

Targeting of Infrastructure in the Middle East Database: Methodology Overview

Project Statement and Coding Overview

The Targeting of Infrastructure in the Middle East project tracks discrete targeting of civilian and environmental infrastructure by various actors involved in selected acute and protracted conflicts in the Middle East and North Africa (MENA).

This document provides an overview of how project researchers collect, verify, and clean incident data.

Overview of Data Collection Process

Who does the coding?

Incidents recorded in the project's TIME database are coded and collected by graduate and undergraduate students at Duke University and the University of New Hampshire following search and source protocols established by the faculty primary investigators (PIs).

Where is the data collected from?

The project codes a variety of primary and secondary sources, which includes human rights reports, UN documents, nongovernmental organizations (NGOs), international organizations (IOs) documents, think tank reports, and published news sources (including both Arabic and English media.) While there are some common sources across countries (e.g., UN-OCHA reports), incidents for specific countries are not collected from the same or uniform set of sources. Sources used vary by country due to differences in coverage for specific conflicts, which vary over time as well. For example, coverage in Yemen expanded over time with organizations becoming more systematic in data collection as the war became increasingly protracted.

How do we collect the data?

We do not collect data in real time; instead, we collect data for selected periods of time specific to each conflict case. For Yemen, Libya, and Syria, the database focuses on the period from 2011 to the present, when significant destruction of infrastructure took place in the wars following the regional uprisings. For longer, protracted conflicts, our data collection time frame varies. For Iraq, we look at incidents dating back to the 1991 Gulf War. For Palestine, we start collecting incidents in 2000 following the second *Intifada*. While both conflicts have longer time frames, we have chosen these time frames for incident collection of targeted infrastructure.

In a separate database, we compile aggregate reporting on the targeting of infrastructure through other means and over longer time periods. We use Rob Nixon's term "slow violence" to capture the long-term impacts of blockades, sieges, restrictions on reconstruction and maintenance and various other forms of restrictions on infrastructure development. The time period varies by availability of sources.

Overview of the Coding Process

How are incidents coded?

Each incident is coded according to a standardized set of guidelines on how to code actors involved in targeting, the infrastructure targeted, when the attack occurred, intentionality of the targeting, operational state of the infrastructure, and length of time for the infrastructure to be repaired. The guidelines ensure the consistency and validity of data and enable comparisons across countries and time. For each incident additional information is recorded where available, including incident IDs numbers, location of the attack, numbers of fatalities and injuries, event descriptions, and citations for data source.

Where can the data and more information be found?

Additional information on how we code actors involved in infrastructure targeting can be found in the document on Actor Methodology. Additional information on the types of infrastructure and sectors that we track and how we code them can be found in the document on Infrastructure Coding Methodology.

Overview of the Data Cleaning and Confirmation Process

The project's data cleaning process involves cleaning data to correct for variations in coding terminology and duplicated incidents. Coded incidents are reviewed to ensure consistency of coding within countries and comparability of data among countries. Additionally, errors in spelling that occur in incident coding are also corrected by the researcher cleaning the data, which occurs with names transliterated in different ways.

In addition to differences in the way incidents are reported among sources, the collection of incident data from different sources can result in duplicate incidents being reported in the database. To control for duplicates, researchers cleaning the data filter the incidents according to the incident date recorded. Researchers check incidents reported as having occurred on the same day for similar attributes like actors, infrastructure type, means of destruction, incident description, and location. If incidents share most of these attributes, they are treated as duplicates. Once duplicate incidents are identified, the incident with the greatest detail is selected to remain in the database and the source information for the duplicate incident(s) is recorded under the remaining incident in the Sources column. This process protects against overcounting attacks on infrastructure by various actors.

When sourcing incident data from databases, researchers perform a process of incident verification to ensure the accuracy of incidents reported. To verify reports of infrastructure destruction, an internet search is conducted using key details such as the location and date of the incident reported in the database. If no matching reports are found in English, an Arabic search is conducted. If new details about the incident are uncovered, they are incorporated into the coded information. The citation for the source used to verify the incident is listed in the Sources column alongside the citation of the database.

Data Uses/Implications

How is this data being used?

We use the data to better understand patterns of targeting infrastructure, including in what sectors, by whom, and with what impacts. We are interested in better analyzing the direct and indirect effects of targeting infrastructure on civilians, including health, livelihoods and development prospects, as well as ecosystem wellbeing. The dataset allows a more accurate sense of the frequency and scope of infrastructure targeting.

Who might be interested in using this data?

Our audiences include academics interested in the study of war and humanitarian action, legal scholarship on the protection of the environment/infrastructure during war and NGOs, human rights organizations, and policymakers interested in protecting civilians and the environment during and after war.