at <https://www.justice.gov/usao-dc/capitol-breach-cases> (visited 28 November 2020).

52 *Wickrematunge v. Rajapaksa*, US District Court Central District of California, Case No. 2:19 CV-02577-R-RAO.

53 K. Aksamitowska, ‘Evidence in Domestic Core International Crimes Prosecutions: Lessons Learned from Germany, Sweden, Finland and The Netherlands’, in this Special Issue of the *Journal*.

54 Government of the Netherlands, *Government Informs UN Security Council of MH17 Trial*, Ministry of Foreign Affairs, 3 March 2020, available online at <https://www.government.nl/topics/mh17-incident/news/2020/03/06/government-informs-un-security-council-on-mh17-trial> (visited 28 November 2020).

55 Warrant of Arrest, *Al-Werfalli* (ICC-01/11-01/17-2), 15 August 2017, §§ 11–22.

56 Judgment and Sentence, *Al Mahdi* (ICC-01/12-01/15-171), Trial Chamber, 26 September 2016, §§ 31–41.

57 ‘Cameroon Atrocity: Finding the Soldiers who Killed this Woman’, *BBC*, 23 September 2019, available online at <https://www.bbc.co.uk/news/av/world-africa-45599973> (viewer discretion on start of video) (visited 4 March 2021).

open source — has also ensured convictions in the Democratic Republic of the Congo for rebel militia of the Democratic Forces for the Liberation of Rwanda.⁵⁸

This non-comprehensive list is but one indication of the opportunities offered by new technologies with regards to the investigation of international law violations, including atrocity crimes. As others have pertinently observed, ‘the future of [accountability] will [inevitably] be intertwined with the advancement of technology’.⁵⁹ For this reason, it is imperative that anyone seeking to leverage these new technologies to document atrocities — whether states, UN mandates, courts and tribunals, or civil society groups — at a minimum understands the specific security needs and vulnerabilities of this kind of information. Operational protocols and information management tools will be necessary to allow the confident handling of information new and open-source technologies can offer.⁶⁰

In the following section, we turn to discussing some of the challenges presented by this material, and how those challenges can impact the ‘life cycle’ of digital information and evidence in support of accountability.

3. Important Challenges Presented by Digital Open-Source Material

If, as mentioned above, the opportunities offered by open source and other digital materials seem boundless, it is equally important to bear in mind that these information and documentation tools also present challenges. These range from the potential widening of the digital divide; to possible biases (towards the digitally visible, or emanating from intrinsic characteristics such as the mandates or identities of information providers); to challenges to authenticity, integrity and many others.⁶¹ The passing of this information between parties, and the need to rely on the cooperation of third parties (including, often, private sector entities designed to profit from data) — can significantly

58 Witness, *Use of Video Evidence Leads to Justice in Democratic Republic of Congo*, September 2018, available online at <https://www.witness.org/video-evidence-helps-lead-to-historic-conviction-in-democratic-republic-of-congo/> (visited 28 November 2020).

59 E. Piracés, ‘The Future of Human Rights Technology’, in M.K. Land and J.D. Aronson (eds), *New Technologies for Human Rights Law and Practice* (Cambridge University Press, 2018) 289–308 (cited by A. Koenig, ‘“Half the Truth is Often a Great Lie”: Deep Fakes, Open Source Information, and International Criminal Law. Symposium on Non-State Actors and New Technologies in Atrocity Prevention’, 113 *American Society of International Law* (2019) 250–255, at 251).

60 Significant contributions have already been made by the Berkeley Protocol. UC Berkeley Human Rights Center and OHCHR, *Berkeley Protocol*, *supra* note, 20.

61 Since it is open *only* to anyone with the right equipment and access to the Internet, digital open-source investigation is very much on one side of the ‘digital divide’. Many vulnerable people lack access to the tools necessary to capture and disseminate digital evidence is determined by social factors including security, gender, income, literacy and Internet-familiarity. See Y. McDermott Rees, A. Koenig and D. Murray, ‘Open-Source Information’s Blind Spots: Human and Machine Bias in International Criminal Investigations’, in this Special Issue of the *Journal*; A. Koenig and U. Egan, ‘Power and Privilege: Investigating Sexual Violence with Digital Open Source Information’, in this Special Issue of the *Journal*.

compromise the value of this information. Appropriate measures and standards are critical to ensure the integrity of information across its lifecycle, from collection to transmission, analysis, disclosure and long-term storage. The breach or failure of protocols and other safety measures can be fatal to criminal accountability efforts.

A. Collection and Sharing of Evidence

The challenges posed by new and documentary technologies begin with collection. Access to the Internet — critical to information sharing — is subject to economic and political dynamics, including prohibitive pricing and Internet-shutdowns. These impediments are of very real concern to UN mandates: in September 2019, for example, the Bangladeshi government banned 3G and 4G Internet and the use of SIM cards by Rohingya refugees in Cox's Bazar, citing security concerns, restoring Internet access only on 29 August 2020.⁶² This caused a gap in collection of evidence potentially crucial to the UN Independent Impartial Mechanism for Myanmar (IIMM).⁶³ Unequal access to Internet services can also foster biases in collection, in the same vein as collectors' agendas and mandates.

For example, in determining what is collected, civil society organizations may be influenced or constrained by funding conditions, for example, encouraging particular focus on certain violations over others. In addition, over-reliance on 'the observable' — a bias towards the visual in documentation — may neglect less visible violations such as gender-based violence, starvation, or crimes against children. Simultaneously, online misinformation and disinformation campaigns may shape understandings and narratives of situations, and thus influence witnesses' 'perceptions of events'⁶⁴ in ways that veer from the 'truth'. This means that deliberate strategies must be deployed not only to ensure that evidence-collection represents the landscape of a situation as comprehensively as possible, but that any potential for bias is addressed and accounted for. A strategic approach to investigations that pays particular attention to 'less visible' crimes, to monitoring of local reporting, and to reliable eyewitness and survivor accounts that can be accessed securely can assist in compensating for blind spots and asymmetric power balances. As when

62 S.M. Najmus Sakib, 'Internet, Mobile Network Restored for Rohingya Refugees', *Anadolu Agency*, 29 August 2020, available online at <https://www.aa.com.tr/en/asia-pacific/internet-mobile-network-restored-for-rohingya-refugees/1957098> (visited 28 November 2020).

63 There is, in fact, a real risk that the very existence of an investigation might prompt authorities to shut the Internet down, with clear ramifications for the effective provision of services as well as access to information.

64 The so-called Hawthorne effect. See also Y. McDermott, D. Murray and A. Koenig, 'Digital Accountability Symposium: Whose Stories Get Told, and by Whom? Representativeness in Open Source Human Rights Investigations', *Opinio Juris*, 19 December 2019, available online at <http://opiniojuris.org/2019/12/19/digital-accountability-symposium-whose-stories-get-told-and-by-whom-representativeness-in-open-source-human-rights-investigations/> (visited 28 November 2020).

locating, selecting and analysing any evidential material, vigilance regarding structural and cognitive biases, including awareness and care regarding algorithm design and input,⁶⁵ and a scrupulous scepticism regarding the authenticity of digital material⁶⁶ will all contribute to ensuring that information collected and shared with accountability mechanisms maintains its independence and integrity. Verification processes designed to minimize the introduction of bias into forensic analyses are being developed by specialized investigators.⁶⁷ International accountability institutions, including UN mandates, can play a valuable role in promoting such best practices.

Incorporating digital safeguards at the evidence collection stage is key, though not always practicable. The development of apps such as 'eyeWitness' goes a long way to overcome issues of security and integrity of material, allowing the instant sharing of on-the-ground data from the most inaccessible places. Our data show, however, that barriers of digital access, digital literacy and trust persist.⁶⁸ The risks of taking and uploading such material may be very high. For this reason, safety and security protocols — and the 'do no harm' principle — should always be of guidance.⁶⁹ In addition, much digital material is lost on damaged, stolen or seized devices, or deleted. This calls for 'contingency plans' to ensure that an authenticated copy (or original) of the material remains available.⁷⁰ For example, YouTube's community standards mean that 'violent or graphic' content will be removed; while exceptions apply to 'newsworthy' content, without back-up, it is inevitable that material potentially valuable to criminal investigators will be lost.⁷¹ For

65 A. Richardson, 'Elbowed off the Pavement', *London Review of Books*, 20 August 2020, available online at <https://www.lrb.co.uk/blog/2020/august/elbowed-off-the-pavement> (visited 28 November 2020); K. Hill, 'Wrongfully Accused by an Algorithm', *The New York Times*, 24 June 2020, available online at <https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html> (visited 28 November 2020).

66 Bellingcat, 'Ghost in the Machine: From Chad, A Case Study on Why You Shouldn't Blindly Trust Tech', 17 July 2020, available online at <https://www.bellingcat.com/resources/2020/07/17/ghost-in-the-machine-from-chad-a-case-study-on-why-you-shouldnt-blindly-trust-tech/> (visited 28 November 2020).

67 See, for example, European Network of Forensic Science Institutes, *Best Practice Manual for Forensic Image and Video Enhancement* (2018), available online at <https://enfsi.eu/wp-content/uploads/2017/06/Best-Practice-Manual-for-Forensic-Image-and-Video-Enhancement.pdf> (visited 28 November 2020).

68 Several of the evidence-collecting organizations we interviewed explained that security concerns prevented them from relying on Internet-reliant communication tools, such as eyeWitness, or even communicating with potential witnesses online. At the other end of the spectrum, some international agencies have developed their own software for the collection, storage and protection of information. In the highly sensitive circumstances of the Syrian conflict, the Syrian Civil Defence (White Helmets) carry and wear cameras; this evidence has been targeted by belligerents.

69 D'Alessandra et al., *supra* note 15.

70 *Ibid.* See also UC Berkeley Human Rights Center and OHCHR, *supra* note 20.

71 M. Pizzi, 'The Syrian Opposition is Disappearing from Facebook', *The Atlantic*, 4 February 2014, available online at <https://www.theatlantic.com/international/archive/2014/02/the-syrian-opposition-is-disappearing-from-facebook/283562/> (visited 28 November 2020); M. Ingram, 'Critics say Facebook is Erasing Pieces of History by Deleting Pages about the War in Syria',

material uploaded to social media sites, the retention and creation of metadata — including that of unpublished material — would be a simple means of safeguarding authentication ability.

B. Authentication and Verification Processes and Information Analysis

1. General Challenges

In a ‘post-truth’ world, the camera often lies. Like other forms of potential evidence, digital material may be manipulated or destroyed, and authorship hidden, denied or falsified. Disinformation campaigns (‘the systematic use of deliberately distorted information to manipulate an adversary’s decision-making elite, or public opinion’) exploit existing social divisions and biases, relying on ‘unwitting agents’ to spread false narratives.⁷² This became evident in the Syrian conflict: despite — or perhaps because of — being one of the most heavily recorded conflicts in history. Disinformation campaigns have used faked, doctored and plagiarized material to malign and discredit various actors.⁷³ International mechanisms scrutinizing conflicts involving major states are also likely to be targeted with disinformation and ‘false flag’ campaigns to distort and impede investigative efforts. Clarifying the truth in such an environment is extremely difficult and requires well-informed understanding of the digital realm and of local dynamics, as well as sophisticated wariness incorporated into review processes and the staffing practices of monitors.⁷⁴ By carefully examining metadata (where available), and/or conducting reverse image searches, investigators may be able to detect misinformation.

One example is the case of images published by international media that purported to show Buddhist monks burning Rohingya victims. The authenticity of the images was soon debunked when it was demonstrated they actually depicted the cremation of victims of China’s 2010 earthquake.⁷⁵ As noted above, however, news spreads fast on social media sites. Before fake information can be found, challenged and removed, it might be leveraged to incite

GIGAOM, 5 February 2014, available online at <https://gigaom.com/2014/02/05/critics-say-facebook-is-erasing-pieces-of-history-by-deleting-pages-about-the-war-in-syria/> (visited 28 November 2020).

72 K. Starbird, ‘Disinformation Campaigns are Murky Blends of Truth, Lies and Sincere Beliefs – Lessons from the Pandemic’, *The Conversation*, 23 July 2020, available online at <https://theconversation.com/disinformation-campaigns-are-murky-blends-of-truth-lies-and-sincere-beliefs-lessons-from-the-pandemic-140677> (visited 28 November 2020).

73 J. Guay and L. Rudnick, ‘Open Source Investigations: Understanding Digital Threats, Risks and Harms’, in Dubberley, Koenig and Murray (eds), *supra* note 24, 293–313, at 302.

74 B. Daragahi, ‘Misinformation Maligns Syrian Uprising’, *The Financial Times*, 2 April 2014, available online at <https://www.ft.com/content/9629a4b9-0b90-34e3-b308-5a3bc9961311> (visited 28 November 2020).

75 G. Koettl, D. Murray and S. Dubberley, ‘Open Source Investigation for Human Rights Reporting’, in Dubberley, Koenig and Murray (eds), *supra* note 24, 12–31, at 24.

hatred, or otherwise stoke the conditions for mass violence.⁷⁶ It is then critical evidence in itself.⁷⁷

In addition, 'deep fake'⁷⁸ technology presents extraordinary opportunities for disruption, falsification or fabrication of convincing 'footage' implicating or exonerating individuals, groups or militaries. As 'President Obama' (impersonated by comedian Jordan Peele) explained in a BuzzFeed video, 'we [have entered] an era in which our enemies can make anyone say anything at any point in time'.⁷⁹ Deep fake detection benefits from artificial intelligence (AI) that can forensically analyse digital material. For example, by training machine learning systems with ISIS propaganda videos until the AI could 'recognise' their distinctive features (perhaps imperceptible to human detection), machines have been taught to detect ISIS propaganda.⁸⁰

The high volumes of material required to train the system make this a feasible option for detecting falsified or fraudulent material in data-rich situations, such as Syria, but less so for situations with lower volumes of suitable 'training' material, such as the examples of 'information black holes' mentioned above. It is a pressing concern, early attempts to develop AI detection have not been wholly successful and are already outdated,⁸¹ and the Internet's pursuit of deep fake detection tools must be monitored closely.⁸² As international and domestic courts increasingly rely on digital open-source material, authentication protocols are needed that will allow actors at every stage of proceedings to ensure and assess that information's integrity and reliability. AI and other technological aides will be helpful for collecting, preserving, processing, authenticating and verifying digital material. However, algorithms and

76 Human Rights Council, *Report of the Independent International Fact-finding Mission on Myanmar*, A/HRC/39/64, 12 September 2018; Human Rights Council, *Report of the Detailed Findings of the Independent International Fact-Finding Mission on Myanmar*, A/HRC/39/CRP.2, 17 September 2018.

77 *The Gambia v. Myanmar*, *supra* note 46.

78 S. Cole, 'AI-Assisted Fake Porn is Here and We're All Fucked', *Motherboard*, 11 December 2017, available online at <https://www.vice.com/en/article/gdydm/gal-gadot-fake-ai-porn> (visited 28 November 2020); K. Hao, 'Memers are Making Deepfakes, and Things are Getting Weird', *MIT Technology Review*, 28 August 2020, available online at <https://www.technologyreview.com/2020/08/28/1007746/ai-deepfakes-memes/> (visited 28 November 2020).

79 BuzzFeedVideo, 'You Won't Believe What Obama Says in This Video!', *YouTube*, 17 April 2018, available online at <https://www.youtube.com/watch?v=cQ54GDm1eL0> (visited 28 November 2020).

80 N. Schick, *Deep Fakes and the Infocalypse* (Octopus Publishing, 2020), at 195; Faculty, *Stopping the Spread of Online Daesh Propaganda* (2020), available online at <https://faculty.ai/ourwork/identifying-online-daesh-propaganda-with-ai/> (visited 28 November 2020).

81 S. Cole, 'Glycat's AI Solution for Fighting Deepfakes Isn't Working', *Motherboard*, 19 June 2018, available online at <https://www.vice.com/en/article/ywe4qw/glycat-spotting-deepfakes-fake-ai-porn> (visited 28 November 2020); M. Land 'Democratizing Human Rights Fact-Finding,' in P. Alston and S. Knuckey (eds) *The Transformation of Human Rights Fact-Finding* (Oxford University Press, 2016) 399–417, at 399.

82 Deepfake, *Deepfake Detection Challenge* (2020), available online at <https://www.kaggle.com/c/deepfake-detection-challenge> (visited 28 November 2020); Jigsaw, *Creating Future-Defining Technology* (2020), available online at <http://jigsaw.google.com/issues/> (visited 28 November 2020).

machine-learning tools will have to be designed with acute care to minimize biases, and be subjected to frequent rigorous examination to ensure they remain as objective as possible.

2. Criminal Justice Challenges

International legal mechanisms must accept that the online realm and the nature of modern evidence will impact their structure, staffing, processes and equipment. In this, they must follow the lead of those pioneering open-source investigations, including intelligence agencies and journalists. Unless material is authenticated and verified at the earliest stages, it will potentially be impossible to rely on it at trial. Worse, it could contaminate proceedings. Criminal procedural standards — and the ethical obligations of lawyers and criminal investigators — are higher and stricter than the obligations upon journalists, intelligence services or ‘freelance’ investigators. Investigative mechanisms, international and domestic courts, and other judicial authorities will be able to satisfy their mandates only if material is collected, preserved and analysed in accordance with these evidential standards. For legal actors, verification is a process of weighing probative value, not simply a binary assessment of ‘true or false’. Non-digital evidence must also be proactively pursued, not least because the ‘voices’ of victims must not be eclipsed by over-reliance on the digital. Verification processes must still aim for the triangulation of documentary, physical and testimonial evidence.⁸³ This applies to accountability mechanisms across the board.

As courts adjust to digital open-source material, new norms may emerge regarding indicia of authenticity and reliability. For example, in a step hailed as a breakthrough for digital open-source investigations, the ICC relied heavily on videos posted on social media sites to issue an arrest warrant against Mahmoud al-Werfalli in August 2017.⁸⁴ In the second arrest warrant issued against him, the Pre-Trial Chamber held that ‘an expert report [...] prepared by a renowned, independent institute’ which ‘concluded that there no traces of forgery or manipulation in relation to locations, weapons or persons shown in the video’, and was corroborated by witness evidence, amounted to ‘sufficient indicia of authenticity’.⁸⁵ Similarly, in view of the varying evidentiary thresholds of national authorities and international tribunals,⁸⁶ the recently established international investigative mechanisms for Iraq, Syria and Myanmar are

83 A. Koenig, “‘Half the Truth is Often a Great Lie’: Deep Fakes, Open Source Information, and International Criminal Law”, Symposium on Non-State Actors and New Technologies in Atrocity Prevention’, 113 *American Society of International Law* (2019) 250–255, at 251.

84 Warrant of Arrest, *Al-Werfalli* (ICC-01/11-01/17-2), Pre-Trial Chamber I, 15 August 2017, at 11–22.

85 Second Warrant of Arrest, *Al-Werfalli* (ICC-01/11-01/17-13), Pre-Trial Chamber I, 4 July 2018, at 18.

86 In some jurisdictions, material will not be admitted into the case unless it meets very high standards of credibility, while in others all material will be admitted and probative value formally ascribed at a much later stage in proceedings.

inevitably being tasked with (and our data indicate often expected to)⁸⁷ assess the probative value of the material they gather. As ‘core processors’ of digital information, these mechanisms have begun to incorporate protocols and ‘state-of-the-art’ tech-infrastructure designed to support all types of accountability measures. Other mandate holders should be given the same instruments, as the work of the Myanmar FFM demonstrates, by working to the highest judicial standards, such bodies will be able to optimize the utility of their assessments and investigations, feeding directly into judicial proceedings.⁸⁸

As a potential approach to these issues, and taking instruction from open-source investigation groups,⁸⁹ we submit that one option would be to adopt the ‘tiered categories of probative weight’ approach to digital open-source material used for investigations. A tiered approach is more conscious of information asymmetries, digital divides and blind spots that can define some of the power structures of digital and open source material, recognizing that unverifiable sources or material may be an indication of a problem with an investigation rather than the credibility of the material.⁹⁰ Such an approach would be most helpful to national authorities, in part because it would provide a provisional assessment of probative value and would accommodate different judicial systems and rules of evidence. For example, information ‘conceded’ (i.e. the veracity of which is acknowledged by the belligerent/organization/individual to whom it is attributed) could be flagged as ‘confirmed’. Where multiple credible sources, including, for example, biographical information, photographic and/or video material forensically authenticate the same information, this could be assigned ‘high’ probative value. One tier below could be ‘fair’, where multiple credible sources have corroborated the information, but not to the strength and standard that would assign it ‘high’ probative value. Finally, where information comes from a single source, this should not be disregarded necessarily, but ‘kept’ on the back-burner and assigned ‘weak’ probative value — leaving only ‘discounted’ materials, that is, those for which allegation or preservation is not considered credible — for the ‘dustbin’. Our interviews indicate that domestic authorities would welcome such an approach.

C. Archiving and Storage

The archiving and preservation of digital and open-source material, of course, present its own challenges to integrity, and to loss protection (either physical, or of probative value). This raises challenges concerning the ‘security of premises and archives’, some of which are technological, operational and structural,

87 D’Alessandra et al., *Part i*, *supra* note 10.

88 *The Gambia v. Myanmar*, *supra* note 46.

89 Airwars, ‘Methodology’, 2020, available online at <https://airwars.org/about/methodology/> (visited 28 November 2020).

90 S. Dyer and G. Ivens, ‘What Would a Feminist Open Source Investigation Look Like?’ 1 *Digital War* (2020) 5–17.

and others which will require legislative and political solutions.⁹¹ For example, digital material held by private companies incorporated in various countries raises issues regarding ownership, vulnerability to state access — for example, through ‘search and seizure’ orders — transparency, privacy and public accountability. Material sourced electronically may also be monitored, raising very real security concerns for activists. Benefitting from UN privileges and immunities, the archives of UN mandates are ‘inviolable’ to states interference.⁹² Although, they must of course ensure that technical, physical and procedural safety measures are in place wherever this information is stored. These issues require urgent consideration, and call for a dedicated study and bespoke solutions. While exhausting these issues lies beyond what we can achieve here, we use the next section to address a few.

4. The Potential of Investigative Mechanisms and other UN Mandates

The legitimacy (both legal and political) of accountability measures relies on their accuracy and impartiality. This is particularly the case when criminal accountability is involved: the political, legal and reputational ramifications for trials and convictions based on false evidence can be devastating. The novel UN investigative mechanisms for Syria, Daesh and Myanmar — as well as, increasingly, other UN mandate holders — play an important role in this ecosystem of information as they centralize the exchange of data, whether the data originate from civil society sources or domestic and international authorities. For this reason, the mechanisms stand to play a crucial role in promoting the secure collection of authentic material; analysing and storing evidence; and encouraging and enabling the effective pursuit of accountability measures, including beyond the courtroom. By developing frameworks for co-operation with judicial and non-judicial mechanisms, these bodies can encourage reconciliation, or feed into reparation and restoration processes. Our interviews with evidence-providing organizations operating in unstable regions suggest an appetite for rule of law and judicial capacity building initiatives as a step towards more comprehensively combatting impunity.⁹³ However, it is

91 E. Fry, ‘The Nature of International Crimes and Evidentiary Challenges: Preserving Quality While Managing Quantity’, in E. van Sliedregt and S. Vasiliev (eds), *Pluralism in International Criminal Law* (Oxford University Press, 2014) 251–272.

92 United Nations, *Chapter III. Privileges and Immunities, Diplomatic and Consular Relations, etc.*, 13 February 1946, available online at https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg_no=III-1&chapter=3&clang=en (visited 28 November 2020).

93 D’Alessandra et al., *supra* note 10. For example, Syrian CSOs emphasized the long road ahead and the need to build local judicial and legal capacities in order to establish the foundation for fair, confident and transparent justice and accountability measures in Syria. Similarly, a South Sudanese organization noted the importance of demonstrating, to both the public and the authorities, the advantages of a fair and effective domestic system of accountability for human rights violations, emphasising that transitional justice must be locally owned.

especially with regard to individual accountability that these mandates can perform a precious — and unique — function.

Mandate holders also stand to serve a crucial role as repositories entrusted with the storage of material obtained at great personal risk, and tasked with its navigation, analysis and archiving. The consistent development of protocols with civil society actors would optimize the mapping of violations and abuses, and ensure the safe and secure provision of information within their possession. At the time of writing, a great step in this direction is constituted by the Protocol of Cooperation between the International, Impartial and Independent Mechanism and Syrian Civil Society Organizations participating in the Lausanne Platform. The protocol clarifies and encourages active engagement between the IIIM and Syrian NGOs, in order to facilitate the shared goals of pursuing justice, accountability and redress for victims.⁹⁴ The 'two-way dialogue' allows all parties to focus on maximizing their contributions, whether the IIIM providing support to civil society documentation and analysis, or affording the evidence providers better understanding as to what the most valuable evidentiary material is, and how to obtain and provide it in a way that meets criminal legal standards. Protocols, tailored-Memorandums of Understandings, and collaborative relationships may go a long way in establishing and maintaining trust to overcome security concerns regarding the transmission of digital material.

Downstream, that is, facing the 'consumer' end of this flow of information, the novel mechanisms — and other UN mandate holders with the support of OHCHR — could fulfil an equally important function. The development of protocols or cooperation agreements with national and international investigation and prosecuting authorities encourages and streamlines engagement between these parties, and ensures that the information collected, received and analysed by UN mandates supports the independent work of forensic investigators and criminal prosecutors. Our data indicate that awareness of the role that can be played by UN mandates is increasing because of the performance of novel investigative mechanisms, and that national authorities believe these mechanisms can play a useful role in providing linkage evidence and authoritative contextual analyses to guide their own case-building.⁹⁵ In view of the likely volumes of digital material (whether 'originally' digital or 'hard' evidence digitally reproduced), the ability of such bodies to provide rapid, useful responses to Requests for Assistance will be reliant on sophisticated legal software, meticulous digital archiving and well-considered 'tagging'. Our data also suggest that information provided to national court proceedings by UN bodies may — by virtue of its perceived 'stamp of UN approval' — have been accorded a higher level of trust than it would otherwise had it come from a

94 See The International, Impartial and Independent Mechanism, *Protocol of Cooperation Between the International, Independent and Impartial Mechanism and Syrian Civil Society Organisations Participating in the Lausanne Platform*, 3 April 2018, available online at <https://iiim.un.org/engagement-with-stakeholders/> (visited 26 March 2021).

95 European Network of Forensic Science Institutes, *supra* note 67.

different source.⁹⁶ This, of course, carries intrinsic risks where UN mandates do not employ proper authentication and verification procedures. At the same time, if such capabilities were provided by re-purposing some of the ‘state-of-the-art’ technology possessed by the novel investigative mechanisms — which our data indicates is a possibility⁹⁷ — these needs could be met with minimal effort, and the initial investment into such technical capabilities would be put to larger (and therefore more impactful) use. How this could occur is a question ripe for scrutiny.

A. Third Party Control of Digital Information and Evidence: The Thorny Question of Social Media Companies

The challenges posed to the ‘lifecycle’ of digital evidence and information by commercial third parties who possess such material but are not oriented towards using this material to support accountability are also significant. Indeed, the identification, retention and collection of material held by social media and other tech companies (SMTCs) is a pressing concern for accountability efforts, and must be addressed with urgency.⁹⁸ The case of Myanmar is illustrative. Despite being described as an ‘information black hole’ due to the Myanmar authorities’ stark exclusion of independent observers from Rakhine State, digital open-source material has enabled researchers to investigate remotely: Google Earth satellite imagery has confirmed attacks on specific villages; NASA’s active fire data recorded episodes of live fire as villages were razed; and images and videos uploaded to social media could be geo-located with satellite imagery and cross-corroborated with eyewitness accounts from Rohingya refugees.⁹⁹

Since its introduction in Myanmar in 2010, the Internet (predominately Facebook) became the medium for a wildfire of misinformation, hate speech and incitement to violence against the long-persecuted Rohingya minority, and is considered to have significantly contributed to the latest wave of extreme, genocidal, violence.¹⁰⁰ Facebook eventually announced that it had taken steps towards the ‘proactive detection and removal of hate speech’, removing 64,000 pieces of content in the third quarter of 2018 alone.¹⁰¹ However, to many, this was too little and too late.

⁹⁶ *Ibid.*

⁹⁷ *Ibid.*

⁹⁸ Human Rights Center, *Digital Lockers: Precedents for Archiving Social Media Evidence of Atrocity Crimes* (forthcoming 2021).

⁹⁹ *The Gambia v. Myanmar*, *supra* note 46.

¹⁰⁰ Human Rights Council, *UN Fact-Finding Mission on Myanmar Report*, A/HRC/39/64, 1310–1311, at 1345, which notes hundreds of social media accounts, pages and groups ‘regularly spreading hate speech’, and ‘particularly influential’ accounts with followers ranging from 10,000 to over a million, high follower engagement and frequent posting (daily or hourly).

¹⁰¹ Facebook, ‘An Independent Assessment of the Human Rights Impact of Facebook in Myanmar’, 5 November 2018 (updated 26 August 2020), available online at <https://about.fb.com/news/2018/11/myanmar-hria/> (visited 28 November 2020).

In addition, the deletion of such information, in particular, when detected by algorithms, has become an obstacle to accountability. Incendiary rhetoric posted online is primary evidence of obvious value to those investigating both state and individual responsibility for the apparent commission of international crimes. Its preservation is necessary for accountability, but is in the hands of private SMTs. While reporting that it is 'coordinating' with the IIMM 'to provide relevant information',¹⁰² Facebook remains entangled in legal wrangling with The Gambia, which seeks removed content for its ICJ case against Myanmar for failure to prevent genocide.¹⁰³ The plausibility of the capture, preservation and transmission of this material is indicated by the commitments Facebook has made to cooperate with the IIMM.¹⁰⁴ However, it remains to be seen how effective this is, and what improvements might be necessary.

In another example, of 7,872,684 videos removed by YouTube in July–September 2020 for violating its Community Guidelines, 7,390,963 (93%) were removed after automatic detection. It remains unclear, however, if and how this information is being preserved to support accountability related to the commission of atrocities.¹⁰⁵ Digital open-source material is promoted or deleted subject to the internal rules and processes designed for commercial purposes. Reliance on algorithms to regulate public speech invokes obvious concerns regarding freedom of expression, the opacity of algorithmic design (and biases contained therein), and the state-like powers left in the hands of private entities. To go back to the above-mentioned January 2021 US Capitol attacks, in the wake of the violent insurrection, platforms such as Facebook, Twitter, and Instagram (which is owned by Facebook) suspended the accounts of then-US President Donald Trump, alleging his role in inciting the riots.¹⁰⁶ The suspension has been sharply criticized both by Trump's political allies and supporters (who accuse the platforms of restricting freedom of speech and silencing

102 *Ibid.*

103 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 28 November 2020).

104 Human Rights Council, *supra* note 100.

105 Google Transparency Report, 'YouTube Community Guidelines Enforcement', July–September 2020, available online at <https://transparencyreport.google.com/youtube-policy/removals?hl=en> (visited 28 November 2020).

106 Whereas Twitter's suspension of Donald Trump's account is permanent, Facebook referred the suspensions of Trump's account on platforms owned by the giant to its Oversight Board, a recently establishment independent oversight mechanism with the power to issue binding decisions on the company. See Twitter, Inc., 'Permanent suspension of @realDonaldTrump', 6 January 2021, available online at https://blog.twitter.com/en_us/topics/company/2020/suspension.html (visited 28 November 2020); Facebook, 'Referring Former President Trump's Suspension from Facebook to the Oversight Board', 21 January 2021, available online at <https://about.fb.com/news/2021/01/referring-trump-suspension-to-oversight-board/> (visited 28 November 2020).

conservative voices), and by some European countries, including Germany and France.¹⁰⁷

Beyond this, however, a significant question is whether and to what extent Internet platforms can identify and preserve potentially relevant material that can serve accountability purposes. Of particular concern to accountability efforts is the algorithm-determined removal of content that contravenes ‘user standards’, or is illegal.¹⁰⁸ Material that would otherwise have been in the public domain is shielded from the view of investigators. This raises a series of urgent questions for accountability efforts and SMTCs alike: How is material of evidentiary interest and value to international criminal proceedings to be identified and retained? Should the platforms be obliged to alert law enforcement authorities upon receipt of material that crosses a defined threshold, and, if so, which? How, and with whom, should the platforms share this material? Can such material be protected by immunities or subject to state search and seizure orders? If so, subject to which jurisdictions? For how long should platforms be asked to store tranches of material that might be of evidentiary value to a possible, distant future legal process? Which legislative tools (e.g. akin to CLOUD Act Agreements, or Mutual Legal Assistance Treaties) need to be developed to ensure that international mandate holders can efficiently and comprehensively access material held by third parties while respecting data rights?

While some resources, such as the above-mentioned Berkeley Protocol, are making crucial contributions to this rapidly developing field, our findings indicate that the development of further, discrete guidance will be required as the technology and field evolves, and challenges become more apparent. As the gatekeepers to much of the digital realm, SMTCs must be part of the conversation, alongside lawyers, ethicists and lawmakers. Because they will ultimately control how their algorithms and policies work, it is imperative that they are engaged early and promptly by (and, in turn, themselves engage with)

107 J. Guynn, “‘They Want to Take your Speech Away,’ Censorship Cry Unites Trump Supporters and Extremists after Capitol Attack’, *USA Today*, 15 January 2021, available online at <https://eu.usatoday.com/story/tech/2021/01/15/censorship-trump-extremists-facebook-twitter-social-media-capitol-riot/4178737001/> (visited 28 November 2020); R. Hart, “‘Problematic’ and ‘Perplexing’: European Leaders Side With Trump Over Twitter Ban’, *Forbes*, 11 January 2021, available online at <https://www.forbes.com/sites/roberthart/2021/01/11/problematic-and-perplexing-european-leaders-side-with-trump-over-twitter-ban/> (visited 28 November 2020).

108 M. Pizzi, ‘The Syrian Opposition is Disappearing from Facebook’, *The Atlantic*, 4 February 2014, available online at <https://www.theatlantic.com/international/archive/2014/02/the-syrian-opposition-is-disappearing-from-facebook/283562/> (viewed 28 November 2020); 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 20 November 2020); BSR, *Human Rights Impact Assessment: Facebook in Myanmar*, 5 November 2018, available online at <https://www.bsr.org/en/our-insights/blog-view/facebook-in-myanmar-human-rights-impact-assessment> (visited 20 November 2020).

accountability actors across the board. Some efforts are already underway,¹⁰⁹ but more are needed. The continued development of international guidance and best practices for those operating at every stage of the evidentiary flow is crucial to optimizing the probative value of material generated and gathered by those caught up in the world's most atrocious crimes. This is, we believe, today's biggest challenge in the field of accountability and international justice.

5. Conclusions

As our data indicate, a 'third wave' of institutional developments in international justice — prompted by what we have termed the 'accountability-turn' — is reflected in the shared goals and increasing symbiosis between evidence providers on the ground and accountability actors around the world. The spread of OSINT, GEOINT, FININT and documentary technologies beyond the control of state apparatuses has revolutionized how evidence and data move around information systems, forcing accountability actors to contend with new realities. Effectively harnessing this digital revolution will be the key to the success of efforts to pursue public transparency, and to holding individuals and states accountable for the world's worst atrocities.

As this article has discussed, the opportunities are extensive. However — from disinformation and false flag attacks, to verification and authentication of digital evidence, to the potential for biases in and vulnerability to manipulation, to challenges to the assessment of digital information's probative value, to secure information management and archiving — the challenges offered by new technologies and digital evidence are significant.

UN mandates increasingly sit at the heart of the 'life cycle' of digital information and evidence. As such, such institutions stand to play an important role in overcoming the many challenges that this changing landscape and ecosystem presents. Our Oxford project 'Anchoring Accountability for Mass Atrocities' aims to contribute to this discussion. This includes understanding what could be within UN mandates' capabilities (and remit) with regards to the verification and safe storage of material, and how this can be shared with other accountability actors and institutions. However, urgent challenges persist beyond these mandates. These range from the standardization of collection, verification and security processes for documenters on the ground, to third party control of information. While important steps have already been taken, additional international guidance and best practices are needed for the many actors now seeking to leverage new technology and digitally derived evidence in international justice and accountability processes. We believe that finding ways to work with tech companies, in particular, to identify and preserve

109 For example, the Oxford Programme on International Peace and Security is partnering with the Berkeley Human Rights Center and the International Bar Association to develop guidance and best practices aimed towards the removal, storage and disclosure of information originating from social media platforms that might be or become of interest to investigations and prosecutions of atrocities and other types of violent crimes.

information that can assist accountability efforts, is one of the greatest next challenges for this field. The challenge is ripe for scrutiny, and we welcome the opportunity to participate in these discussions, sometimes by asking hard questions — though always with the hope of supporting and improving the delivery of justice for the world's worst crimes.