Harnessing the power of Artificial Intelligence to uncover patterns of violence

May 25, 2023, Analysis / Artificial Intelligence and Armed Conflict / Humanitarian Action / New Technologies

🕐 9 mins read

Fiona Terry
Head, ICRC Centre for Operational Research and Experience

Fabien Dany
Results-Based Management Adviser, ICRC

This month saw the successful conclusion to a year-long collaboration between the ICRC and the Swiss Data Science Centre – a joint venture between Switzerland's two federal institutes of technology, EPFL[1] and ETHZ[2] – on a research project to track patterns of violence.[3] The collaboration produced an algorithm, developed to reclassify open-source data according to international legal norms and trained by machine learning, that will permit the ICRC to have deeper insights into patterns of violence by armed forces and armed groups. The project's findings were launched during an event on the 9th of May at the ICRC Humanitarium.

In this post, Fiona Terry, head of the ICRC's Centre for Operational Research and Experience (CORE), and Fabien Dany, CORE adviser, describe the creation of this tool and how it will enable the ICRC to have faster and more accurate insights into who did what to whom and when, which will enhance its protection work and its analysis of violent events that threaten the safety of humanitarian personnel.

ICRC Humanitarian Law & Policy Blog · Harnessing the power of Artificial Intelligence to uncover patterns of violence

This research project to track patterns of violence is one of twelve projects included in the *Engineering for Humanitarian Action* initiative that was launched in December 2020. The initiative fosters the sharing of expertise of Switzerland's two federal institutes of technology to the benefit of humanitarian action. In a competitive process, ICRC personnel propose a humanitarian challenge that might have a technical solution, and a research proposal is

developed with teams from EPFL and/or ETHZ and submitted as a candidate for funding. Scientific committees from both universities review and rate the proposals for their academic credentials and innovative contributions to scientific knowledge, and ICRC personnel review them and rate them according to their importance to the ICRC, the feasibility of the research methodology proposed from the ICRC's perspective, and the potential for implementation of the results. The duration of projects must not exceed two years.

The challenge discussed here – to uncover 'patterns of violence' by armed forces and armed groups over time – stemmed directly from a recommendation of the ICRC's 2018 *Roots of Restraint in War* study. The study suggested that the ICRC ought to expand its monitoring of behaviour by armed forces and groups beyond IHL violations, but also to instances of restraint. Identifying genuine restraint can help the ICRC explore who or what might have influenced that restraint; provides an indication of the level of control commanders have over their soldiers/fighters; and permits the ICRC to interrogate its potential role in influencing that restraint through its confidential, bilateral dialogue with representatives of armed forces and groups.

Restraint, however, is a counterfactual and as such is difficult to identify. One way of doing so, it was suggested,[4] is to track patterns of violence to identify peaks and troughs in the pattern and then interrogate the possible factors behind the surge and drop in violence. The drop in violence might not constitute genuine restraint – it might just be the start of the rainy season. But it is only by seeing the change and asking the question that we can start to understand factors that might have influenced this behaviour, including the ICRC's own actions.

How was the challenge approached?

Due to the confidential nature of the ICRC's own data on the behaviour of armed forces and groups, we decided to use the *Armed Conflict Location and Event Data (ACLED)* dataset for this project because it is open-source, widely-used, and has broad coverage of political violence and protests around the world. In the ACLED dataset, each event is entered with a short description of what happened, is classified according to its 'event type', and includes details such as the number of fatalities and who was involved.

Since ACLED 'event type' classifications are not based on categories matching the rules of international humanitarian law or international human rights law, the first task at hand was to devise an automated method of reclassifying the data in accordance with these legal frameworks. The second task was to dig more deeply into the details of the events, because the ACLED dataset associates only one event type to each database entry, regardless of how many distinct episodes of violence the event might have contained. An attack on a village, for example, might result in four people killed, ten abducted and twenty houses burned. The ACLED dataset may only classify the event as 'attack' while other descriptions such as 'abduction,' 'property destruction' and 'killings' are equally relevant. This provides the dataset with a finer granularity. The data can then be visualized on charts plotted against time to show changes in the intensity and types of violence.

Event coding such as the one performed by ACLED is done by humans and is resource intensive. Given the volume of data to be parsed for reclassification, the process needed to be automated. So, the scientific team used a novel approach called 'prompt-entailment,' a process described by them in this *scientific publication*. In addition to reclassifying event types, we wanted to identify who perpetrated the violent act; who was the target; and whether we could distinguish between civilian and military casualties. For this part of the algorithm, 'extractive question answering', 'named entity recognition' and 'binary classification techniques' were used. The algorithm was then tested using ACLED data from Colombia, Ethiopia and Ukraine, representing several thousand events.

The results led to the development of a 'proof of concept' that takes a codebook as an input, through which domain experts can specify the types of events they are looking to reclassify. This codebook is then used with two standard, open-source, Natural Language Processing models and the algorithm outputs a reclassified dataset that also contains additional information on who was involved in the events. This approach offers greater flexibility to current methods, since event types can be added or modified, and requires limited training data which means a lower cost to changing the model. The code can be easily adapted to additional languages. An extra benefit to this approach is that any performance gains on the pre-trained open-source models used will improve the classification process, without extra workload.

In a nutshell, this rich collaboration between political and computer scientists from the ICRC, EPFL and ETHZ produced a model that uses artificial intelligence to give us a clearer and faster picture of evolutions in the use and type of violence by armed forces and groups. It will improve accuracy in the classification of events – using categories more closely related to the ICRC's mandate – and enrich the data with more granularity. The time savings associated with automating this process will allow ICRC colleagues to more fully triangulate events with other data sources, including the ICRC's own, to obtain greater accuracy. The visualization of patterns of violence will permit us to overlay other information such as political events which might have provoked violence or brought restraint and allow us to observe whether the topics discussed in the ICRC's confidential bilateral dialogue with commanders correlates with any changes in behaviour. The model can also be used to track patterns of violence that might threaten our field teams, and thus support decision-making processes on safety and security.

What now?

Not surprisingly, there is considerable henthusiasm at the ICRC for this model to become part of the data analysis tools at our disposal. But for this to happen we need to do two things: first, define with precision our use-cases to allow our analysts to use this reclassification tool, and subsequently develop the software architecture; and second, refine the codebook to overcome some of the current limitations noted with the identification and naming of actors of violence.

Fortunately, the Engineering for Humanitarian Action (EHA) initiative recently introduced a new workstream to facilitate the 'piloting and implementation' of the work achieved by the Humanitarian Action Challenges to ensure that the technology has real-world application to the challenges the ICRC faces in carrying out its humanitarian mandate. So we shall continue to adapt and test the algorithm and explore what it tells us about the behaviour of armed forces and groups and how we might influence it to comply with IHL.

[1] École polytechnique fédérale de Lausanne (Federal Institute of Technology, Lausanne)

[2] Eidgenössische Technische Hochschule Zürich (Federal Institute of Technology, Zürich)

[3] With significant contributions from other members of EPFL and ETHZ

[4] The suggestion came from Francisco Gutierrez-Sanin and is elaborated in this article: Francisco Gutierrez Sanin & Elisabeth Jean Wood, 'What Should We Mean by "Pattern of Political Violence"? Repertoire, Targeting, Frequency, and Technique,' *Perspectives on Politics* 15 (1), March 2017: pp. 20-41.

See also

- Fiona Terry, *Taking action, not sides: the benefits of humanitarian neutrality in war*, June 21, 2022
- Christopher Chen, *The future is now: artificial intelligence and anticipatory humanitarian action*, August 19, 2021
- Fiona Terry & Brian McQuinn, *Behind the scenes: The Roots of Restraint in War study*, June 18, 2018

Tags: ACLED, artificial intelligence, big data, Colombia, EPFL, Ethiopia, ETHZ, patterns of violence, Ukraine

You may also be interested in:

War on the coastline: mitigating civilian harm in the littorals

🕐 12 mins read

Analysis / Artificial Intelligence and Armed Conflict / Humanitarian Action / New Technologies Prashant Kahlon

With 80 per cent of the world's population living within 100 kilometers of a coastline, …

International humanitarian law and policy in Colombia: progress and challenges

🕐 12 mins read

Analysis / Artificial Intelligence and Armed Conflict / Humanitarian Action / New Technologies Mariana Chacón Lozano & Elizabeth Rushing

Over the last several years, while attention was turned to tackling the impact of the …