Yale MACMILLAN CENTER
Genocide Studies Program

Mass Atrocities in the Digital Era Initiative (MADE) *Working Paper No. 1*

March 2021

Social Media Evidence of Alleged Gross Human Rights Abuses: Improving Preservation and Access Through Policy Reform

Olivia Mooney, Kate Pundyk, Nathaniel Raymond and David Simon

Potentially crucial digital evidence of gross human rights violations that occur outside the United States is being lost. The absence of a specific legal mandate and protocol by which this evidence could be routinely preserved and accessed is a problem that the United States Congress will need to help solve. This paper builds on the Yale Genocide Studies Program's Mass Atrocities in the Digital Era (MADE) initiative's three-month consultation process with a diverse range of civil society stakeholders working to improve preservation of digital evidence. It considers how U.S. potential liability has limited sharing social media data with stakeholders in the human rights community and presents three potential legal processes to address this issue. This work promotes justice and accountability for alleged gross human rights abuses.

Table of Contents

Executive Summary

Evidence from social media is increasingly central to human rights investigations. In some cases, this content is being deleted before investigators from civil society and international justice mechanisms can examine it. In other cases, potential evidence relevant to international investigations has been preserved within companies, but investigators cannot access it. This document outlines a proposal for addressing this increasingly critical and complex problem.

Ensuring social media evidence of alleged gross human rights abuses occurring outside the United States is able to be preserved and shared will require amendment to current U.S. law, the creation of a process for companies who share this data to be granted a liability waiver, and the designation of a U.S. government focal point for coordinating requests for such waivers. While action by the United States Congress is required to address this issue, this proposal also provides recommendations for how civil society and platform companies can coordinate more effectively on this issue.

This paper comes to the following conclusions:

1. Sharing court-admissible data with international justice mechanisms can be possible through amendments to U.S. law
2. Preservation and sharing access for potential social media evidence of alleged international gross human rights abuses can be granted to civil society organizations through a liability waiver granted through new legislation. An interagency focal point could be established within the government to approve these liability waiver requests. Adjacent to this process, there could be an Evidence Review Board (ERB) to address ethical and legal considerations and provide oversight.
3. Civil society and companies could create a non-governmental coordinating entity to address broader issues of standards and future advocacy. Creating this body outside of government will create a venue for ongoing dialogue and access/preservation requests in cases where waivers are not needed.

Methodology

The conclusions are informed by feedback from a confidential consultation process with stakeholders from leading organizations in open-source investigations, human rights advocacy, academic research, technology law, and international justice mechanisms, representing approximately 18 organizations in total. The process revealed vastly different opinions and interpretations on the suitability of a US-based legislative fix, but participants nonetheless agreed that something must be done on access to data relating to gross human rights abuses: alleged genocide, mass atrocities, crimes against humanity and war crimes. Nevertheless, this working paper should not be taken to be a "conference proceedings" document, nor does it suggest endorsement from any of the consulted stakeholders or their organizations. While reflecting the feedback of those consulted and having been informed by those consultations, this paper's conclusions are solely the position of the authors, on behalf of MADE. Additionally, this proposal only relates to data on gross human rights abuses that a) do not take place on U.S. soil and b) do not apparently involve U.S. citizens as alleged victims or perpetrators.

Background: An American Problem with International Implications

Major social media platforms are almost exclusively US-registered corporations, including Google (YouTube), Facebook, Twitter, and others. Thus, this paper focuses on issues of US-liability for American companies engaged in preservation and sharing of social media evidence of alleged abuses.

Companies often cite multiple reasons, including various elements of U.S. law, for not sharing content with investigators outside a U.S. or applicable international court order. However, the key statute frequently cited by companies as a reason for not fulfilling access requests is the Stored Communications Act (SCA). This paper does not take a position on the

validity of these SCA claims but rather seeks to build a system wherein human rights evidence can be preserved and accessed through a system of clearly articulated legal obligations and bounds. Addressing the barriers that are interpreted as precluding sharing under U.S. law is a key step to ensuring this data is accessible to both US-based and international human rights actors.

Objectives: Preservation and Access

This working paper addresses two issues at the heart of this problem: preservation of evidence and access to evidence. While the two issues are interconnected, each has distinct factors and legal issues at play. Preservation of evidence is not just about past evidence that has been or will be deleted but about prospective evidence retention when mass atrocities scenarios appear to be unfolding. Alexa Koenig proposes the creation of "evidence lockers" within companies to hold preserved content in a way that is outlined in law, with clearly defined obligations for the protection of privacy, intellectual property and national security concerns.[1] After preservation within such a vault, there are two key access issues: (1) access for international justice mechanisms, and (2) access for civil society.[2]

Currently, U.S. law does not have a waiver program or specific provisions that would allow evidence to be shared with international justice mechanisms, such as the US-funded International, Impartial and Independent Mechanism (IIIM) investigating violations of international law in Syria. Fixing this will require amending U.S. law, such as the Stored Communications Act, to ensure that court-admissible data can be shared while complying with privacy and due process obligations.

Civil society organizations and academics are often the first responders when it comes to documenting and preserving records of alleged gross human rights abuses and they also often serve a role in long-term memorialization. What may begin as an NGO-initiated investigation, might become central to formal international justice proceedings at a later point. This ecosystem of actors is crucial to international justice processes and proposals should be created with these existing processes in mind. The creation of a focal point in the U.S. government to coordinate sharing with civil society presents companies with a clear legal avenue and requirements for complying with requests. However, not all cases involve potential US-liability, so a parallel nongovernmental coordination entity could help ensure access is possible in such circumstances. This nongovernmental coordination body could also be able to serve as a convening point for further digital evidence discussions unresolved/unaddressed in this proposal, such as details of digital evidence lockers, proactive holds on data, and additional metadata requests.

Finally, this proposal is focused on addressing US-liability barriers, but other jurisdictions may choose to mirror this approach. MADE's hope is that this proposal, and the resulting U.S. mechanism, can help shape EU reform, as well as create venues for further cooperation on digital evidence issues.

[1] Alexa Koenig, "Big Tech Can Help Bring War Criminals to Justice," November 11, 2020, https://www.foreignaffairs.com/articles/united-states/2020-11-11/big-tech-can-help-bring-war-criminals-justice

[2] See Glossary on page 10 for proposed definitions

I. Background

Evidence obtained from social media has proven crucial in bringing gross human rights abusers to justice in recent years. A German citizen was convicted of war crimes by a German court in 2016 after Facebook photos showed him posing with severed heads in Syria.[3] A similar case in Sweden resulted in a conviction on war crimes charges based on photos he posted online.[4] Social media data, telecommunications data and metadata, and other user-generated content has become essential to investigating, corroborating, and strengthening prosecutions in international proceedings.[5]

Often, this content is deleted from public view and for good reason. Today, social media companies employ increasingly sophisticated algorithms to detect and delete content that infringes on their terms of service. Such content may be classified as terrorist-related, violent, extremist, hateful, sexually explicit or harassing. However, while this data is often rightly removed from public view, records of the posts are often not retained, even though they could be used as evidence in a court of law or for memorialization purposes.

There are different categories of content: content that was deleted before it was known (often prior to any viewers seeing it), content which was deleted after being available publicly (often after user complaints), content that was deleted by a user themselves and associated metadata. Even with the most sophisticated open-source intelligence (OSINT) techniques, some data and metadata are unreachable by investigators.

When a post that is intended to be made public is blocked algorithmically at the time of posting, investigators are unlikely to know that this potential evidence ever existed.[6] This type of immediate deletion is more and more common, according to platforms. For example, almost 95% of YouTube's deleted content between October-December 2020 was deleted by automatic flagging, and over a third of the deleted content had not yet been viewed at all, and more than another third had fewer than ten views.[7] But even if investigators initially see the evidence, it can be deleted at a later date. Human Rights Watch (HRW) recently discovered that 11% of the content from YouTube, Twitter, and Facebook cited in their reports has since been deleted.[8] Amnesty International has also identified this as an issue for years.[9] There are a variety of archival organizations, including the Syrian Archive, the Yemeni Archive, the Sudanese Archive, the Rohingya Archive, and others who exist to preserve this valuable evidence. Syrian Archive, for example, estimates that 21% of their archive of nearly 1.75M YouTube videos archived up to Jun 2020 can no longer be accessed.[10]

---

[3] "Social-Media Platforms Are Destroying Evidence of War Crimes," *The Economist*, September 21, 2020, https://www.economist.com/international/2020/09/21/social-media-platforms-are-destroying-evidence-of-war-crimes.

[4] Nadim Houry, "A Move to Restore Dignity to Syria's Victims," Human Rights Watch, September 15, 2017, https://www.hrw.org/news/2017/09/15/move-restore-dignity-syrias-victims.

[5] Lindsay Freeman, *Digital Evidence and War Crimes Prosecutions: The Impact of Digital Technologies on International Criminal Investigations and Trials*, 41 Fordham Int'l L.J. 283 (2018). Available at: https://ir.lawnet.fordham.edu/ilj/vol41/iss2/1 p. 8-9

[6] "Social-Media Platforms Are Destroying Evidence of War Crimes"; Avi Asher-Schapiro Barkawi Ban, "'Lost Memories': War Crimes Evidence Threatened by AI Moderation," *Reuters*, June 19, 2020, https://www.reuters.com/article/us-global-socialmedia-rights-trfn-idUSKBN23Q2TO.

[7] "YouTube Community Guidelines Enforcement – Google Transparency Report," accessed March 17, 2021, https://transparencyreport.google.com/youtube-policy/removals?total_removed_videos=period:Y2020Q4;exclude_automated:all&lu=videos_by_country&videos_by_views=detection_sources:ALL&videos_by_country=period:Y2020Q4;region:;p:3.

[8] HRW, "'Video Unavailable' Social Media Platforms Remove Evidence of War Crimes," Human Rights Watch, September 10, 2020, https://www.hrw.org/report/2020/09/10/video-unavailable/social-media-platforms-remove-evidence-war-crimes.

[9] "YouTube Removals Threaten Evidence and the People That Provide It," accessed March 9, 2021, https://www.amnesty.org/en/latest/news/2017/11/youtube-removals-threaten-evidence-and-the-people-that-provide-it/.

[10] "Social-Media Platforms Are Destroying Evidence of War Crimes," *The Economist*, September 21, 2020, https://www.economist.com/international/2020/09/21/social-media-platforms-are-destroying-evidence-of-war-crimes.

Companies have been unwilling to share with civil society and international tribunals because they argue that such sharing does not comply with the SCA. When HRW requested access to the now-deleted content cited in previous HRW reports, Twitter responded: "Pursuant to the U.S. Stored Communications Act (18 U.S.C. 2701 et seq.), Twitter is prohibited from disclosing users' content absent an applicable exception to the general bar on disclosure."[11] When the Gambia requested Facebook content related to alleged serious human rights abuses in Myanmar in relation to their ICJ case Facebook denied the request, stating: "Absent a statutory exception, the SCA strictly prohibits Facebook from disclosing the contents of communications on its platform."[12] These two examples demonstrate how policy is endangering access and identify a need for compelling processes to facilitate liability waivers under U.S. law.

The recommendations in this document are in concert with research and advocacy addressing other facets of the problem. The recently released *Berkeley Protocol on Digital Open Source Investigations* provides meaningful guidelines on how open-source information can be used as evidence in international criminal and human rights investigations.[13] Additionally, dominant proposals for evidence preservation are an archive models known as 'evidence lockers' of information held either within companies or third-party groups.[14] Often, these evidence locker models[15] would function as a restricted-access library of content for a range of stakeholders.[16] Two initiatives relating to online violent extremist content specifically include Tech Against Terrorism's Terrorist Content Analytics Platform (TCAP)[17] and the Global Internet Forum to Combat Terrorism's (GIFCT) Hash Sharing Consortium, an industry-based hash-sharing database used to identify and remove terrorist and violent extremist content.[18] MADE's proposal focuses on addressing U.S. liability that currently prevents social media companies from sharing the content. Addressing such U.S. liability is a prerequisite for expanding existing mechanisms and proposals which seek to address this issue.

II. Proposal: A Model for Access and Preservation

Based on feedback from MADE's consultation, this paper offers proposals regarding how U.S. law could be changed to facilitate the sharing of social media data that could serve as evidence in international justice proceedings and other mechanism of transitional justice and memorialization. It splits the problem into three distinct realms: (1) access to court admissible data for international justice mechanisms, (2) civil society access to data that requires liability waivers to be shared, and (3) coordination of data access requests between civil society and companies in cases where liability is not a concern. Each realm requires a separate solution, as described and explained below.

---

[11] HRW, "'Video Unavailable' Social Media Platforms Remove Evidence of War Crimes."

[12] "Facebook's Opposition to Petitioner's Application Pursuant to 28 U.S.C. § 1782" The Republic of the Gambia v. Facebook, Inc., United States District Court for the District of Columbia, case 1:20-mc-00036-JEB-DAR, August 4, 2020.

[13] Human Rights Center UC Berkeley School of Law and United Nations High Commission on Human Rights, "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" (United Nations and the Human Rights Center at the University of California, Berkeley, School of Law, 2020), https://www.ohchr.org/Documents/Publications/OHCHR_BerkeleyProtocol.pdf. "Resources," bellingcat, accessed March 17, 2021, https://www.bellingcat.com/category/resources/.

[14] Alexa Koenig, "Big Tech Can Help Bring War Criminals to Justice," November 11, 2020, https://www.foreignaffairs.com/articles/united-states/2020-11-11/big-tech-can-help-bring-war-criminals-justice; Joan Donovan and Gabrielle Lim, "The Internet Is a Crime Scene, And After the Capitol Riots We Need Better Information Governance Rules for Treating It Like One.," *Foreign Policy*, January 20, 2021, https://foreignpolicy.com/2021/01/20/internet-crime-scene-capitol-riot-data-information-governance/.

[15] Koenig et al., (*Forthcoming)*

[16] Koenig, "Big Tech Can Help Bring War Criminals to Justice"; HRW, "'Video Unavailable' Social Media Platforms Remove Evidence of War Crimes."

[17] Terrorist Content Analytics Platform, "About Us," https://www.terrorismanalytics.org/about

[18] "About," *GIFCT* (blog), accessed March 17, 2021, https://gifct.org/about/.

Additionally, as with most policy proposals, establishing useful definitions within the proposal is a key challenge. Determining what constitutes an alleged 'gross human rights abuse' under this mechanism, featured prominently during MADE's consultations. Grounding for the relevant core international crimes — genocide, crimes against humanity, and war crimes — is found in international law and U.S. law. These three crimes are defined as "atrocities" by the U.S. Elie Wiesel Genocide and Atrocities Prevention Act of 2018.[19] Genocide, as defined in the UN's Convention on the Prevention and Punishment of the Crime of Genocide (1948), is defined in 18 U.S. Code § 1091.[20] Crimes against humanity, as referenced in the Elie Wiesel Act, is grounded in the international tribunal definition of the crime. Finally, the definition of "war crime" is found in the Jones War Crime Act of 1996 and applies to a "grave breach" of the Geneva Conventions.[21]

Amendments to U.S. law could allow international justice mechanisms to request and receive court-admissible social media data

U.S. law currently does not explicitly allow prosecutors at international justice mechanisms to request access to digital evidence held by platforms. Many platforms have refrained from sharing certain social media content with international mechanisms due to their perceived liability under the Stored Communications Act (SCA). Domestic law enforcement and approved national-level prosecutors (as defined in the CLOUD Act) already have access to this data and use it in court frequently. [22]

Congress needs to build a relevant exception that allows sharing with international justice mechanisms. This will need to include amendment to the Stored Communications Act. Importantly, this would not be the first time the SCA has been amended to address human rights-based concerns. The Act includes specific carveouts, such as one around child sexual assault material archived at the National Center for Missing and Endangered Children. [23]

Amending SCA to allow international justice mechanisms to access data would allow the Congress to define an official pathway and legislative boundaries for crucial discovery and evidence-access to bring those guilty of international human rights abuses to justice. Official pathways would increase the predictability and transparency of the system and allow companies to have more clearly defined legal liability and protection. Congress would likely need to consider other legislative clarifications to ensure that privacy and due process are respected in the sharing process.

One other recent amendment to the SCA that is important to mention is the Clarifying Lawful Overseas Use of Data (CLOUD) Act.[24] It amended the SCA to allow federal law enforcement to compel the sharing of social media data, even if it is stored on servers on foreign soil. It also allows data to be shared with "qualifying foreign governments" who have an executive data-sharing agreement with the United States,[25] providing an alternative route to a mutual legal assistance treaty, or MLAT. Outside of CLOUD agreements, MLATs allow the two or more countries to assist each other with criminal investigations partially based on each country's domestic law. The MLAT process for accessing social media

[19] Benjamin L. Cardin, "Text - S.1158 - 115th Congress (2017-2018): Elie Wiesel Genocide and Atrocities Prevention Act of 2018," webpage, January 14, 2019, 2017/2018, https://www.congress.gov/bill/115th-congress/senate-bill/1158/text.
[20] "18 U.S. Code § 1091 - Genocide," LII / Legal Information Institute, accessed March 17, 2021, https://www.law.cornell.edu/uscode/text/18/1091.
[21] "H.R.3680 - 104th Congress (1995-1996): War Crimes Act of 1996," legislation, August 21, 1996, 1995/1996, https://www.congress.gov/bill/104th-congress/house-bill/3680.
[22] Stephen P Mulligan, "Cross-Border Data Sharing Under the CLOUD Act" (Congressional Research Service, n.d.), https://fas.org/sgp/crs/misc/R45173.pdf.
[23] "18 U.S. Code § 2258A - Reporting Requirements of Providers," accessed March 23, 2021, https://www.law.cornell.edu/uscode/text/18/2258A.
[24] Doug Collins, "H.R.4943 - 115th Congress (2017-2018): CLOUD Act," webpage, February 6, 2018, 2017/2018, https://www.congress.gov/bill/115th-congress/house-bill/4943.
[25] Doug Collins, "Text - H.R.4943 - 115th Congress (2017-2018): CLOUD Act," webpage, February 6, 2018, 2017/2018, https://www.congress.gov/bill/115th-congress/house-bill/4943/text.

data takes time, so CLOUD Agreements were introduced to streamline the process. The CLOUD Act also formalizes a process for companies to challenge a request for data and creates restrictions to address privacy concerns.[26]

The CLOUD Act is important because it created a pathway for foreign governments to bypass the slow MLAT process to access court-admissible data from U.S. companies. As written, however, the CLOUD Act does not include international justice mechanisms. One suggestion that came up in MADE's consultation is to amend the CLOUD Act so international justice mechanisms, such as the ICJ, could request social media data through the same process as a non-US national court.

Regardless of what combination of amendments are required, there are three key metrics that need to be addressed. First, this change can only be deemed successful if international justice mechanisms are able to formally make a request for data and have USG approval that allows companies to comply with the request. Second, the amendments must ensure that there is transparency on how sharing decisions are made, as well as, appeal pathways. Finally, the process must include consideration for user privacy and data security once the data has been shared. Ensuring that these three areas are addressed in the amendment process will be crucial in measuring success but also for preventing abuse.

Preservation and sharing of social media evidence of alleged international gross human rights abuses can be granted to civil society organizations through a liability waiver granted through new legislation. An interagency focal point could be established within the government to approve waiver requests. Adjacent to this process, an Evidence Review Board (ERB) could resolve ethical and legal considerations and provide oversight.

Sharing social media data of alleged gross human rights abuses with civil society may that companies receive a waiver from certain laws including the Stored Communications Act. This USG focal point must consider (1) what information is eligible to be shared, (2) for what duration, (3) what liability waivers are required to allow sharing, (4) what organizations can receive shared data, and (5) for what purpose can a receiving organization use that information.

Answering these questions will require interagency cooperation, but likely will require sign-off by the Department of Justice on any liability waivers. The proposed focal point will process requests from civil society and issue relevant waivers under determined legal parameters. To protect company intellectual property, as well as data security, companies should hold the data and then transfer it directly to approved organizations without an intermediary once a transfer is approved by the focal point. By being interagency, this model remains flexible to the development of new legislation as well as potentially novel forms of potential evidence of alleged gross human rights abuses. Importantly, the focal point can be a place for companies to seek clarity on what their legal obligations are when it comes to sharing.

MADE proposes this focal point also possess an Evidence Review Board composed of both government and external stakeholders that determines parameters and procedures of data sharing under relevant statute, including setting disclosure policies and appeal mechanisms. These parameters should include:

- Countries, scenarios, and content which qualify relating to international, gross human rights abuses;
- Technical Standards of secure storage, retention, and chain of custody;
- Clear guidelines for accessing the content in accordance with privacy rights and data protection standards;
- Terms of sharing after data is received;
- Appeal mechanisms for both companies and requesting organizations; and

[26] Stephen P Mulligan, "Cross-Border Data Sharing Under the CLOUD Act" (Congressional Research Service, n.d.), https://fas.org/sgp/crs/misc/R45173.pdf.

- Measures to ensure the protection of proprietary intellectual property for companies.

The legislation creating the focal point would need to begin to address some of the above concerns. To resolve current and future unaddressed legal and ethical considerations, the Evidence Review Board could be empowered to convene domestic and international efforts to advance and professionalize the field of digital investigations in multiple relevant sectors, not simply the investigations covered in this proposal. The focal point, in conjunction with the Board, could also help advise Congress and the Executive Branch on issues related to domestic and international online investigations and evidence collection.

Congress might decide that these unresolved issues are too numerous for the Evidence Review Board and may require the creation of an ad hoc or standing commission. The Executive Branch has the discretion and responsibility to make a determination on this matter, as well as one on the issue of what actions amount to judicial versus nonjudicial, whether through a presidential commission, a DOJ task force, independent study group, the National Academy of Sciences, or other convening entity.

Civil society and companies could create a non-governmental coordinating entity to address broader issues of standards and future advocacy. Creating this body outside of government would create a venue for ongoing dialogue and access/preservation requests in cases where waivers are not needed.

Not all potential social media evidence will require USG-granted waivers for data to be shared with civil society organizations. Not all data sharing requires government involvement and this proposal does not seek to intervene in processes and use agreements where data is already being shared. Importantly, though, there is currently no forum for social media companies and civil society to coordinate and share data related to alleged gross human rights abuses. Other entities exist, but they focus on more narrow content issues. For example, the Global Internet Forum to Combat Terrorism (GIFCT) is focused on terrorist and violent extremist content only,[27] and while it partners with the UN's Tech Against Terrorism, this model is thematically limited.[28] On the advocacy front, while there are many organizations that bring tech companies and civil society together, such as the Global Network Initiative (GNI),[29] there is currently no group of companies and advocates focused squarely on preserving and sharing online evidence to document and investigate gross human rights abuses. While having a government focal point will be important to address U.S. liability hurdles, there will no doubt be more international issues that cannot be addressed by a group housed in the U.S. government.

Repeatedly in MADE's consultation, participants pointed to the need for a third-party organization for this purpose. This civil society-company coordinating entity could bring together international stakeholders, with a focus on supporting regional, grassroots groups who might not have the means to liaise with companies on their own. Existing social media data sharing involves bilateral sharing between organizations and companies. It is envisioned that this entity can be a space to expand and facilitate those types of agreements in an equitable manner.

[27] "About," *GIFCT* (blog), accessed March 17, 2021, https://gifct.org/about/.; "GIFCT: Possibly the Most Important Acronym You've Never Heard Of," Just Security, September 30, 2020, https://www.justsecurity.org/72603/gifct-possibly-the-most-important-acronym-youve-never-heard-of/.
[28] "Tech Against Terrorism," September 6, 2017, https://www.techagainstterrorism.org/.
[29] "Global Network Initiative," January 16, 2018, https://globalnetworkinitiative.org/.

Conclusion

Many stakeholders within civil society, companies and government have identified social media data access as a problem that needs to be addressed. This working paper outlines recommendations focused primarily on overcoming the hurdles in U.S. law. Addressing U.S. liability is a prerequisite for other proposed mechanisms to allow companies to comply with preservation and access requests in the first place.

One benefit of the case-by-case nature of this proposal is its ability to adapt to circumstances and technological advancement. While the Stored Communications Act is the clearest avenues of liability (both real and perceived), other issues may arise with current and future privacy regulations, laws against material support for terrorism and technology legislation. Furthermore, the case-by-case nature of the mechanism allows it the flexibility to promote relevant data protection parameters and use agreements according to the nature of each request. Addressing U.S. liability allows companies to share data with stakeholders in the human rights community and thereby strengthens human rights investigations and accountability internationally.

Preventing Abuses

Any policies on this issue must build requirements for accountability, remedy and oversight to prevent potential abuse. Sharing should be narrowly limited to alleged gross human rights abuses and the focal point process should also include processes to verify and certify civil society organizations, particularly in terms of cybersecurity and privacy protections. The Evidence Review Board described on page 8 will be an essential mechanism for oversight, transparency, limitation, and accountability to prevent abuse. Essential to this is the creation of a robust appeal process for both companies and organizations. The ERB should also institute processes to limit fishing expeditions. Finally, they should be required to publish scheduled public reports that disclose the general nature and quantity of sharing that has occurred.

Impact

The impact of this proposal can be measured through the quantities of requests on preservation, access, and the initial process of defining standards. MADE expects that with a government focal point to issue waivers to social media companies, companies will feel comfortable preserving and allowing civil society to access the data. With such changes, civil society should be able to access and report on social media data of alleged gross human rights abuses without worrying that the data they have access to will disappear forever. Instead of being unwilling or unable to share data without a subpoena, companies will share potential evidence of alleged gross human rights abuses under parameters constrained by statute and guidance from a review board.

Comments and feedback on this working paper can be sent to olivia.mooney[at]yale.edu

Glossary

Civil Society Organizations: refers to U.S. and non-US nongovernmental organizations and academic groups. This includes large multinational entities but also smaller grassroots organizations. In the eyes of this proposal, academic researchers seeking access to data would also fall into this category.

Gross Human Rights Abuses: reference the alleged core international crimes: genocide, crimes against humanity and war crimes. The scope of this proposal only relates to data on gross human rights abuses that do not take place on U.S. soil and also do not apparently involve U.S. citizens as alleged victims or perpetrators. Further clarification of the legal basis for this definition can be found on pages 5-6.

International Justice Mechanisms: includes justice and accountability mechanisms that cannot presently access court-admissible social media data through the CLOUD Act agreements or MLATs. This includes international tribunals, UN fact-finding missions and international dispute mechanisms such as the International Court of Justice.

Liability Waiver: a U.S. government-granted liability waiver for companies who preserve and/or share data relating to alleged gross human rights abuses within defined parameters. Most commonly, this liability waiver will likely be based on exemptions to relevant requirements the Stored Communications Act.

Preservation: the retention of data within social media companies. Data that is preserved may have been deleted from the public-facing platform but is retained in internal archives (commonly called "evidence lockers").

Prospective retention: The flagging of a potential or ongoing gross human rights abuse situation for proactive data preservation within the social media company. For example, if a specific region of the world is erupting into crisis, a proactive retention strategy would result in social media companies holding on to related data, as opposed to deleting it in everyday data minimization activities.

Sharing: the transfer of data from a social media company to an international justice mechanism or civil society organization.

Social Media Companies: include platforms where there is public sharing of user generated content. As such, content posted publicly to a platform such as Facebook or Twitter could be accessed through this mechanism. However, direct communications through messaging services like WhatsApp would not be accessible.

Stored Communications Act (1986): prevents service providers from sharing customers' communications. It creates tiered legal request requirements to allow for sharing both content and non-content (metadata). The SCA reflects an attempt to create privacy and search and seizure protections for users of (private) network providers and bans some types of voluntary disclosure.[30] While the implications of sharing with governmental entities are more definitively constrained by the SCA, the act does not draw clear lines around how it affects sharing with nongovernmental third-party organizations

[30] Orin S. Kerr, "A User's Guide to the Stored Communications Act, and a Legislator's Guide to Amending It," *The George Washington Law Review* 72, no. 1208 (August 5, 2003): 36, https://doi.org/10.2139/ssrn.421860.