WHAT CAN GO RIGHT?

Positive use cases for science and technology in human rights investigations

Access Accountability

WHAT CAN GO RIGHT?

Positive use cases for science and technology in human rights investigations

WHAT CAN GO RIGHT?
Positive use cases for science and technology in human rights investigations

Contributing Authors
Megan Price
Jorge Ruiz Reyes
Raïss Tinmaung
Arthur Gill Green

Editors
Scott Stevens
Yurim Choi

Report Design
Eui-hwan Cho
Sook-yi Oh

Recommended Citation
Access Accountability. *What Can Go Right? Positive use cases for science and technology in human rights investigations*. Seoul: Transitional Justice Working Group, 2021.

Publisher
Transitional Justice Working Group (TJWG) | Access Accountability (AA)
Seoul, Republic of Korea
Website: en.tjwg.org | accessaccountability.org
Email: info@tjwg.org | info@accessaccountability.org
Twitter: @TJWGSeoul | @TJWG_AA

ISBN 979-11-975620-0-6

WHAT CAN GO RIGHT?

Positive use cases for science and technology in human rights investigations

Access Accountability

About Transitional Justice Working Group

The Transitional Justice Working Group (TJWG) is a Seoul-based NGO founded by human rights advocates and researchers from five countries in 2014. TJWG aims to pursue advanced methods for addressing grave human rights violations and advocating justice for victims in pre- and post-transition societies. We collaborate and share our practices with other organizations and individuals concerned with the pursuit of accountability for mass atrocities and human rights abuses.

About Access Accountability

Access Accountability is an initiative of the TJWG. Access Accountability arose from TJWG's experience as a start-up NGO seeking resources and expertise to guide the development of our structure, projects and systems. We provide training and resources for human rights documentation groups globally by identifying areas of need and matching them with the expertise required to achieve their goals. Our aim is to assist groups involved in monitoring and documenting human rights abuses in any region, but particularly those who may be looking ahead to a transitional justice process in their local context.

Sponsor

NED
NATIONAL ENDOWMENT FOR DEMOCRACY
SUPPORTING FREEDOM AROUND THE WORLD

Acknowledgment

RightsCon Online 2020

Contents

Introduction

Advances in remote sensing technologies, machine learning, and information management are already being applied to the field of human rights documentation. Dialogue surrounding these tools has shifted from their potential to help to their potential to harm.

Both are possible.

This publication is an attempt to sketch out a rubric for human rights documenters to better identify and determine which is more likely in a given scenario and how to assess which tools will support their goals. To do this, we will examine the development of two projects as case studies: one using machine learning to help identify potential locations of mass grave sites in Mexico, and another looking at how remote sensing work can serve multiple goals in documenting the genocide of the Rohingya people.

This document has been written with the intention to show both the power of these tools, and also reasonable limits for their use by non-expert, civil society practitioners. The publication describes how the civil-society led collaborations were carried out, and lessons learned for other non-experts looking to integrate remote sensing and analytic tools into their human rights documentation and advocacy work. Civil society groups and activists will be able to examine these cases and decide if and how to apply similar investigative methodologies in different contexts.

Lastly, this document aims to identify areas for future work by interested experts and avenues for future collaborations with activists. At the same time it will also increase stakeholder understanding (including donors) of the resources necessary (time, human, financial) to overcome the technical and organizational challenges involved in creating these successful projects.

This discussion evolved from a RightsCon Online (July 27-31, 2020) presentation featuring experts and practitioners from:

- **Universidad Iberoamericana,** Mexico City, Mexico
- **Rohingya Human Rights Network,** Toronto, Canada
- **University of Guelph, Guelph,** Canada
- **Human Rights Data Analysis Group (HRDAG),** San Francisco, USA
- **Human Rights Information and Documentation Systems (HURIDOCS),** Geneva, Switzerland
- **Transitional Justice Working Group (TJWG),** Seoul, South Korea

This publication was coordinated by Access Accountability, an initiative of the Transitional Justice Working Group (TJWG) based in Seoul, South Korea. Access Accountability provides training and resources for human rights documentation groups globally by identifying areas of need and matching them with the expertise required to achieve their goals. The project's aim is to assist groups involved in monitoring and documenting human rights abuses in any region, but particularly those who may be looking ahead to a transitional justice process in their local context.

POSITIVE USE CASES FOR SCIENCE AND TECHNOLOGY

Predicting the Location of Hidden Graves in Mexico

Megan Price
Jorge Ruiz Reyes

Enforced disappearances and involuntary disappearances are one of the major public concerns in Mexican society. According to official figures, there are 83,000 missing persons in the country. More than 95% of these cases have been documented since late 2006. This is due to the strategy implemented by the federal government to counter drug-related activities. Military and federal police units were deployed in central and northern states of the country to fulfill public security duties. The latter increased confrontations and excessive use of force, not only against armed non-state actors, but also against the civilian population.[1] After 15 years, violence and grave violations of human rights remain extended in the majority of the states of Mexico.

Through a collaboration between the Human Rights Program of

1 Galindo, Carlos, Mara Gómez, Raúl Zepeda, and Roberto Castellanos. "Seguridad interior: elementos para el debate [Internal Security: Elements for Discussion]". *Temas estratégicos*, no. 39 (enero 2017): http://bibliodigitalibd.senado.gob.mx/bitstream/handle/123456789/3344/Reporte39_SeguridadInterior_DistDigital.pdf?sequence=1&isAllowed=y.

Universidad Iberoamericana (HRP), the Human Rights Data Analysis Group (HRDAG) and Data Cívica we have documented discoveries of hidden graves in Mexico to approximate the magnitude of the phenomenon and to assist authorities and groups of families with missing persons in the development of search programs in different regions using **machine learning** models.[2]

Machine learning is the use of computer algorithms that improve automatically through experience and interaction with data

We have created databases of discoveries of graves with observations or registries from official and non-official sources. Official information is acquired through freedom of information requests to local and federal attorney's offices. These requests are made using the National Freedom of Information Platform (Plataforma Nacional de Transparencia)[3] in different periods of time. Hidden grave reports are delivered in different formats such as PDFs or CSVs. However, not all the authorities send information about this phenomena and remains classified, due to a lack of transparency.

Press observations are acquired using specialized websites that gather notes by national and local press websites, utilizing specific search terms such as "hidden grave". Press reports are returned in the form of PDFs that have to be analyzed to extract information such as the year, geographic description of the finding, and the number of bodies exhumed.[4] After this, cases are verified and duplicates are removed from the dataset. More than 3,000 press notes have been gathered from 151 press outlets in this period.

By 2019, we were able to document almost 2,000 hidden graves in the country. However, these figures do not report the full extent of the phenomenon since both sources have biases produced by geographic, economic, and political

2 This is a collaboration that started in march of 2017, during a workshop organized by Universidad Iberoamericana about statistics and human rights. HRDAG provided the main lecture and work sessions were carried out for a week with the three organizations. During this week, the objective of the project was defined and the first results of our machine learning model were introduced.

3 The National Freedom of Information Platform (Plataforma Nacional de Transparencia) is a website from the Mexican government where citizens can request public information from federal and local authorities, according to the *General Law on Transparency and Access to Public Information* in Mexico. Authorities are obliged to respond to the information requests within 20 days. Information can be granted or reserved, depending on the sensitivity of the information requested or the authority's interpretation.

4 Since 2016, Ibero has had a group of 20 volunteers that review these press notes.

factors. To overcome this, we have used machine learning; the goal is to develop a tool that allows us to identify municipalities in Mexico that have high probabilities of having unreported graves; thus, it will help to answer the question of where to look for the disappeared?

Providing tools to answer this question is essential for two reasons. The first is that authorities must start search and exhumation programs according to the *General Law Against Enforced Disappearances* in Mexico. This law was created in 2017 by the Mexican government to encourage investigations into missing persons.

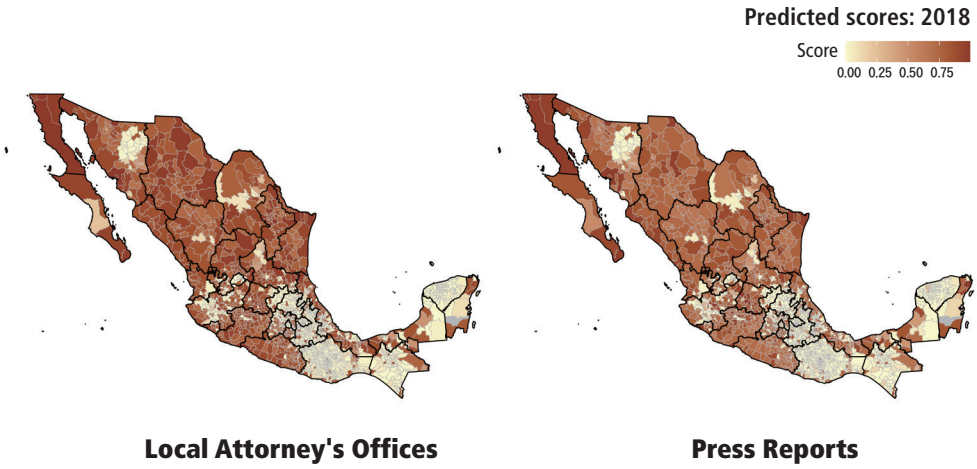
The second reason is that groups of families with missing persons have created their search brigades in different regions of the country to search for their missing loved ones due to the lack of response from the authorities. We aim to assist these groups in pressuring authorities to conduct new search programs based on their knowledge and our results.

The model[5] works by building a list of the 2,458 municipalities in Mexico. We then classify each municipality according to three categories: 1) municipalities where official sources or the press have observed graves; 2) municipalities where we believe it is unlikely hidden graves exist due to the social and geographic characteristics of the municipality and 3) municipalities where we don't know if hidden graves may be found. The third category is the one that interests us the most since these are the sites we believe should be prioritized to conduct new searches.

After this, we include 56 predictor variables that describe characteristics of the municipalities. The data includes geographic, sociodemographic, and violence related information. Then we divide the data into a training and testing set. By doing this, the model "learns" or identifies the characteristics of the municipalities in which hidden graves have been discovered from 2009 to 2018. The model then recognizes the municipalities from our third category and assigns a score – similar to a probability – to all municipalities. We then see

5 We use a machine learning algorithm called Random Forest. It is a regression and classification algorithm that is robust due to its simplicity, ability to deal with missing values and to avoid overfitting. The model is executed in the R programming language.

Figure 1.1 | The maps visualize the predicted scores for the 2, 458 municipalities in Mexico in 2018. Darker municipalities have higher probabilities of having unreported graves.



how the model works with unseen data, using our testing set.

Our results have been highly consistent since we started this project in 2017. We have corroborated that our model correctly identifies municipalities where new hidden graves could be found. For example, in 2011, our model showed that 573 municipalities had an estimated score greater than 0.5. By 2012, a hidden grave was discovered in 30.5% of these municipalities.

Furthermore, our analysis has had a very positive impact on authorities and groups of families. Advocacy strategies to share our results are done by the three organizations. We have presented our results in meetings with groups of families with missing persons and in specialized forums with authorities.

Groups of families and independent forensic teams have used our predictions to pressure authorities or to prepare forensic diagnostics of some states in the country, such as Chihuahua and Nuevo León. In May of 2020, the Argentine Forensic Anthropology Team[6] requested a policy brief about

6 Equipo Argentino de Antropología Forense (EAAF). Accessed June 22, 2021. https://eaaf.org/.

the situation of hidden graves in the state of Veracruz, Mexico. Moreover, in September of 2020, we published along with five groups of families with missing persons a report about the situation of hidden graves in the state of Guanajuato, Mexico.[7]

This report highlighted the emergency of that state and provided detailed information of where graves could be found according to our model. Between October and December of 2020, authorities found graves in two of the predicted municipalities.[8]

We believe this approach provides a valuable advocacy tool to families and organizations in their search for the disappeared and we are optimistic that it could be applied in other contexts. This modeling approach requires two primary types of data: 1) lists of "labeled" locations (i.e., municipalities where hidden graves have been found and municipalities that are unlikely to contain hidden graves) and 2) descriptive information about those locations (such as population that has social security, human development index, homicide rates, etc.). Therefore, groups interested in utilizing these kinds of models will need access to this kind of data and the ability to fit machine learning models.

In our experience, the positive collaboration between HRP, HRDAG, and Data Cívica has created the necessary combination of on the ground knowledge, access to and understanding the context and possible biases of the data, and implementation, evaluation, and interpretation of analytical models. Replicating this approach in other contexts would likely benefit from similar interdisciplinary collaborations.

---

7 Reyes, Jorge Ruiz, Fabrizio Lorusso and Óscar Elton. "Informe sobre la situación de fosas clandestinas en el estado de Guanajuato [Report on the situation of clandestine graves in the state of Guanajuato]." Accessed June 18, 2021. https://fosas-guanajuato.datacivica.org/#intro.

8 Reyes, Jorge Ruiz, Óscar Elton and Fabrizio Lorusso. "Fosas clandestinas en Guanajuato: la violencia que continúa en 2021 [Clandestine graves in Guanajuato: the violence that continues in 2021]." Last modified March 16, 2021. https://poplab.mx/article/FosasclandestinasenGuanajuatolaviolenciaquecontinuaen2021.

Documenting crimes against humanity in Rakhine State, Myanmar

Raïss Tinmaung
Arthur Gill Green

In this section we address efforts to monitor and document human rights violations while outside of a region that cannot be readily accessed or wherein clandestine information gathering must be undertaken. As a case study, we examine the Rohingya experience from 2017-2021.

In August 2017, the Myanmar military conducted a brutal campaign to drive the Rohingya off their homeland in northwestern Rakhine State.[9] The campaign, called "clearance operations", saw an extreme level of savagery that included mass executions, gang rapes, and burning of children alive.[10] An exodus of nearly 800,000 people to neighboring Bangladesh ensued in a matter of weeks, thereby forming what we now know as the world's largest

9 Beech, Hannah, Saw Nang and Marlise Simons. "'Kill All You See': In a First, Myanmar Soldiers Tell of Rohingya Slaughter." *New York Times*, September 8, 2020. https://www.nytimes.com/2020/09/08/world/asia/myanmar-rohingya-genocide.html.

10 Mason, Claire and Mark Kaye. *Horrors I will Never Forget*. London: Save the Children International, 2017. https://reliefweb.int/report/myanmar/horrors-i-will-never-forget-stories-rohingya-children.

refugee camp.[11] Two years later, the government of Myanmar was brought to the International Court of Justice for breach of the the Convention on the Prevention and Punishment of the Crime of Genocide.[12]

In the months following the 2017 massacre, UNITAR Operational Satellite Applications Programme (UNOSAT) and the Australian Strategic Policy Institute released its findings confirming that nearly 400 villages were completely or partially burned to the ground during the 2017 clearance operations.[13] Human Rights Watch's **remote sensing** team also released its findings that confirmed the above, in addition to the fact that the burned village remains were later bulldozed flat to make way for new construction.[14]

Remote sensing refers to collecting data without on-site access. Usually this involves satellite or aircraft-based sensor technologies that can detect objects or features

Each village that is burned to the ground destroys decades, if not centuries of the history of its inhabitants. Each time burned village remains are bulldozed, all kinds of legal evidence and documentation that could be presented at a potential tribunal are erased.

To exacerbate matters, the region of northwestern Rakhine has been under strict prohibition of access for several decades.[15] International journalists are usually denied access or are only provided access under close surveillance by Myanmar authorities.[16] Local journalists are charged and imprisoned for anything done in northwestern Rakhine that supposedly violates the

---

11 "Bangladesh to build one of world's largest refugee camps for 800,000 Rohingya." *The Guardian*, October 6, 2017. https://www.theguardian.com/world/2017/oct/06/bangladesh-build-worlds-largest-refugee-camps-800000-rohingya.

12 United Nations. "Top UN court orders Myanmar to protect Rohingya from genocide." Last modified January 23, 2020. https://news.un.org/en/story/2020/01/1055841; International Court of Justice. "Application of the Convention on the Prevention and Punishment of the Crime of Genocide (The Gambia v. Myanmar)." Accessed June 25, 2021. https://www.icj-cij.org/en/case/178.

13 "Satellite images show Myanmar's 'minimal preparations' for Rohingya return: think-tank." *Reuters*, July 24, 2019. https://www.reuters.com/article/us-myanmar-rohingya-idUSKCN1UJ0D8.

14 Human Rights Watch. "Burma: Scores of Rohingya Villages Bulldozed." Last modified February 23, 2018. https://www.hrw.org/news/2018/02/23/burma-scores-rohingya-villages-bulldozed.

15 "Myanmar Rakhine: Inside the Closed Rakhine Region." *BBC*, December 7, 2016. https://www.bbc.com/news/av/world-asia-38232929.

16 International Federation of Journalists. "Myanmar: Journalist Assaulted in Rakhine State." Last modified May 18, 2020. https://www.ifj.org/media-centre/news/detail/category/press-releases/article/myanmar-journalist-assaulted-in-rakhine-state.html.

government's Official Secrets Act.[17] Top UN officials including the Special Rapporteur on Myanmar are barred from accessing Rakhine State.[18]

With ground access rendered impossible, one of the few options that remain to document the grave human rights violations that took place in northwestern Rakhine is the use of Geographic Information System (GIS) and remote sensing technology. Historic aerial and satellite imagery can be utilized to document areas of criminal investigation interest in and around villages that have been burned to the ground. Such areas would include but not be limited to sites of mass executions, mass burials, killings, and arson.

GIS and remote sensing technology can also be used to document historic landmarks in northwestern Rakhine, particularly those that are of interest to the survivor community. Aerial and satellite imagery, corroborated with information from survivor testimonies, can be leveraged to memorialize sites of social, cultural, and religious significance such as schools, hospitals, mosques, and graveyards that once existed on Rohingya lands to preserve history for the surviving and future generation of Rohingya.

Below, we document some techniques used and current lessons learned from a collaborative project led by Rohingya refugees with Canadian academics, forensic geospatial experts, and Access Accountability. The goals of this collaboration are to:

1. Document human rights violations.
2. Document sites of cultural, religious, and social importance.
3. Create a website to make this information available to the public.
4. Provide training in open-source investigation geospatial tools to human rights workers and refugees.

Methods focus on using reliable, open-source tools to collect, organize,

17 Lewis, Simon and Shoon Naing. "Two Reuters Reporters Freed in Myanmar After More Than 500 Days in Jail." *Reuters*, March 7, 2019. https://www.reuters.com/article/us-myanmar-journalists/two-reuters-reporters-freed-in-myanmar-after-more-than-500-days-in-jail-idUSKCN1SD056.

18 "UN's Yanghee Lee Denied Access to Rohingya Villages." *Aljazeera*, January 15, 2017. https://www.aljazeera.com/news/2017/1/15/uns-yanghee-lee-denied-access-to-rohingya-villages.

analyze, visualize, and share evidence for legal proceedings and for public campaigns.

The first step we undertook was defining project purpose to strategically plan the life cycle of the project and the ways in which data is able to be ethically collected, organized, analyzed, and visualized.[19] This planning time considered logistical and technical considerations for sensitive data and evidence handling. This planning time is meant to avoid future time lost by volunteers and staff attempting to post hoc standardize interview data and fix technical problems rather than pursuing further documentation and analysis efforts. Identification of secure protocols for communication and transfer of the data as well as creating firewalls between confidential data and public facing data were established. This project intends to collect data that can be used for public advocacy and/or as legal evidence in one or more courts of law, so proper documentation steps were considered in relation to jurisdictional standards. The project leaders assigned a timeline to the four goals of the project and related data procedures.

Data acquisition and organization strategies include remote associates gathering information on the ground and project participants working from outside the region with free and paid imagery sources to document burned villages, atrocities, witness statements, and sites of public interest for Rohingya history. A training in KoboToolbox was undertaken to facilitate on the ground data collection. KoboToolbox is an open-source software that allows groups to set up secure servers or use free, secure public instances for hosting surveys that people on the ground can access via mobile phone applications to upload multimedia, multiple survey question types, automatic geographic location information from a phone GPS, and additional geographic information (such as points, polygons, and lines). The application is able to remotely upload data and immediately erase local data stored on a phone so that there is less risk to people on the ground. Currently, a survey based on the project goals is being developed so that data can be entered either remotely on phones or by project

19 Dyer, Sophie and Gabriela Ivens. "What would a feminist open source investigation look like?." *Digi War* 1, (2020): 5–17.

participants on a website. Protocols and permissions are being developed that will, for instance, allow segmented participants to have administrative rights to enter data but not necessarily access or modify other raw data on the server. Within the region, future interview methodology training and informed consent forms are being prepared.

For gathering data outside of the region, we focused on using free (when possible) aerial and satellite images. Through two workshops we studied ways to access free imagery; use online and offline imagery analysis tools to detect changes in infrastructure, soils[20], and environmental variables; digitize and georeferenced hand drawn maps; and handle, load, and visualize data in simple maps using open-source software. We found that given the technical sophistication required to use desktop software and to program custom algorithms for detecting changes in land cover, **software as a service (SaaS)** platforms were often better able to immediately help novice users remotely document changes. For example, in examining a new refugee camp created on a silt island (Bhasan Char in Bangladesh) activists were able to document likely areas of unstable soil that may flood during major storms using SaaS platforms (see Figure 2.1). While SaaS is easier to start with, desktop software and downloaded data help with replication, security, and to build a more reliable evidence base for advocacy and legal proceedings. These workshops on data processing for surveys and aerial and satellite imagery were filmed and shared publicly under a CC-BY-SA license on Access Accountability's website and YouTube.[21]

Software as a service (SaaS) is a subscription based and sometimes free method of software delivery from a centrally hosted server. This allows data to be accessed from any device on the internet and lowers required data processing and storage needs for end users

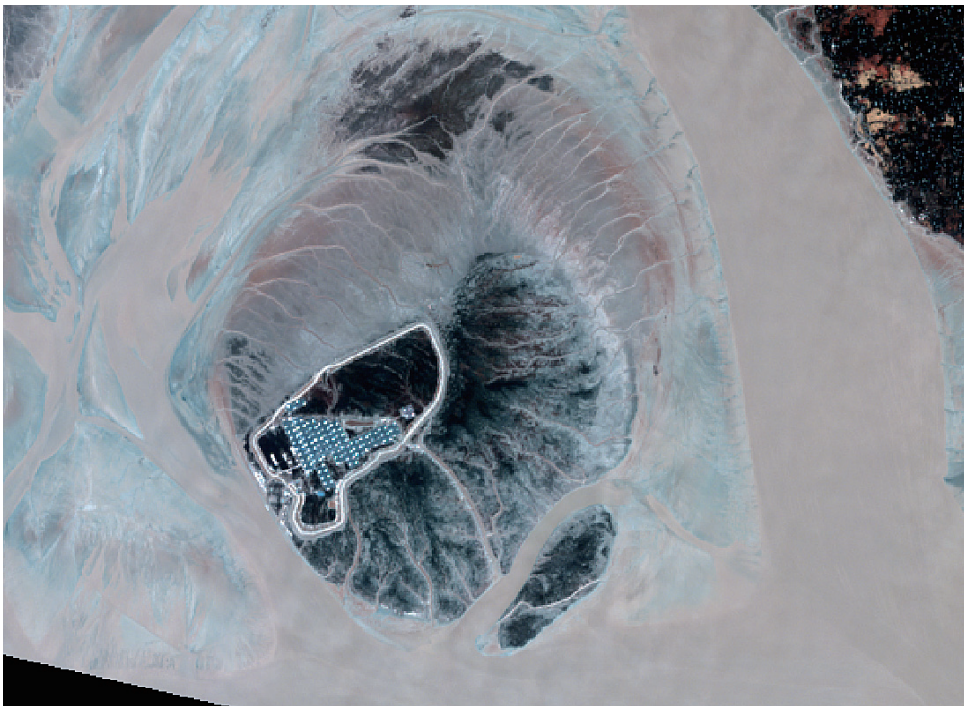
20 Soil spectral signatures refer to the unique ways in which different soils and surfaces reflect light. These wavelengths within multispectral, hyperspectral, and other forms of imagery can be used to understand soil disturbances and soil types.

21 Access Accountability. "Video Tutorials on Satellite Imagery for Human Rights Monitoring (Part 1 of 2)." Resources. Last modified January 15, 2020. https://accessaccountability.org/index.php/2020/01/15/video-tutorials-on-satellite-imagery-for-human-rights-monitoring-part-1-of-2/?et_fb=1&PageSpeed=off; Access Accountability. "Video Tutorials on Satellite Imagery for Human Rights Monitoring (Part 2 of 2)." Resources. Last modified January 28, 2020. https://accessaccountability.org/index.php/2020/01/28/video-tutorials-on-satellite-imagery-for-human-rights-monitoring-part-2-of-2/.

We are currently targeting data acquisition. Yet, future data analysis will involve student volunteers from the University of Guelph, geospatial experts, and other trained volunteers. These participants will utilize partner organizations' data sets and satellite imagery to map known burned villages. As well, using historical imagery the participants will identify large soil disturbances that correlate with witness testimonies regarding mass graves. Collaboratively with people familiar with Rohingya history in local areas, participants will identify Rohingya sites of public interest that need to be documented and, in some cases, may have been destroyed. Lastly, participants will examine ways in which the data suggests repetitive criminal behaviors, criminal intent, and evidence of coverup or clandestine intent. In addition to correlating testimonies with other data sets, this last objective may use spatial statistical models to infer areas of investigation. Currently we are building a method of assigning a level of reliability to both data points (e.g., observation of a burned village) and data sources (e.g., triangulation of witness testimony

Figure 2.1 | Using SaaS EO Browser with prebuilt and custom algorithms to detect flooding vectors and soil stability.

Figure 2.2 | Using QGIS with remote sensing plugins to combine multiple forms of imagery for time series monitoring of soil change.



reliability with imagery and media reports). For example, testimony is assigned a score based on if the event was experienced directly, seen first hand, heard from a reliable source, or heard as a rumor.

Data organization, data visualization, and data sharing objectives required identifying and evaluating **content management systems (CMS)** and types of databases. We determined that, ideally, the CMS would come with a spatial enabled relational database or would easily be upgraded to PostGIS/PostgreSQL to allow spatial visualization and queries. We are in the process of evaluating two CMS and several spatial database options. These database options including the creation of a custom web map server (using a Geoserver based stack) or use of SaaS providers such as ESRI's ArcGIS Online, Mapbox, CartoDB, and QGIS Cloud to lower the

A content management system (CMS) is an application that manages web content, user permissions, and database interactions

technical burden of maintaining a database and map server that allow data visualization and query.

Our criteria for the database and CMS included:

1. Allow documents and other forms of testimony to be securely uploaded by users that can be assigned a set of administrative permissions.
2. Have the capacity to establish, visualize, and query spatial and relational ties between victims, perpetrators, and documented evidence so that when possible these can be publicly navigated.
3. Allow easy customization of both the frontend and backend of the website.
4. Allow implementation of a news feed that can provide updates on the project.
5. Be low cost, free, or open source.
6. Have a low level of technical expertise required to manage CMS and data infrastructure.
7. Able to be independently hosted on a secure server.

The two open source and free CMS options considered were Uwazi[22] and WordPress. To experiment with CMS functionality, the domain https://rohingyamaps.org was obtained and now is owned by a Rohingya refugee human rights organization. While we found that Uwazi has many benefits for human rights documentation and spatial queries, the project ultimately decided to work with WordPress due to more customizability overall and team members' level of familiarity with the CMS.

One recurring theme throughout the early stages of this project is organizational capacity. Small human rights organizations benefit from collaboration with other organizations, academics, and volunteer professionals because the wide array of technical skills and professional

22 Uwazi. Accessed June 22, 2021. https://www.uwazi.io/.

knowledge required to conduct **open-source intelligence (OSINT)** investigations, gather legal evidence, and undertake website and database management. This array of skills is unlikely to be within one organization and can be a project-stopping hurdle in terms of finances and time commitments. While an organization may have essential trustworthy connections and relevant sociopolitical, economic, and cultural expertise, without collaboration a project may not move forward or, possibly worse, make mistakes along the way that expose marginalized people to risk of harm, miss opportunities, and lose trust. Another theme that has become evident is that sharing materials created and the process of navigating the project are an important part of building community, avoiding repetition of mistakes, and finding novel and powerful ways to leverage new technologies for human rights monitoring.

Open-source intelligence (OSINT) investigations use publicly available (open) data sources to inform understanding about events, people, and processes. OSINT investigation techniques require knowledge of quantitative and qualitative methods, and often require high levels of legal knowledge and technical knowledge about imagery and other data sources

As a group, we have limited time and a few financial resources. So, we are actively recruiting volunteers that are willing to be trained or have already been trained in the use of geospatial technologies, are able to translate Burmese or Rohingya language, and who are familiar with or able to document historical sites. Refugee leaders working on this project must travel to conduct interviews and manage teams of interviewers in Bangladesh to collect original and verifiable accounts of crimes against humanity. These costs of raw data collection require funding. While we have attempted to rely on open source and free software, this approach is time consuming and ultimately requires high levels of technical knowledge. So, we are evaluating ongoing costs for maintaining the public website, the survey server, and a database/map server provider. Further legal analysis work is needed to ensure the documentation meets different jurisdictional standards. We believe that this project would benefit from further guidance of an advisory board. As well, our financial limitations for ongoing costs to gather data would greatly benefit from an

organization or individual financial backer interested in technology-forward approaches to human rights monitoring and evidence gathering for one of the worst human rights crises in world history. This work can make the difference of bringing those responsible for crimes against humanity to legal consequences.[23]

23 Human Rights Watch. "Myanmar Events of 2019." Accessed June 21, 2021. https://www.hrw.org/world-report/2020/country-chapters/myanmar-burma#; Bard Wilkinson. "UN Official Convinced of Myanmar Rohingya 'Genocide'." CNN, March 12, 2018. https://www.cnn.com/2018/03/12/asia/myanmar-rohingya-un-violence-genocide-intl/index.html.

GUIDANCE FOR CIVIL SOCIETY IMPLEMENTATION

Below are workflow charts for each of the projects discussed. While each project will have its own unique organizational and data collection challenges, the hope is that civil society documenters can view these charts to get a better understanding of the steps involved to successfully apply these tools. At the same time, experts, technologists, and other stakeholders including donors, can see where resources are needed to ensure the success of other projects using these tools.

In reality, the strict sequencing of the steps and the timing of the resources needed has been simplified. Each project's unique situation will require flexibility when approaching each of the challenges.

Workflow and Resources Map 1:

Predicting the Location of Hidden Graves in Mexico

| Collect information about hidden graves in Mexico |

Core funding for the project was possible through the support of the Human Rights Program at Universidad Iberoamericana.
HRDAG funded their portion of the work through core (or unrestricted or general operating) funds via grants from the Oak Foundation, Open Society Foundations, and MacArthur Foundation.

1. Prepare definitions and concepts to document the phenomena. The objective is to establish what we will understand as a hidden grave, considering the advances in literature and technical characteristics from the forensic and criminal sciences and international human rights standards.

2. Identify the sources of information. Our project includes two primary sources of information: 1) hidden graves reported or observed by national and local press and; 2) official records from local and federal attorney's offices. The press reports are obtained through search engines using different keywords such as "Mexico," "Hidden Graves," or "Disappeared." The official reports are obtained via freedom of information requests. Sometimes this information has to be litigated.

3. Identify the formats and verification of the data. Press reports are obtained in PDF format or spreadsheets with the complete text and link to the original note. The press notes are verified to establish the veracity of the report, date, state, and municipality of the site, description of the geographic characteristics, number of graves, and bodies recovered from the site.

Since 2016, 20 volunteers have helped classify and systematize the press notes. All of them have been undergraduate students from Universidad Iberoamericana. The majority of them worked during the first phase of the project (2016-17). After this period, four of them worked part-time at the human rights program to continue updating our records.

Official reports are obtained in PDF or spreadsheets. Often, this information is not structured, so the analyst has to read the documents and identify the section about the description of the sites.

The main challenge is to train the volunteers to read the press notes and classify the information.

4. Store the data. The data is stored in spreadsheets, separating each source of information. We also keep the original files in their different formats in order to access them if necessary.

| Conduct the statistical analysis | [24]

1. Share the information in a single repository with the three main partners. Hidden grave findings are shared in a list with municipalities of Mexico, indicating whether hidden graves were found in each municipality in a given year. Also, other lists are shared with information obtained by Data Cívica about all these municipalities (independent variables). This data is stored in a Github repository hosted by HRDAG.

This work has been done by 7 members of Data Cívica, since 2017.

2. Prepare the information for the analysis. The data is merged into a single file, where each row is a municipality in Mexico, with information about grave findings and their social or geographic characteristics. Missing data is also imputed in order to conduct the statistical analysis.

3.Implement a machine learning algorithm known as Random Forest to predict municipalities with hidden graves. We use machine learning to identify municipalities with similar characteristics to the ones that the press or official sources have observed. We use a classification algorithm called Random Forest. Predictions are assessed to minimize the identification of false positives.

This is done by HRDAG.

24 The analysis is conducted using an open source programming language called R.

| Sharing our results |

1. Prepare reports of our results and implications of our results to the general public. Blog posts are shared through the websites or the three partner organizations and through national press outlets to explain our results and explain how this tool could support the search strategies of authorities and families with missing relatives.

2. Prepare specific policy reports for authorities and groups of families with missing persons. Reports detailing our results have been prepared for local attorney's offices and groups of families with missing persons in specific regions.

Workflow and Resources Map 2:

Documenting Crimes against Humanity in Rakhine State, Myanmar

| Define project purpose/project planning |

1. Define project goals and stakeholders.

Project definition and stakeholder identification was crucial to launch the project in the first place.

2. Strategically plan the life cycle of the project and identify ways in which data is able to be ethically collected, organized, analyzed, visualized, and shared (e.g., what must be public/private; what legal venue and which authorities).[25]

3. Consult with experts on the above steps (e.g., data scientists, each of these above decisions determines key directions the project should take in the following steps).

For a small volunteer-run organization with limited resources, engaging with subject matter experts may be the only way to realize the project.

25 Dyer, Sophie and Gabriela Ivens. "What would a feminist open source investigation look like?." *Digi War* 1, (2020): 5–17.

| Data acquisition and organization strategies |

1. Identify data collection objectives, variables, and analysis strategies.

2. Create codebook and training manual for coders and surveyors.

3. Identify data collection platform and the required data formats for analysis, interoperability, visualization, and sharing.

4. Multiple coding exercise to establish inter-rater reliability for qualitative coding.[26] Revise codebook and survey tool.

5. Create survey tool on secure platform.

6. Define administrative roles on the secure platform (determines what data can be edited, by whom, and at what stage of the project).

7. Pilot survey tool. Revise codebook and survey tool.

8. Engage in any necessary training of coders and surveyors.

9. Administer survey tool and other data acquisition processes.

Administering the survey tool requires human resources on the grounds. Personnel hired need to have the capability to administer the survey tool and follow the proper protocols for data acquisition, recording and transmitting. Plus, hiring of personnel requires cost considerations that the organization will need to budget ahead of time.

26 Power, Robert and Brian Williams. "Checklists For Improving Rigour In Qualitative Research." *BMJ: British Medical Journal* 323, no. 7311 (2001): 514.

Data analysis

1. Recruit and train volunteers when possible and applicable.

2. Clean survey data.

3. Code secondary materials and survey data.

4. Acquire and analyze remote sensing imagery.

5. Validate and triangulate different data sources to cluster events in time and space.

6. Assign a level of reliability to different data points (observations) and data sources.

7. Using qualitative approaches and spatial statistics models examine ways in which the data suggests repetitive criminal behaviors, criminal intent, and evidence of coverup or clandestine intent.

| Data visualization and sharing |

1. Decide on visualization strategy: audience, narrative/story arc, objectives, platform.

2. For web-based design – determine hierarchy of pages, wireframe design the pages, and look for ways to make actionable and shareable points for advocacy.

3. Use multimedia storytelling strategies to best position types of data collected in public engagement.

4. Prepare reports for public dissemination.

> The output data can be public or classified. If it is classified then we will simply need to ensure that the data is in a format that can be useful to the target agency/group. If it is public then we need to be cognizant of our target audience. Is it the public at large, is it human rights groups, is it college/school children, is it Rohingya school children living and going to schools in the refugee camps, etc.

LIST OF CONTRIBUTING AUTHORS

If you would like to connect with any of the report authors, or are interested in other human rights documentation training, please send inquiries to info@accessaccountability.org

Megan Price, Ph.D.

Executive Director | Human Rights Data Analysis Group (HRDAG)

As the Executive Director of the Human Rights Data Analysis Group, Megan Price designs strategies and methods for statistical analysis of human rights data for projects in a variety of locations including Guatemala, Colombia, and Syria. Her work in Guatemala includes serving as the lead statistician on a project in which she analyzed documents from the National Police Archive; she has also contributed analyses submitted as evidence in two court cases in Guatemala. Her work in Syria includes serving as the lead statistician and author on three reports, commissioned by the Office of the United Nations High Commissioner of Human Rights (OHCHR), on documented deaths in that country. Megan is a member of the Technical Advisory Board for the Office of the Prosecutor at the International Criminal Court and a Research Fellow at the Carnegie Mellon University Center for Human Rights Science. She is the Human Rights Editor for the Statistical Journal of the International Association for Official Statistics (IAOS) and on the editorial board of Significance Magazine. She earned her doctorate in biostatistics and a Certificate in Human Rights from the Rollins School of Public Health at Emory University. She also holds a master of science degree and bachelor of science degree in Statistics from Case Western Reserve University. From 2013 through 2015, Megan was the Director of Research at HRDAG; on December 1, 2015, she became Executive Director.

Contact: meganp@hrdag.org

Twitter: @StatMegan | @hrdag

Link: https://hrdag.org/mexico/

Jorge Ruiz Reyes

Member | Human Rights Program of Universidad Iberoamericana

Jorge Ruiz Reyes has an undergraduate in Political Science and Public Administration from Universidad Iberoamericana in Mexico City. He has also taken specialized courses in international criminal law, statistics, and techniques to search for missing persons. He was a member of the Oaxacan Truth Commission in Mexico, and in 2016 joined the Human Rights Program of Universidad Iberoamericana, where he has led the research regarding enforced disappearances and clandestine graves findings. He is now the coordinator of the line of research on grave violations of human rights.

Contact: jorge.ruiz@ibero.mx

Twitter: @jorgerure

Links: https://ibero.mx/files/informe_fosas_clandestinas_2017.pdf

https://ibero.mx/files/2019/violencia-y-terror.pdf

https://ibero.mx/files/2019/fundenl-informe.pdf

Raïss Tinmaung

Executive Director | Rohingya Human Rights Network (RHRN)

Raïss is a Rohingya from Toronto, and originally from Akyab (Sittwe). He is the founder and director of Rohingya Human Rights Network (RHRN). He has led campaigns, petitions, peaceful rallies, and new chapter formations of RHRN across Canada. His writings have been published in several newspapers including the Toronto Star, Le Soleil, Montreal Gazette, Vancouver Sun, The Hill Times, Free Malaysia Today, etc. Following the 2017 massacre of the Rohingya, Raïss spent 1 month at the refugee camps; after which he presented at the Canadian Senate, as well as at the House of Commons. Raïss is an Aerospace Systems Engineer by profession, but his passion lies in development work on the grounds - he has lived and volunteered in long term development projects in South America, the Carribean, Southern Africa, and the Middle East. Apart from leading the Rohingya Human Rights Network, Raïss also leads a network of schools and vocational training centers at the Rohingya refugee camps and villages.

Contact: raiss.tinmaung@yahoo.com | info@rohingya.net

Twitter: @rohingyanetwork

Link: https://rohingyanetwork.com/

Arthur Gill Green, Ph.D.

Data Scientist | Department of Geography, Environment and Geomatics, College of Social and Applied Human Sciences, University of Guelph

Dr. Green is a geospatial scientist with expertise in environmental governance and human rights issues. He is also an open education advocate and a consultant with over 20 years experience leading projects on forensic geospatial science, Sustainability & GIScience curriculum development, and agricultural land management. Green's research includes GIScience applications for human rights projects - primarily, spatial statistics, remote sensing analyses, and the design of forensic data acquisition and management systems. Green is the Professor of Environmental Management and Geomatics at Okanagan College. Currently, he is a visiting Data Scientist examining cyber threats to national security and agricultural data systems at the University of Guelph in Ontario Canada.

Twitter: @greengeographer

Link: http://greengeographer.com

Spatial Databases and CMS Resource List

- **Joomla!**
  https://www.joomla.org/

- **PostgreSQL**
  https://www.postgresql.org/

- **PostGIS**
  https://postgis.net/

- **The Engine Room. "Tech Tools for Human Rights Documenters."**
  https://www.theengineroom.org/tech-tools-for-human-rights-documenters/

- **Uwazi**
  https://www.uwazi.io/
  https://huridocs.org/technology/uwazi/

- **Wordpress**
  https://en-ca.wordpress.org/

- **KoBoToolbox**
  https://www.kobotoolbox.org/

- **Bayanat**
  https://www.bayanat.org/

Universidad Iberoamericana

Human Rights Data Analysis Group

Rohingya Human Rights Network

Human Rights Information and Documentation Systems

University of Guelph

Transitional Justice Working Group

ISBN 979-11-975620-0-6