

SOCIAL & TECHNOLOGICAL HUMANITARIAN STUDIES

DEPARTMENT OF GEOGRAPHY
UNIVERSITY OF CALGARY

What happens when open-source data and citizen evidence becomes a resource for human rights investigators?

Angela Ambrose

Social & Technological Humanitarian
Studies Working Paper #1

Published Open Access on SocArXiv

An essential dimension of humanitarian work is human rights investigations to identify violations and crimes. Human rights investigation organizations, in the digital age, are taking advantage of the growing prevalence of online citizen evidence and extractable data from what they often refer to as 'open sources' and social media

The Atlantic Council and Bellingcat are guilty of war propaganda. As @ian56789 wrote to me in a message:

"The members of the Atlantic Council and DFRLab should be indicted as accomplices to War Crimes, for providing actual material support to al-Qaeda terrorists, and for Treason (actively supporting official enemies of the US & UK). They should be spending the rest of their lives in jail and fined every penny they've got."

And those abusing and exploiting Bana al-Abed in their ongoing war propaganda should join them.

From <https://www.rt.com/op-ed/431128-bana-alabed-bellingcat-atlantic-council> Eva Bartlett, June 29, 2018.

platforms. For the purpose of this discussion, we make use of the term ‘open source’ as it is specifically used by the organizations discussed here – we acknowledge that ‘open source’ as a term is often used in problematic ways in place of what is simply extractable, publicly available data – the term open source refers to accessible and editable software source code and in this paper’s context the term often misleadingly refers to datasets that have come at a high cost to the organization that procured them. Human rights investigations labs’ mission statements advocate for the dispelling of disinformation, distinguishing between verifiable information relating to human rights violations and unverifiable, potentially harmful and sensationalist media claims. *Disinformation*, not to be mistaken with *misinformation*, is intentionally false or inaccurate information - often spread deliberately for a particular purpose. Bartlett, a freelance journalist, wrote the following for *Russia Today* on two investigation organizations, Atlantic Council’s Digital Forensics Research Lab and Bellingcat, both of whom were involved in Syrian conflict investigations. These two labs had used information from a controversial Twitter account allegedly run by Bana al-Abed, a Syrian child based in Aleppo who documented her experience. Bartlett supported the idea that members of Atlantic Council and Bellingcat should be tried for war crimes as a result of their support for the Twitter account that had been criticized for being exploitative and widely accused of being war propaganda. Interestingly, however, Twitter had given al-Abed the ‘blue check’ of a verified account, but this did nothing to dispel – and perhaps even ignited further – significant backlash against the use of this account for documenting the ongoing siege in Syria:

While media sources certainly hold their fair share of biases, there is room for analysis on how organizations go about their investigations, who’s involved and how they differ from one another. Bartlett’s claim could very well be an exaggeration of reality, or one supported by concrete evidence. From this situation, myriad questions arise about the ethics and utility of the information generated by human rights investigation organizations:

- * **From where are citizen evidence and open source data acquired?**
- * **Who’s responsible for securing, managing and using it ‘appropriately’?**
- * **Is this type of evidence an efficacious ethical and advocative method for achieving justice in human rights cases?**

International, national and regional NGOs that specialize in human rights investigations have all, within the past decade, demonstrated the utility of citizen-generated evidence and data from open sources. Such projects can range from the investigation of reported violations at the United States-Mexico border (WITNESS Media Lab), to using change-detection techniques as a method to identify major oil spills that can result in complex environmental justice cases (Amnesty Decoders), to analyzing remote-sensed satellite imagery of locations of airstrikes in Syria in an effort to understand how civilians were impacted (Bellingcat).

Here, we take a non-exhaustive sample of human rights investigation labs and contrast them with one another in terms of their techniques, persons involved, and types of projects undertaken. Indeed, the investigation labs analyzed here have not only different *raison d’être*, but disparate guiding principles as well: Some promote themselves as investigators of individual human rights abuses, and others specialize in what appear to be larger-scale investigations to detect civilian impact in conflict zones or human-induced environmental damage that results in environmental impacts *on* humans. These are all human rights violations in one way or another, and we consider this in our analysis.

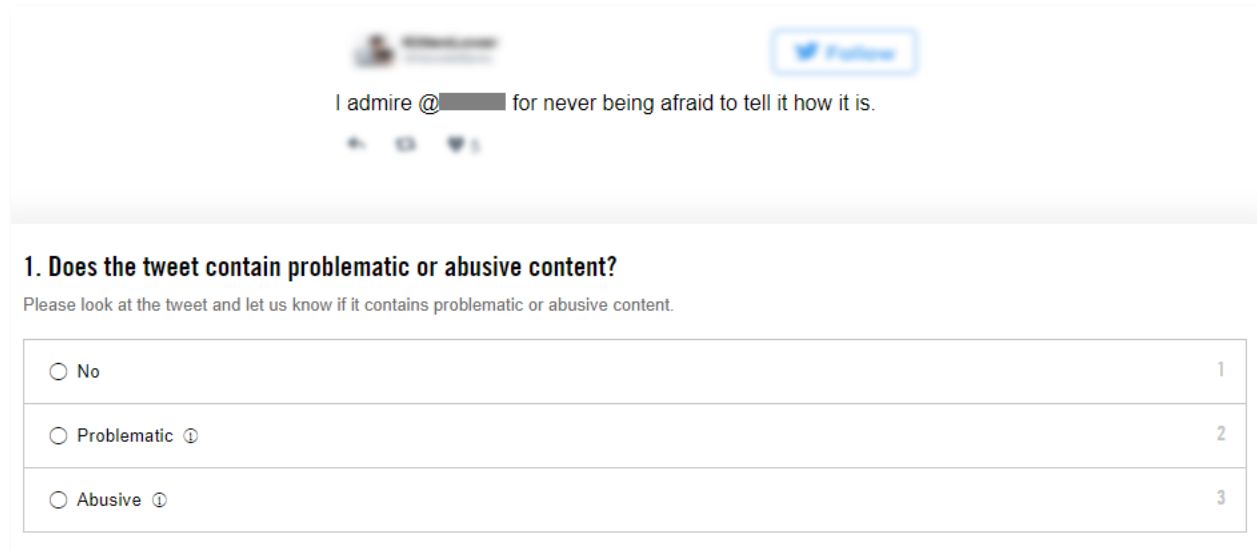
Amnesty International

Amnesty International, a highly recognized international NGO, employs at least two organizational projects with the aim of verifying citizen evidence to identify and use open-source data to support evidence of human rights abuses: *Amnesty Decoders* and the *Digital Verification Corps* (DVC).

Amnesty Decoders

Amnesty Decoders is, unlike most other human rights investigation organizations, comprised of crowdsourced volunteers who support and provide information to Amnesty International’s researchers. It operates using a micro-

tasking platform that is designed for ease of access, so more volunteers can easily contribute to the project. A current project, 'Troll Patrol', asks volunteers to read randomized Twitter posts (tags are removed for anonymity) and categorize them as problematic, abusive or neither (Amnesty International, 2018).



I admire @ [redacted] for never being afraid to tell it how it is.

1. Does the tweet contain problematic or abusive content?

Please look at the tweet and let us know if it contains problematic or abusive content.

<input type="radio"/> No	1
<input type="radio"/> Problematic ⓘ	2
<input type="radio"/> Abusive ⓘ	3

Amnesty Decoders' *Troll Patrol* project, 2018. <https://decoders.amnesty.org>

Other projects often involve image analysis using open-source satellite imagery. The same location is compared between dates in an effort to detect changes. A project executed in 2016, Decode Darfur, was initiated to detect evidence of attacks by the Sudanese government and allied militias. Other Decoders projects have included oil spill detection in Nigeria, impacting nearby residents.

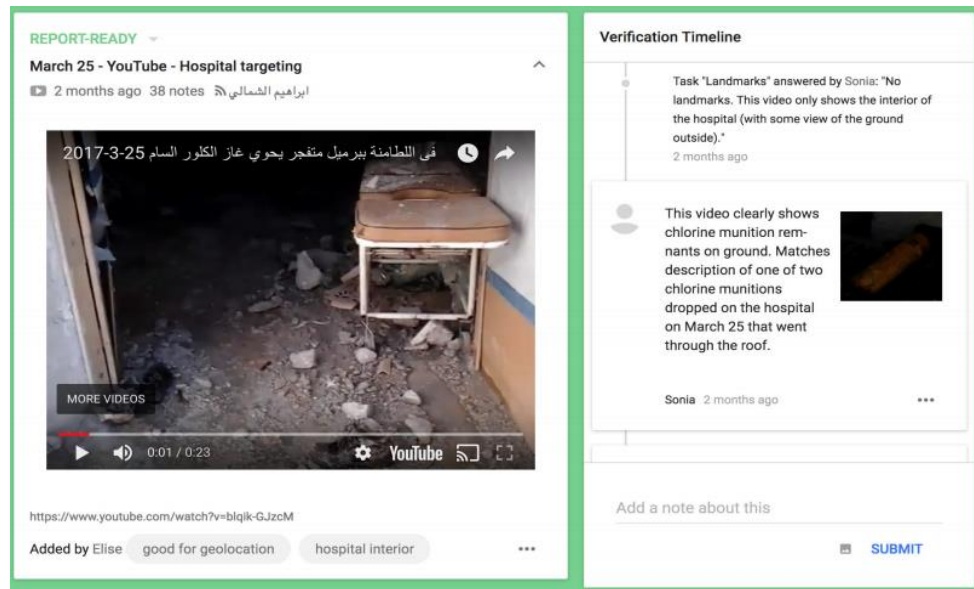
Digital Verification Corps

While Amnesty International's Digital Verification Corps (DVC) do not appear to have a central dedicated website, they seem to be an emerging key player in the field of human rights investigations. Founded by Sam Dubberley, the DVC is in place as a response to the growing volume of data. By employing student volunteers, Amnesty researchers would technically have more resources to work with in their investigations. Comprised of a student volunteers and interns across a network of international universities and their human rights centres – UC Berkeley, University of Essex, University of Pretoria, University of Toronto and, more recently, the University of Cambridge, students with a background in law are trained in verification of citizen evidence and open source investigation methods to identify human rights violations, often using social media platforms (Verhaert, 2017). Data, in past projects, has often been provided by Syrian Archive, which is an organization that compiles and preserves citizen-uploaded media related to the ongoing conflict in Syria before it is taken down (Verhaert, 2017).

UC Berkeley's Human Rights Investigations Lab

UC Berkeley's Human Rights Investigation Lab, headed by Alexa Koenig, works as part of the network of Amnesty International's DVC but also undertakes projects independently. Founded in 2016, the organization is, by majority, student-run. Students come from a range of disciplinary fields of study and linguistic backgrounds. The Human Rights Investigation Lab, like the DVC, partners with Syrian Archive. Other partners include The Center for Justice & Accountability, ProPublica, Archer and Meedan (Berkeley Law, 2018).

In March 2017, UC Berkeley students undertook an investigative project leveraging publicly available social media posts from YouTube, Twitter and Facebook and using mentorship from the Syrian Archive, a human rights collective. The goal was to use this data to verify whether there were chemical strikes on al-Lataminah, Syria. Syrian Archive collected the media data and distributed it to the students at the Human Rights Investigation Lab for verification. The team leveraged geolocation methodologies by analyzing video content to identify landmarks and cross-reference it with satellite imagery. They also used a tool to preserve the data in case the videos were systematically removed from the platform during the project.



Part of the verification process (on a video extracted from YouTube during the al-Lataminah chemical strike project. <https://www.law.berkeley.edu/wp->

DFRL claimed that social media content verification played a pivotal role in determining whether chemical attacks occurred in al-Lataminah – reports cite the use of citizen evidence to more efficaciously identify human rights abuses such as civilian attacks when there’s little-to-know popular media coverage or acknowledgement by formal institutions.

Atlantic Council’s Digital Forensic Research Lab, or #DigitalSherlocks

The Atlantic Council, a Washington, D.C.-based think-tank involved in international affairs, operates and funds the Digital Forensic Research Lab (DFRL). The DFRL employs figures they refer to as ‘experts’, rather than volunteers or students, contrasting with Amnesty International’s associated human rights investigation organizations. They include Bellingcat founder Eliot Higgins, former senior advisor in the Obama Administration, Naz Durakoglu, and former White House and National Security Council adviser, Graham Brookie. Techniques used by the DFRL are unclear apart from the general use of what they call ‘open-source’ (derived from satellite imagery, Google Earth or similar) and social media data (Facebook, YouTube and Twitter primarily) – their published reports demonstrate evidence of their findings, but not necessarily how that evidence was found.

Most published reports on completed projects are centered on ‘debunking’ disinformation by monitoring viral trends in social media and using image analysis techniques. DFRL, sometimes in collaboration with Bellingcat, has had a number of recent investigations involving the ongoing conflict in Syria. Most recently, a recent image analysis project resulted in locating a mass grave in Raqqa, Syria.

Unfortunately, a persistent challenge lies in locating published detailed information on how DFRL undertakes and carries out investigations, what technologies they use and their methodological approach. Their website states that

they use what they refer to as open data, social media data and ‘digital forensic research’ to investigate war crimes, although what entails ‘digital forensic research’ remains unclear – published reports exude a ‘leave-it-to-the-experts’ disposition in which there’s minimal discussion of what comprises the research process and more of the determinations.

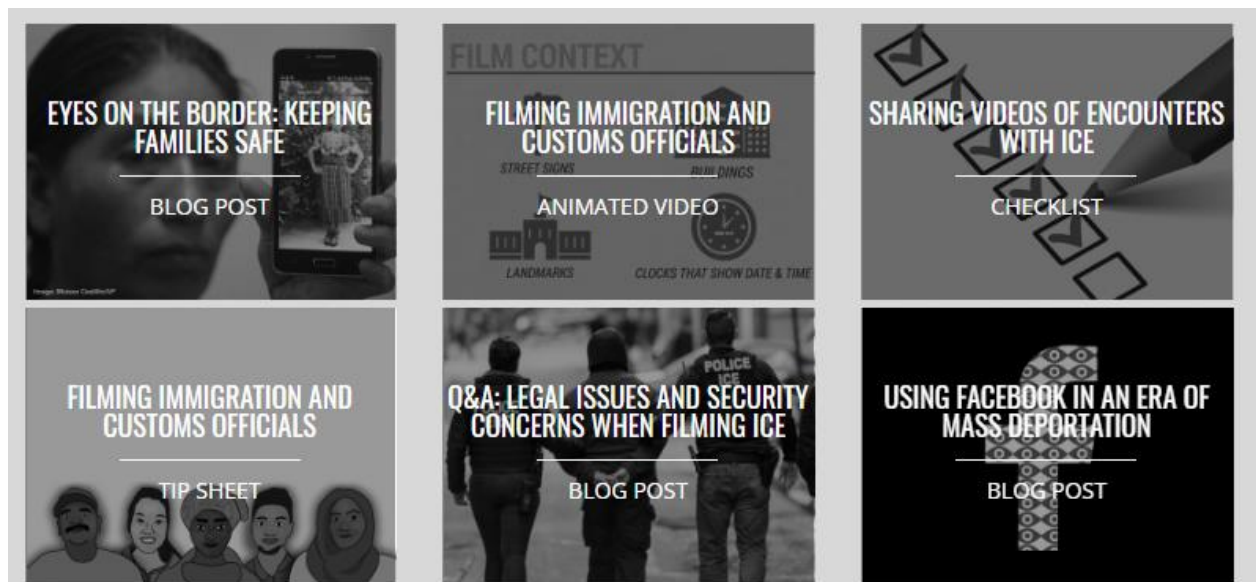
Bellingcat

Bellingcat is a UK-based organization initially funded using a Kickstarter crowdfunding campaign, although current sources of funding are unclear – they do not publish annual reports or lists of financial contributors. The organization has worked closely with the DFRL, sharing several figures – the most prominent and well-known being Eliot Higgins, founder of Bellingcat and nonresident senior fellow of the DFRL’s Future Europe Initiative (Atlantic Council, 2018). Bellingcat also has a team comprised of both full-time investigators and volunteers (where volunteers are recruited from is unclear). The organization’s website doesn’t publish official, detailed reports (or none that were viewable by the public at the time of search) as much as information in the form of blog-style posts that are categorized by geographic region. There are how-to resources geared towards journalists, analysts and researchers working independently of Bellingcat that often discuss new verification techniques using social media. For example, one such resource published in July 2018 entitled *How to Digitally Verify Combatant Affiliation in Middle East Conflicts*, provided links to weapon and camouflage pattern databases for individuals to cross-reference with potential evidence one witnesses in social media posts – techniques like this help in identifying likely locations of conflict.

Bellingcat is well-known for its investigation of the downing of Malaysian Airlines Flight 17 (MH17) and identifying a ‘key person of interest’ using mobile traffic data that revealed contact information and traced phone calls of the individual. More recently is Bellingcat’s (and DFRL’s) work on the conflicts in Syria, including investigations on the bombing in Raqqa, Syria using data of satellite imagery that was date-checked and cross-analyzed with information posted on Twitter by the Russian Ministry of Defence (Bellingcat, 2018).

WITNESS Media Lab

While not a human rights *investigation* organization, WITNESS Media Lab is in place with the goal of training and equipping citizen activists to “expose injustices through video” - supporting human rights investigation organizations through helping to generate citizen evidence that leads to the identification and follow-up on violations (WITNESS Media Lab). Their *Cameras Everywhere* report discusses the “current challenges and opportunities at the intersection of human rights, video and technology”. The challenges are related to privacy and safety, ethics, overabundance of ‘information’, and preservation. Almost all components of their organization, according to the website, are directed toward training the public to responsibly film human rights violations in ways that don’t subsequently create ethics, safety or privacy violations for the activist or those involved in the situation. According to their Frequently Asked Questions, WITNESS was instrumental in encouraging YouTube to create its face blur tool so that activists could upload citizen evidence without revealing the identities of people shown in their footage (WITNESS Media Lab).



WITNESS Media Lab's *Eyes on ICE*' project resources and updates <https://lab.witness.org/projects/eyes-on-ice/>

WITNESS Media Lab's current projects are on the U.S.-Mexico Border crisis and racial profiling in the United States. They also operate in an international capacity, including analysis of citizen video evidence documenting civil unrest during Morocco's media blackout (WITNESS, 2018). WITNESS' *Eyes on ICE* project, focusing on the border crisis, includes a series of training videos, resources and articles on the current situation.

Cross-Organizational Contrasts

All of these organizations leverage what they refer to in their publications as open-source and social media platforms to extract data and use it throughout their investigatory process, all with the aim of minimizing disinformation. Atlantic Council's DFRL and Bellingcat tend to focus projects on identifying perpetrators and persons of interest in war crimes rather than Amnesty International's predominant focus on the implications of human rights violations. Many of these projects do focus on the investigation of war crimes that have direct impact on civilian populations in conflict zones. Contrastingly, Amnesty International's investigation organizations, Decoders and Digital Verification Corps (as well as UC Berkeley's Human Rights Investigation Lab) focus on human rights violations with the mentality of providing aid.

While each organization operates in differential ways, there are informational exchanges between most of them – at an Amnesty International Digital Verification Corps summit held in 2017, Eliot Higgins, founder of Bellingcat and non-resident fellow of Atlantic Council's DFRL delivered the keynote address and discussed the importance of crowdsourcing, critical analysis of media sources, and learning to recognize 'false imagery' that can serve as disinformation to serve particular agendas (Human Rights Center, 2017). Bellingcat and the DFRL often work in collaboration with one another, and WITNESS operates as a 'training' facilitator to produce citizen evidence that can, importantly, be used in human rights investigations regardless of investigation organization.

Conclusion

Human rights investigations labs adopt a range of approaches, with implications for conceptions of social justice: some focus on investigations with the focus of providing and distributing aid (Amnesty International, WITNESS Media Lab), while others, such as DFRL and Bellingcat, seem to focus on identifying persons and parties of interest as perpetrators for the purpose of political intervention and legal process as a way to solve large-scale operations that can result in human rights violations.

One may question whether an organizations' projects are motivated, perhaps at least in part, by their partners and financial supporters. Atlantic Council, self-proclaimed as non-partisan, accepts funding from organizations, governments (one of their largest funding supporters is United Arab Emirates), ministries of defense and foreign affairs (Norway, Sweden), NATO and oil companies (Atlantic Council, 2018). Considering such major financial supporters, it isn't a far stretch to think that the Atlantic Council and DFRL may have vested interests in prioritizing some investigatory projects over others.

From where are citizen evidence and open source data acquired? Across this cross-section of investigation organizations, citizen evidence through social media platforms and open source data are heavily leveraged. Facebook, YouTube and Twitter are easily the most used social media platforms by these organizations, and open source data ranges from satellite imagery for image analysis to databases of mobile traffic acquired by 'digital forensic' means.

Who's responsible for securing, managing and using it (the data) 'appropriately'? It's challenging to contrast organizations with a lack of information on how data are secured, managed and used throughout the investigatory process. There's a big, wide world of human rights investigation organizations out there, and it's ultimately up to each one to decide how they'll extract, manage and leverage the data for a particular project. It seems that, by-and-large, it's the responsibility of the citizen to produce evidence in a method that protects their safety and anonymity (and those involved in the potential violation).

Is this type of evidence an efficacious ethical and advocative method for achieving justice in human rights cases? Each organization operates differently – this question could only be answered by understanding *how* the evidence is used by a given organization, and, equally importantly, how citizen evidence and open source data are acquired, secured, and managed. Organizations such as WITNESS function in order to support the use of citizen evidence by providing resources that improve the 'usability' of it in an ethical and secure way. This could result in an investigation going two ways: first, a citizen attempting to capture evidence through video may not capture *enough* information to be used in an investigation due to necessarily following guidelines set out by WITNESS (or other such organizations). Second, and perhaps more favorably, there could be an increased volume in citizen evidence leading to the successful identification of a human rights violation. Another consideration is whether this will result in the dramatic increase of 'usable' evidence which would, in turn, require greater time and resources to be spent on filtering through information – do these human rights investigation organizations have the means for this? These, by majority, are not organizations who crowdsource their investigators.

The Twitter account of Bana al-Abed is a specially telling example of the controversy that surrounds leveraging social media data as citizen evidence. Some of the questions that commonly arose surrounded the legitimacy of the account and the claims made on it, the potential for it to be used as propaganda and, certainly, the potential for the online exploitation of a young child. It's in this sense that more data doesn't always produce beneficial results or 'more' justice if the content either cannot be verified or interpreted in beneficial ways – this data could be manipulated and interpreted to serve many purposes and agendas. Importantly, it is with a deeply critical eye that citizen evidence and open source data must be analyzed. There are as many praxes of data analysis for human rights violations as there are human rights investigations organizations themselves – their purposes and ideas of justice vary, and so too may their interpretations of the data they use.

References

- Atlantic Council. (n.d.). *Eliot Higgins*. Retrieved July 18, 2018, from <http://www.atlanticcouncil.org/about/experts/list/eliot-higgins>
- Atlantic Council. (n.d.). *Honour roll of contributors*. Retrieved July 14, 2018 from <http://www.atlanticcouncil.org/support/supporters>
- Berkeley Law. (2018). *Human Rights Investigations Lab: where facts matter*. Retrieved July 17, 2018 from <https://www.law.berkeley.edu/research/human-rights-center/programs/technology/human-rights-investigations-lab-internships>
- Human Rights Center (UC Berkeley). (2017). *Digital verification corps student summit: evaluating the first year of university-based open source investigations of human rights*. Retrieved July 17, 2018 from https://www.law.berkeley.edu/wp-content/uploads/2015/04/Summit_report_2017_final4.pdf
- Verhaert, P. (2017, June 19). *Amnesty International's Digital Verification Corps: new networks and methods for human rights research*. Retrieved July 19, 2018, from <https://www.theengineroom.org/digital-verification-corps/>
- WITNESS Media Lab. (n.d.). *Witness press kit*. Retrieved July 19, 2018, from <https://witness.org/about/press-kit/>
- WITNESS Media Lab. (n.d.). *Watching Western Sahara*. Retrieved July 19, 2018, from <https://lab.witness.org/projects/citizen-video-in-western-sahara/>