# ACLED-DS: A Large Multilingual Expert-Annotated Event Dataset for the Real World

**Anonymous ACL submission**

## Abstract

Coding, the method of labeling and organizing qualitative data, is commonly used in social science studies. For example, to provide an understanding of political violence around the world, an international non-profit group, ACLED (Armed Conflict Location and Event Data), has been collecting and coding reports of protests and conflicts for over a decade. Using this high-quality manually collected data, we create ACLED-DS, a dataset of 45,426 armed conflict events spanning 22 languages and 172 countries, with extensive coverage of region-specific entities.

Building on this real-world dataset, we motivate a modification to the traditional event extraction task. We propose the task of *abstractive* event extraction (AEE) and entity linking: events are extracted by a holistic understanding of the entire document, and all event arguments are normalized. This formulation simplifies applications such as aggregating information from diverse sources in different languages to understand global trends and patterns.

We introduce a novel zero-shot AEE system ZEST based on large language models. On ACLED-DS, ZEST achieves 77.6% and 82.9% $F_1$ on event detection and abstractive event argument extraction respectively. ZEST outperforms GoLLIE, a state-of-the-art information extraction model, even when tested on English data and after GoLLIE is fine-tuned on 12,000 examples. For the event linking subtask, ZEST achieves 40.8% compared to 11.0% of OneNet, a strong LLM-based baseline. While these results establish ZEST as a strong baseline for this new dataset, they also highlight important challenges in entity linking within this global and highly specialized context. [1]

[1]A subset of the dataset is submitted as supplementary material. The full dataset and code will be released upon publication.

## 1 Introduction

Event Extraction (EE) is a task that extracts structured information about events and their arguments from unstructured text, such as news articles. EE systems serve as important tools for studying various real-world phenomena, including disease outbreaks (Parekh et al., 2024; Consoli et al., 2024; Min et al., 2021), cybersecurity threats (Satyapanich et al., 2020), political conflicts (Hu et al., 2022), protests (Radford, 2020; Zhukov et al., 2019; Hürriyetoğlu et al., 2022a; Zavarella et al., 2022), and crime (Mostafazadeh Davani et al., 2019).

The goal of this paper is to *understand the requirements of effective EE for studying global real-world phenomena*, and to refine the EE task accordingly. We analyze data collected by the Armed Conflict Location & Event Data (ACLED) organization as a case study. ACLED is a non-profit organization that has been collecting and coding data on violent conflict and protest events around the globe for over a decade (Raleigh et al., 2010). Its data have been used by organizations like the United Nations' International Organization for Migration, the International Rescue Committee and the European Commission for tracking and predicting forced displacements and evaluating humanitarian efforts (ACLED, 2023).

Based on our analysis, we contribute a cleaned and partially reannotated ACLED data set (which we call ACLED-DS) to the NLP research community. We evaluate state-of-the-art EE models and large language models (LLMs) on this dataset, providing insights into the current capabilities and limitations of these systems for real-world event extraction tasks.

Monitoring global socio-political developments is a widely studied yet challenging application of EE (Hürriyetoğlu, 2021; Hürriyetoğlu et al., 2022b, 2023, 2024). ACLED-DS differs from existing event datasets in the following important ways:

**Multilinguality and Geographic Diversity.** To provide a *global* view, the analysis of outbreaks and conflicts extends beyond the global north to include the global south and international settings (Braha, 2012). While most existing EE datasets focus solely on English or Chinese, ACLED-DS, on the other hand, contains events from 172 countries reported in 22 languages.

**Aggregation of Extracted Events.** One of the main uses of event data is trend discovery and aggregate reporting (Li et al., 2019, 2020b, 2021a; Reddy et al., 2023). To simplify data aggregation, all event arguments in the gold labels of ACLED-DS are normalized, regardless of their form in the input document or the document's language. Unlike previous datasets that annotate at the sentence or span level, our annotations require comprehensive *abstractive* document understanding and inference-making. Figure 1 shows an example highlighting this difference.

**Tail Entities.** Entity linking enables the aggregation of event participants. Although commonly used entity databases like Wikidata (Wen et al., 2021) and Wikipedia (Li et al., 2019, 2020a) provide broad coverage, they lack the depth required for specialized domains. ACLED-DS contains a database of 9,777 entities, each with a single-paragraph expert-written description. These entities include Generic terms like "Students", but also entities that do not have a Wikipedia article: Specialized political entities (e.g., "Liwa' Al Hashemiyoun" - a political militia in Syria since 2023), Regional organizations (e.g., "NNO: Nagorik Nari Oikya" - the women's wing of Nagarik Oikya in Bangladesh), Institutional subdivisions (e.g., "Police Forces of Italy - Road Police" and "Police Forces of Italy - Guardia di Finanza"). These tail specialized entities make ACLED-DS particularly challenging for entity linking systems. This is especially important given that most LLMs are pretrained on Wikipedia entities and memorize a large portion of them.[2]

**Expert Annotations.** High-quality annotations are crucial for EE systems used in high-impact policy decisions, such as international peacemaking efforts (Andrea Ruggeri and Dorussen, 2011), as poor annotation quality can lead to biased inferences. Raleigh et al. (2010); Caselli and Huang (2012) argue that domain expert annotations are essential for this reason. In contrast, most existing document-level EE datasets rely on crowdsourcing (Ebner et al., 2020; Liu et al., 2024a; Ren et al., 2024; Wang et al., 2020), students (Li et al., 2021b), or weakly supervised methods (Li et al., 2023). ACLED data is annotated by approximately 200 regional experts, with multiple review rounds (See Section 4.1 for more information on the annotation process).

### 1.1 Contributions

This paper addresses the requirements of event extraction in real-world applications through the following contributions:

**The *Abstractive* Event Extraction (AEE) Task.** We propose a new AEE task that extracts events from whole documents according to a code book, where all event arguments must be normalized to either a numerical or categorical value, or a subset of the known entities in a given database. The formal definition of the task is in Section 3.

**The ACLED-DS Dataset.** We present ACLED-DS, a new document-level dataset for the multilingual AEE task. ACLED-DS distinguishes itself from other document-level EE datasets through its size, language coverage, and geographic diversity. See Table 4 for a comparison with several other datasets.

**Models for the Multilingual AEE Task** We evaluate several models from the literature on ACLED-DS. In addition, our ZEST, a novel zero-shot system for multilingual AEE which includes a novel zero-shot entity-linking component, achieves 77.6 and 82.9 $F_1$ on the event detection and event argument extraction subtasks of AEE, significantly outperforming all baselines.

We also show that ZEST outperforms the state-of-the-art combination of a zero-shot information extraction system GoLLIE (Sainz et al., 2024) and entity linker OneNet (Liu et al., 2024b) by 29.8% on entity linking on this dataset.

## 2 Related Work

Much of today's EE research follows the ACE05 project's (Walker et al., 2006) task formulation, which divides the work into sentence-level subtasks

[2] This issue goes beyond event datasets. For example, Cao et al. (2022) introduced a knowledge base question answering dataset on Wikidata where models without an entity linking module achieve 90% accuracy.

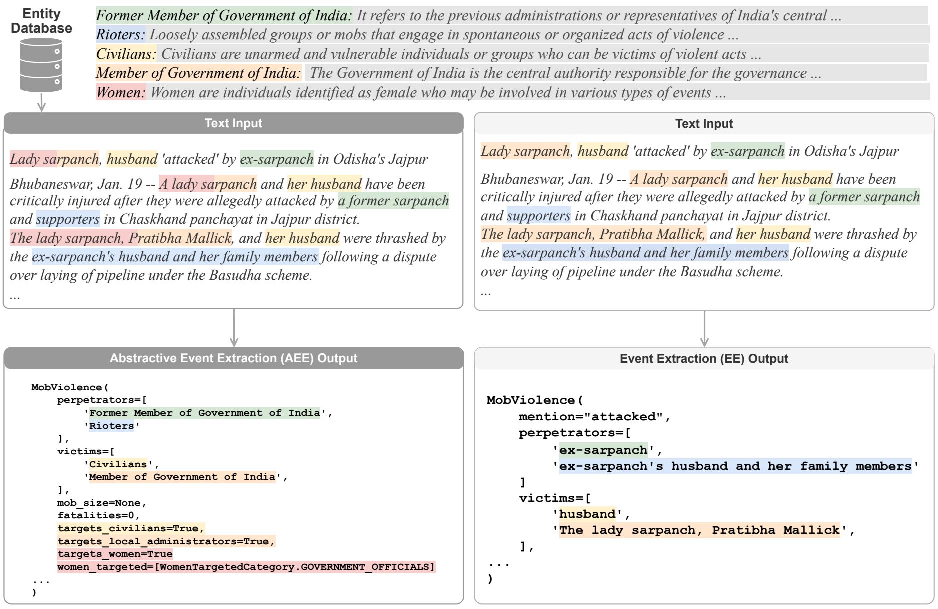



Figure 1: An example from ACLED-DS with its abstractive event annotation. The input text and annotations are summarized for this figure. A hypothetical extractive annotation for the same event is also provided for comparison.

using span-based intermediate annotations. Li et al. (2021b) extended EE to allow for arguments of an event to be from surrounding sentences, and introduced the concept of "most informative span" for arguments. Tong et al. (2022) introduced the DocEE dataset, where event arguments are on average scattered across 10 sentences in the document, fully realizing EE as a *document-level* task.

While most EE datasets focus on English and Chinese (Zhu et al., 2024; Ren et al., 2024; Walker et al., 2006), several datasets have been developed for other languages. These datasets, which widely vary in annotation quality, often focus solely on the simpler event detection subtask. Notable examples include: BKEE (Nguyen et al., 2024) for Vietnamese, InDEE-2019 (Maheshwari et al., 2019) for 5 Indic languages, MEE (Pouran Ben Veyseh et al., 2022) for Portuguese, Spanish, Polish, Turkish, Hindi, Japanese and Korean, Zavarella et al. (2014) for Bulgarian, Romanian and Turkish, Balali et al. (2022) for Farsi, Li et al. (2019) for Russian and Ukranian, Prabhu et al. (2019) for English, Spanish, Italian and French, Saetia et al. (2024) for Thai and Colruyt et al. (2023) for Dutch.

### 2.1 Entity Detection and Linking

Entity Linking (EL) is the task to link an entity mention in the document to a given entity database (Milne and Witten, 2008). EL typically follows the subtask of entity mention detection, where spans that denote an entity are marked. In our work, however, abstract EL (AEL) directly goes from the input text to a list of linked entities.

In this work, we primarily focus on zero-shot EL, which enables linking to a new entity database with no direct supervision. Logeswaran et al. (2019a) demonstrated the effectiveness of pre-trained models in zero-shot EL and introduced the ZESHEL dataset, which has become a popular benchmark for subsequent research. More recently, Xu et al. (2023) proposed a read-and-select framework for key stages of entity disambiguation and fine-tunes a RoBERTa-base model for this purpose. OneNet (Liu et al., 2024b) advanced the state of the art by introducing a three-module pipeline implemented by prompting LLMs.

## 3 The Abstractive EE (AEE) Task

The distinguishing factor in AEE is that it moves away from the surface form of the text and primarily focuses on grounding events in a predefined ontology like an entity database, and categorical, numerical, or string event arguments. The term *abstractive* has been used in other NLP tasks like OpenIE (Pei et al., 2023) and summarization (Radev et al., 2002) to refer to the concept of moving away from the surface form. Although AEE is, by definition, a generalization of EE that allows event arguments to be spans as well, this paper focuses specifically on cases where event arguments are not spans to highlight the key distinctions.

In AEE, we remove both the limitation that arguments must be spans and the requirement that they be explicitly mentioned in the text. Formally, we define the AEE task as follows:

**Definition 3.1.** We define event extraction codebook $C = (T, \mathcal{D}, S)$ where

- $T$ is the set of possible event types,
- Each $D \in \mathcal{D}$ is a domain such as integers, strings, or a set of known entities,
- $S = [(t_1, a_{1,1}, \ldots, a_{1,n_1}), ..., (t_m, a_{m,1}, \ldots, a_{m,n_m})]$ is a list of $m$ event signatures, where $n_i$ is the number of arguments for event type $t_i$, and $a_{i,j}$ is an argument with domain $D_{i,j} \in \mathcal{D}$.

**Definition 3.2.** The *Abstractive Event Extraction* (AEE) task is: given codebook $C = (T, \mathcal{D}, S)$ and writing $w \in W$, extract abstractive event(s) of the form $(t_i, v_1, ..., v_{n_i})$ from $w$, where $t_i$ is the $i$th event type in $T$, $v_j \in D_{i,j}$ and $n_i$ is the number of arguments for event type $t_i$ .

Figure 1 shows an example of AEE annotations from ACLED-DS on the left side, and how it would have been annotated under traditional EE formulation. In the AEE annotation, $t_i$ = `MobViolence` $\in T$, the first two arguments, `perpetrators` and `victims` represent the two sides in the violence, with $D_{i,1}, D_{i,2}$ being the set of all subsets of possible entities from the event database, and the domain of `targets_women`, $D_{i,7}$ is $\{\text{True}, \text{False}\}$. In this example, the AEE annotation of the event captures the following information: 1) Women were specifically targeted 2) The targeted women were government officials 3) No fatalities occurred in the incident 4) The size of the mob was not specified in the article (`mob_size=None`). These event arguments are not present as text spans in the article and must be inferred and captured using using boolean, enum and numerical fields, representing examples of *abstract* event arguments. This significantly simplifies aggregation of event information, for instance to study the extent and types of violence against women around the globe.

This definition unifies several traditionally separate subtasks. Traditional document-level EE (the annotation on the right side of the figure) typically involves the following sequential steps (Huang et al., 2024):

1. Event Detection (ED): Identifying trigger spans and their corresponding event types (e.g., `MobViolence` event) 2. Event Argument Extraction (EAE): For each event, identifying argument spans and their roles 3. Entity Linking (EL): Resolving entity coreferences and linking event arguments to their corresponding entries in an entity database.

AEE eliminates the need for intermediate annotations such as event triggers and entity mentions. It also forgoes annotating multiple spans for the same event argument, as is common practice in traditional EE. Instead, AEE datasets directly annotate events and linked entities, allowing AEE systems to generate and be evaluated against these annotations. This reduces the annotation work for AEE datasets, leading to less noise in labels. It also enables easier evaluation using a simple exact match metric, mitigating the shortcomings of automatic metrics for EE (Lu et al., 2024).

This streamlined approach focuses instead on these three modified subtasks:

- Event Detection (ED): given the writing $w$, find the event type of the event(s) mentioned. $\text{ED}(w, C) = t$
- Abstract Event Argument Extraction (AEAE): find the event arguments for non-entity arguments, given the event type. $\text{AEAE}(w, C, t) = [v_1, \ldots]$ for non-entity arguments.
- Abstract Entity Linking (AEL): given the writing $w$ and a gold event type $t$, find a subset of the entity database involved in that event and assign them to the correct event argument. $\text{AEL}(w, C, t) = [v_1, \ldots]$ for entity arguments.

ED is similar to how it is defined in prior work, AEAE is to identify normalized values for non-entity arguments, given the event type. AEL is the same, but for entity arguments. An end-to-end AEE system needs to tackle the task of ED first, then use the predicted event type(s) (instead of the gold event type) to tackle AEAE and AEL in parallel.

## 4 ACLED-DS, a Multilingual Expert-Annotated AEE Dataset

Originally published in 2010, Armed Conflict Location and Event Data (ACLED) (Raleigh et al., 2010) focused on civil war, subnational and transnational violent events in 50 unstable countries. It has since expanded to track more types of political violence event, as well as civil unrest events, in 243 countries and territories in 100 languages in near real-time (Sam Jones, 2022; ACLED, 2023).

ACLED-DS, a new dataset for AEE, is a subset of ACLED, partially reannotated and processed to be viable as an NLP task dataset. Each writing is a news article, whose main event is annotated. Our setting is similar to document-level single-event setting of DocEE (Tong et al., 2022; Liu et al., 2024a) where only the *main* event is annotated, excluding background or historical events that are typically mentioned in writings to provide more context. Each annotation contains the event type, and a list of entity arguments and non-entity arguments associated with the main event and their role. There are 25 event types in ACLED-DS all in the domain of socio-political developments, covering a wide range from peaceful protests to the use of chemical weapons. See Appendix H for the full list of event types and their arguments.

ACLED-DS covers 22 languages: English, Arabic, Spanish, Portuguese, Turkish, Burmese, Korean, French, German, Indonesian, Italian, Persian, Ukrainian, Russian, Somali, Nepali, Hebrew, Chinese, Polish, Dutch, Hindi, and Japanese. These languages are typologically diverse (Clark et al., 2020), and span several high and low resource languages. To the best of our knowledge, this is the first event extraction dataset for Burmese, Indonesian, Hebrew, Somali and Nepali, and covers the most number of languages than any other event dataset. ACLED-DS contains events from 172 countries. For these reasons, it can be used to evaluate how models perform when processing cultural or region-specific information and entities.

Table 4 shows a summary of differences between ACLED-DS and several other document-level event datasets.

### 4.1 Construction of ACLED-DS

**ACLED Annotation and Review Process.** We start from the publicly available expert annotations of ACLED. ACLED annotations are done by a group of around 200 experts and is updated on a weekly basis. It sources documents from news media, international organizations, NGO and security reports, and local partner organizations and select social media channels.

These writings go through a multi-step review and quality assurance process (ACLED, 2020). The annotation of events is done at a regional level (e.g. the Middle East, Africa etc.) by experts of those regions. These experts have local language skills and knowledge about regional conflicts, and sometimes live within the country they cover. The annotations are then merged by a research manager who reviews these data for inter-coder reliability across the region. Researchers use an annotation tool that provides them with the up-to-date list of entities and locations, and communicate with each other to clarify difficult annotation decisions. After merging regional data, another round of manual reviewing is performed by another expert.

ACLED data is annotated at the event level, with each event potentially spanning multiple news articles. The annotations are done by consulting multiple sources, including maps and images to pinpoint the event's location, and involved entities. The main challenge in creating ACLED-DS therefore, is to ensure document-level annotations only contain information that can be extracted or inferred from the text of one document. Here we briefly describe the steps taken. For a more detailed explanation, see Appendix A.

- We start from the ACLED data from the first 11 months of 2024. Overall, this includes 285,777 events, each paired with one or more URLs pointing to documents about that event, and one event annotation.
- Data filtering and cleaning: We remove articles whose text does not contain enough information related to the event, for instance, social media posts that contain images. We prompt GPT-4o to detect the language of each document. We obtain the full text of writings from the provided URLs, and clean them by removing advertisements etc. using an LLM prompt.
- Simplifying location: We reannotate the location event arguments to match this definition: "*The location argument is the most specific place that is supported by the writing*".
- Schematization: We clean the event schema for the dataset, and convert annotation to Python code, following Wang et al. (2023).

**Entity Database** ACLED annotates entities involved in each event. There are in total 9,777 unique entities. Note that this database is a superset of the 5,087 entities in ACLED-DS and 3,658 in the dev and test splits of it. This means entity linking systems evaluated on this dataset, need to be good at differentiating *distracting* entities.

To make domain knowledge available to models in a realistic way, we also provide a one-paragraph description for each entity. These descriptions are meant to provide entity linking models with enough context and domain knowledge to understand domain entities, especially the long tail (Mallen et al., 2023). This closely follows the formulation of Zeshel entity linking dataset (Logeswaran et al., 2019b). Entity linking systems are expected to use these entity descriptions to find the ones that are related to the input document. Appendix D includes a sample of the entities and their description.

**Data splits** We provide validation and test sets in 22 languages, and a large training set in English. The data split is across time, meaning that the events in the training set are from the first 6 months of 2024, and the events in validation and tests sets are from July to November 2024. This mimics the real-world setting where the distribution of events and entities might change over time. Because of this split, 75.3% of entities in the validation and test sets are not seen in the training set. The split between validation and test sets is random. Appendix C contains the language, event type and geographical distribution of ACLED-DS.

## 5 Systems for AEE

In section 7, we use ACLED-DS's large training set to establish strong baselines in the supervised setting. However, as manual annotation in ACLED is a major undertaking, we also wish to understand if it is possible to handle new domains without training data. Here, the zero-shot models assume no access to training data in any language, and use only information about the schema and the domain provided in the form of an event codebook and an entity database.

### 5.1 Our Models

We propose ZEST, a zero-shot AEE system, consisting an in-context learning ED module, an AEAE module adapted from Code4Struct (Wang et al., 2023), and a novel AEL module.

**In-context learning for ED** For an in-context learning ED model, we prompt an LLM with instructions and in-context examples to select the most suitable event type from the ontology. Given that the list of event types ($T$) is relatively small (25 in the case of ACLED-DS), event detection can be done as a zero-shot in-context learning task. The prompt (Table 12) includes the input writing $w$ and a list of event types and their descriptions. The task is to return the most likely event type $t$. We use chain-of-thought (Wei et al., 2023) for this prompt.

**AEAE using Code4Struct** We modify the instructions of Code4Struct (Wang et al., 2023) to adapt it to the document-level AEAE setting. Specifically, we instruct the LLM to directly output the event arguments and their role given the input article, instead of sentence-level and span extraction which is the setting of the original paper. This is done using a prompt (Table 16) that given $w$ and the event type signature for $t$, outputs all non-entity argument values.

**A Novel Abstractive Entity Linker** For AEL, ZEST does not extract entity spans, but attempts to directly link entities from the given input document. Given an event type, event database, and the input document, the AEL task is to find a subset of the entity database that are involved in the event. Naive in-context learning cannot handle the large number of entities (9,777 in the case of ACLED-DS). Therefore, ZEST tackles this in multiple stages: it first narrows down the number of candidates, and in the second stage more closely filters down the set. Finally, the third stage assigns each entity to the correct event argument, for example assigns "Member of Government of India" to `victims` in Figure 1.

Stage 1, entity retrieval: We first create a vector database by embedding all entities and their descriptions using an embedding model. Then, given the input document, ZEST generates multiple queries to search in the entity database. Each query returns $N$ entities. The set of all returned entities is given to the next stage.

Stage 2, entity filtering: For each candidate entity from stage 1, and their description, this stage uses another prompt (Table 14) to find evidence of each entity's involvement in the event. We then remove the ones for which we cannot find any evidence.

The last step is to match each entity with its correct event argument. For this, we use another prompt (Table 15) that given a list of entities and event argument roles, outputs a mapping between the two.

### 5.2 Comparison with Prior Works

**GoLLIE (Sainz et al., 2024)** For the ED and AEAE subtasks, we compare with GoLLIE (Sainz et al., 2024), a model specifically instruction-tuned for information extraction tasks. GoLLIE achieved state-of-the-art zero-shot generalization results on several event extraction datasets, including (Li et al., 2021b). We note that GoLLIE has been trained on RAMS (Ebner et al., 2020) and ACE05 document-level event datasets as well. Additionally, we fine-tune GoLLIE on ACLED-DS in the supervised setting. Therefore, GoLLIE is a strong candidate to test on ED and AEAE subtasks in the zero-shot setting.

**OneNet (Liu et al., 2024b)** For AEL, we choose OneNet, a state-of-the-art few-shot entity linking model. OneNet uses retrieval and entity descriptions to find the best match, given an entity mention. We use GoLLIE event argument spans as entity mentions.

# 6 Experiments and Results

## 6.1 Experiment Setup

We use greedy decoding for all models. For all experiments involving GPT-4o, we use the version `gpt-4o-2024-08-06`. We also experiment with the 70B-parameter instruction-tuned LLaMA 3.1 (Dubey et al., 2024), since it has shown strong results in multilingual benchmarks.

## 6.2 Metrics

To evaluate a predicted event against a gold event from ACLED-DS, we first normalize the location arguments using a lookup in the OpenStreetMap geographic database. We then use simple string equality to calculate precision, recall and micro-averaged $F_1$ (Manning et al., 2008).

For ED, we compare the predicted event type against the gold event type, and report the micro averaged **ED** $F_1$. For AEAE, we force the gold event type as the first part of the model output, and have it generate event arguments and their values $\{(a'_1, v'_1), ...\}$. We then consider this set as the returned result, and calculate its precision, recall and $F_1$ against the gold $\{(a_1, v_1), ...\}$ and report **AEAE** $F_1$. In other words, two arguments are considered equal if their argument *and* values match.

For entities, we report **AEL** $F_1$, which is the result of comparing the entity IDs between prediction and gold, calculated similar to AEAE $F_1$ but only on entity arguments.

# 7 Results

## 7.1 ED

Table 1 shows the $F_1$ metric for each system and language. In the zero-shot setting, ZEST (GPT-4o), ZEST (LLaMA-3.1-70B) and GoLLIE achieve 77.6%, 53.4% and 25.0% on average across all languages. ZEST with GPT-4o significantly outperforms GoLLIE in all language, including English which GoLLIE has been trained with. ZEST with the weaker LLaMA-3.1-70B also outperforms GoLLIE in Enlish and all other languages, except for German (de) and Chinese (zh).

Interestingly, fine-tuning GoLLIE on the English ACLED-DS training set improves its performance on 17 out of 22 languages, for an average of 19.5%.

Burmese (my), Arabic (ar), Somali (so), Nepali (ne), Hebrew (he), Hindi (he) and Japanese (ja) are among the weakest languages for GoLLIE, some of which imporve after fine-tuning. However, Arabic, Burmese, Somali and Hebrew do not improve much. This can be attributed to the fact that CodeLLaMA has not been pretrained on these scripts, and its tokenizer is especially weak for them. For instance, due to its inefficient tokenizer, the number of tokens in the English codebook plus the input text in Somali is longer than the model's context window for almost all examples.

## 7.2 AEAE

Table 2 shows the results for AEAE. In the zero-shot setting, ZEST (GPT-4o), ZEST (LLaMA-3.1-70B) and GoLLIE achieve 82.9%, 74.2% and 26.0% on average across all languages. Here we observe similar trends to ED in terms of relative performance of our four systems. For AEAE, given that different languages have different distributions of event types, and therefore event arguments, this introduces some additional variance across languages.

Here, the improvement resulting from fine-tuning GoLLIE for Burmese and Somali is slightly more than ED. Overall, Hebrew, Burmese, Arabic and Somali are the languages where the best system struggles with, compared to other languages.

## 7.3 AEL

Table 3 shows the results for AEL. ZEST (GPT-4o) by far performs the best, with 49.1%, and OneNet (GPT-4o) achieves only 13.1%. Both systems use the same underlying entity embeddings and LLM. Therefore, we conclude that the big gap between the two in English is due to the fact that OneNet relies on entity spans extracted by GoLLIE. Note that GoLLIE has been instruction-tuned using several named entity recognition datasets, and achieves state-of-the-art results on those (Sainz et al., 2024). The gap indicates that ACLED-DS cannot be tack-

| Model | en | ar | es | pt | tr | my | ko | fr | de | id | it | fa | uk | ru | so | ne | he | zh | pl | nl | hi | ja | Avg |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Zero-shot** | | | | | | | | | | | | | | | | | | | | | | | |
| ZEST (GPT-4o) | 66.3 | 65.0 | 74.0 | 77.1 | 80.5 | 58.8 | 78.7 | 69.8 | 92.1 | 80.0 | 91.4 | 86.3 | 64.3 | 69.9 | 69.9 | 76.8 | 68.9 | 83.7 | 91.5 | 91.1 | 83.7 | 87.3 | 77.6 |
| ZEST (L-70B) | 47.3 | 54.7 | 57.1 | 68.7 | 52.9 | 45.0 | 27.7 | 53.7 | 76.7 | 48.3 | 61.3 | 71.5 | 57.7 | 45.3 | 54.7 | 25.2 | 56.3 | 30.6 | 65.3 | 65.0 | 40.7 | 68.3 | 53.4 |
| GoLLIE-7B | 29.1 | 10.3 | 20.0 | 43.4 | 18.0 | 0.9 | 4.1 | 29.3 | 77.0 | 14.0 | 53.5 | 15.2 | 33.6 | 26.1 | 0.0 | 0.0 | 3.2 | 54.0 | 57.1 | 44.4 | 7.8 | 9.2 | 25.0 |
| **Full-shot** | | | | | | | | | | | | | | | | | | | | | | | |
| GoLLIE-7B | 66.6 | 8.5 | 57.0 | 25.0 | 61.6 | 4.0 | 87.0 | 49.8 | 61.8 | 73.1 | 58.0 | 37.4 | 16.0 | 29.7 | 2.9 | 34.4 | 7.5 | 46.2 | 62.3 | 70.8 | 51.6 | 66.7 | 44.5 |

Table 1: ED $F_1$ results on ACLED-DS test set.

| Model | en | ar | es | pt | tr | my | ko | fr | de | id | it | fa | uk | ru | so | ne | he | zh | pl | nl | hi | ja | Avg |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Zero-shot** | | | | | | | | | | | | | | | | | | | | | | | |
| ZEST (GPT-4o) | 80.2 | 69.8 | 87.7 | 92.2 | 85.8 | 74.5 | 92.3 | 82.8 | 88.1 | 91.8 | 94.1 | 77.1 | 66.5 | 76.9 | 68.8 | 88.6 | 59.8 | 88.5 | 81.3 | 93.2 | 91.5 | 93.1 | 82.9 |
| ZEST (L-70B) | 73.2 | 59.9 | 83.3 | 86.2 | 73.5 | 64.2 | 81.3 | 74.9 | 80.3 | 86.2 | 81.9 | 68.5 | 54.1 | 57.0 | 63.1 | 80.4 | 57.1 | 80.5 | 71.6 | 81.6 | 84.8 | 87.9 | 74.2 |
| GoLLIE-7B | 45.2 | 14.4 | 32.1 | 42.7 | 8.1 | 2.0 | 6.4 | 48.0 | 52.9 | 27.4 | 46.4 | 22.2 | 27.5 | 34.5 | 0.5 | 0.9 | 5.3 | 61.9 | 47.0 | 32.7 | 4.9 | 9.9 | 26.0 |
| **Full-shot** | | | | | | | | | | | | | | | | | | | | | | | |
| GoLLIE-7B | 69.3 | 11.7 | 63.2 | 36.0 | 61.5 | 26.5 | 47.9 | 58.5 | 70.3 | 78.5 | 70.1 | 32.1 | 18.1 | 42.8 | 16.5 | 44.7 | 18.4 | 55.0 | 59.8 | 76.8 | 61.9 | 62.7 | 49.2 |

Table 2: AEAE $F_1$ results on ACLED-DS test set. ZEST is using Abstract Code4Struct.

| Model | en | ar | es | pt | tr | my | ko | fr | de | id | it | fa | uk | ru | so | ne | he | zh | pl | nl | hi | ja | Avg |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GoLLIE + OneNet (GPT-4o) | 14.2 | 4.6 | 5.1 | 11.8 | 7.1 | 0.8 | 4.8 | 12.4 | 28.5 | 4.2 | 22.7 | 13.1 | 25.9 | 8.9 | 0.0 | 0.0 | 3.8 | 4.2 | 36.7 | 19.9 | 11.2 | 3.0 | 11.0 |
| ZEST (GPT-4o) | 37.5 | 39.3 | 48.5 | 43.9 | 36.6 | 46.7 | 38.3 | 45.0 | 32.0 | 33.6 | 45.6 | 49.1 | 30.3 | 30.9 | 51.4 | 44.0 | 34.8 | 47.8 | 50.3 | 46.1 | 37.9 | 27.5 | 40.8 |

Table 3: AEL $F_1$ results on ACLED-DS test set. All methods are zero-shot.

led using entity spans due to the existence of abstract entities.

## 8 Conclusions

This paper introduces the task of abstractive event extraction (AEE), which more closely matches the requirements of event extraction for real-world applications. We have derived a large high-quality dataset for the AEE task, in 22 languages, from the expert-annotated data created by ACLED.

We introduced ZEST, a novel zero-shot AEE system, that establishes a strong baselines for this new dataset, outperforming strong baselines like GoLLIE from the literature. We hope that including this dataset in future instruction-tuned models like GoLLIE, will help expand their capabilities.

## Limitations

This paper focuses on document-level event extraction. As such, it does not include event coreference resolution across multiple documents (Eirew et al., 2022), which is another necessary component for aggregate analysis of events. Existing event coreference models like Gao et al. (2024) can be adapted to AEE, and we leave that to future work.

ACLED-DS also excludes other common forms of information extraction, like relation extraction, and only contains annotations for event and entity extraction.

Domain of ACLED-DS is violent conflict and protest events and focuses on differentiating relatively minor differences between different event types (for instance, a Peaceful Protest that has or has not been met with excessive force are separate event types in ACLED-DS). Datasets like GLEN (Li et al., 2023) instead, focus on broad coverage of many types of events, from conflict events to sports and more.

## Ethics Statement

No human subjects were involved in this study. We will release ACLED-DS in accordance with the ACLED Terms of Use. ACLED data do not contain personally identifiable information (e.g. names of individuals or mobile device IDs), and cannot be used to track individuals. No crowdsourcing was performed as part of this paper.

## References


ACLED. 2020. Coding review process. Accessed: 2024-09-30.

ACLED. 2023. 2023 impact report. Accessed: 2024-09-28.

ACLED. 2023. Armed conflict location & event data project (acled) codebook. https://acleddata.com/knowledge-base/codebook/. Last updated: 27 September 2024.

Theodora-Ismene Gizelis Andrea Ruggeri and Han Dorussen. 2011. Events data as bismarck's sausages? intercoder reliability, coders' selection, and data quality. *International Interactions*, 37(3):340–361.

Ali Balali, Masoud Asadpour, and Seyed Hossein Jafari. 2022. Cofee: A comprehensive ontology for event extraction from text. *SSRN Electronic Journal*.

Luisa Bentivogli, Pamela Forner, Claudio Giuliano, Alessandro Marchetti, Emanuele Pianta, and Kateryna Tymoshenko. 2010. Extending English ACE 2005 corpus annotation with ground-truth links to Wikipedia. In *Proceedings of the 2nd Workshop on The People's Web Meets NLP: Collaboratively Constructed Semantic Resources*, pages 19–27, Beijing, China. Coling 2010 Organizing Committee.

Dan Braha. 2012. Global civil unrest: Contagion, self-organization, and prediction. *PLOS ONE*, 7(10):1–9.

Shulin Cao, Jiaxin Shi, Liangming Pan, Lunyiu Nie, Yutong Xiang, Lei Hou, Juanzi Li, Bin He, and Hanwang Zhang. 2022. KQA pro: A dataset with explicit compositional programs for complex question answering over knowledge base. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 6101–6119, Dublin, Ireland. Association for Computational Linguistics.

Tommaso Caselli and Chu-Ren Huang. 2012. Sourcing the crowd for a few good ones: Event type detection. In *Proceedings of COLING 2012: Posters*, pages 1239–1248, Mumbai, India. The COLING 2012 Organizing Committee.

Jonathan H. Clark, Eunsol Choi, Michael Collins, Dan Garrette, Tom Kwiatkowski, Vitaly Nikolaev, and Jennimaria Palomaki. 2020. TyDi QA: A benchmark for information-seeking question answering in typologically diverse languages. *Transactions of the Association for Computational Linguistics*, 8:454–470.

Camiel Colruyt, Orphée De Clercq, Thierry Desot, and Véronique Hoste. 2023. Eventdna: a dataset for dutch news event extraction as a basis for news diversification. *Language Resources and Evaluation*, 57(1):189–221.

Samuel Colvin, Eric Jolibois, Hasan Ramezani, Adrian Garcia Badaracco, Terrence Dorsey, David Montague, Serge Matveenko, Marcelo Trylesinski, Sydney Runkle, David Hewitt, and Alex Hall. 2024. Pydantic.

Sergio Consoli, Peter Markov, Nikolaos I. Stilianakis, Lorenzo Bertolini, Antonio Puertas Gallardo, and Mario Ceresa. 2024. Epidemic information extraction for event-based surveillance using large language models. In *Proceedings of Ninth International Congress on Information and Communication Technology*, pages 241–252, Singapore. Springer Nature Singapore.

Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, et al. 2024. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*.

Seth Ebner, Patrick Xia, Ryan Culkin, Kyle Rawlins, and Benjamin Van Durme. 2020. Multi-sentence argument linking. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8057–8077, Online. Association for Computational Linguistics.

Alon Eirew, Avi Caciularu, and Ido Dagan. 2022. Cross-document event coreference search: Task, dataset and modeling. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 900–913, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Qiang Gao, Bobo Li, Zixiang Meng, Yunlong Li, Jun Zhou, Fei Li, Chong Teng, and Donghong Ji. 2024. Enhancing cross-document event coreference resolution by discourse structure and semantic information. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 5907–5921, Torino, Italia. ELRA and ICCL.

Yibo Hu, MohammadSaleh Hosseini, Erick Skorupa Parolin, Javier Osorio, Latifur Khan, Patrick Brandt, and Vito D'Orazio. 2022. ConfliBERT: A pre-trained language model for political conflict and violence. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 5469–5482, Seattle, United States. Association for Computational Linguistics.

Kuan-Hao Huang, I-Hung Hsu, Tanmay Parekh, Zhiyu Xie, Zixuan Zhang, Prem Natarajan, Kai-Wei Chang, Nanyun Peng, and Heng Ji. 2024. TextEE: Benchmark, reevaluation, reflections, and future challenges in event extraction. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 12804–12825, Bangkok, Thailand. Association for Computational Linguistics.

Ali Hürriyetoğlu, editor. 2021. *Proceedings of the 4th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE 2021)*. Association for Computational Linguistics, Online.

Ali Hürriyetoğlu, Osman Mutlu, Fırat Duruşan, Onur Uca, Alaeddin Gürel, Benjamin J. Radford, Yaoyao

Dai, Hansi Hettiarachchi, Niklas Stoehr, Tadashi Nomoto, Milena Slavcheva, Francielle Vargas, Aaqib Javid, Fatih Beyhan, and Erdem Yörük. 2022a. Extended multilingual protest news detection - shared task 1, CASE 2021 and 2022. In *Proceedings of the 5th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE)*, pages 223–228, Abu Dhabi, United Arab Emirates (Hybrid). Association for Computational Linguistics.

Ali Hürriyetoğlu, Hristo Tanev, Surendrabikram Thapa, and Gökçe Uludoğan, editors. 2024. *Proceedings of the 7th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE 2024)*. Association for Computational Linguistics, St. Julians, Malta.

Ali Hürriyetoğlu, Hristo Tanev, Vanni Zavarella, Reyyan Yeniterzi, Erdem Yörük, and Milena Slavcheva, editors. 2023. *Proceedings of the 6th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text*. INCOMA Ltd., Shoumen, Bulgaria, Varna, Bulgaria.

Ali Hürriyetoğlu, Hristo Tanev, Vanni Zavarella, and Erdem Yörük, editors. 2022b. *Proceedings of the 5th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE)*. Association for Computational Linguistics, Abu Dhabi, United Arab Emirates (Hybrid).

Albert Q. Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, Lélio Renard Lavaud, Marie-Anne Lachaux, Pierre Stock, Teven Le Scao, Thibaut Lavril, Thomas Wang, Timothée Lacroix, and William El Sayed. 2023. Mistral 7b. *Preprint*, arXiv:2310.06825.

Manling Li, Sha Li, Zhenhailong Wang, Lifu Huang, Kyunghyun Cho, Heng Ji, Jiawei Han, and Clare Voss. 2021a. The future is not one-dimensional: Complex event schema induction by graph modeling for event prediction. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 5203–5215.

Manling Li, Ying Lin, Joseph Hoover, Spencer Whitehead, Clare Voss, Morteza Dehghani, and Heng Ji. 2019. Multilingual entity, relation, event and human value extraction. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations)*, pages 110–115, Minneapolis, Minnesota. Association for Computational Linguistics.

Manling Li, Alireza Zareian, Ying Lin, Xiaoman Pan, Spencer Whitehead, Brian Chen, Bo Wu, Heng Ji, Shih-Fu Chang, Clare Voss, Daniel Napierski, and Marjorie Freedman. 2020a. GAIA: A fine-grained multimedia knowledge extraction system. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 77–86, Online. Association for Computational Linguistics.

Manling Li, Qi Zeng, Ying Lin, Kyunghyun Cho, Heng Ji, Jonathan May, Nathanael Chambers, and Clare Voss. 2020b. Connecting the dots: Event graph schema induction with path language modeling. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 684–695, Online. Association for Computational Linguistics.

Sha Li, Heng Ji, and Jiawei Han. 2021b. Document-level event argument extraction by conditional generation. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 894–908, Online. Association for Computational Linguistics.

Sha Li, Qiusi Zhan, Kathryn Conger, Martha Palmer, Heng Ji, and Jiawei Han. 2023. GLEN: General-purpose event detection for thousands of types. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 2823–2838, Singapore. Association for Computational Linguistics.

Minghui Liu, MeiHan Tong, Yangda Peng, Lei Hou, Juanzi Li, and Bin Xu. 2024a. DocEE-zh: A fine-grained benchmark for Chinese document-level event extraction. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 637–649, Miami, Florida, USA. Association for Computational Linguistics.

Xukai Liu, Ye Liu, Kai Zhang, Kehang Wang, Qi Liu, and Enhong Chen. 2024b. OneNet: A fine-tuning free framework for few-shot entity linking via large language model prompting. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 13634–13651, Miami, Florida, USA. Association for Computational Linguistics.

Lajanugen Logeswaran, Ming-Wei Chang, Kenton Lee, Kristina Toutanova, Jacob Devlin, and Honglak Lee. 2019a. Zero-shot entity linking by reading entity descriptions. *arXiv preprint arXiv:1906.07348*.

Lajanugen Logeswaran, Ming-Wei Chang, Kenton Lee, Kristina Toutanova, Jacob Devlin, and Honglak Lee. 2019b. Zero-shot entity linking by reading entity descriptions. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3449–3460, Florence, Italy. Association for Computational Linguistics.

Ilya Loshchilov and Frank Hutter. 2017. Decoupled weight decay regularization. *Preprint*, arXiv:1711.05101.

Yi-Fan Lu, Xian-Ling Mao, Tian Lan, Chen Xu, and Heyan Huang. 2024. Beyond exact match: Semantically reassessing event extraction by large language models. *Preprint*, arXiv:2410.09418.

Ayush Maheshwari, Hrishikesh Patel, Nandan Rathod, Ritesh Kumar, Ganesh Ramakrishnan, and Pushpak Bhattacharyya. 2019. Tale of tails using rule augmented sequence labeling for event extraction. *Preprint*, arXiv:1908.07018.

Alex Mallen, Akari Asai, Victor Zhong, Rajarshi Das, Daniel Khashabi, and Hannaneh Hajishirzi. 2023. When not to trust language models: Investigating effectiveness of parametric and non-parametric memories. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 9802–9822, Toronto, Canada. Association for Computational Linguistics.

Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. 2008. *Introduction to Information Retrieval*. Cambridge University Press, USA.

David Milne and Ian H. Witten. 2008. Learning to link with wikipedia. In *Proceedings of the 17th ACM Conference on Information and Knowledge Management*, CIKM '08, page 509–518, New York, NY, USA. Association for Computing Machinery.

Bonan Min, Benjamin Rozonoyer, Haoling Qiu, Alexander Zamanian, Nianwen Xue, and Jessica MacBride. 2021. ExcavatorCovid: Extracting events and relations from text corpora for temporal and causal analysis for COVID-19. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 63–71, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Aida Mostafazadeh Davani, Leigh Yeh, Mohammad Atari, Brendan Kennedy, Gwenyth Portillo Wightman, Elaine Gonzalez, Natalie Delong, Rhea Bhatia, Arineh Mirinjian, Xiang Ren, and Morteza Dehghani. 2019. Reporting the unreported: Event extraction for analyzing the local representation of hate crimes. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 5753–5757, Hong Kong, China. Association for Computational Linguistics.

Thi-Nhung Nguyen, Bang Tien Tran, Trong-Nghia Luu, Thien Huu Nguyen, and Kiem-Hieu Nguyen. 2024. BKEE: Pioneering event extraction in the Vietnamese language. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 2421–2427, Torino, Italia. ELRA and ICCL.

OpenStreetMap contributors. 2017. Planet dump retrieved from https://planet.osm.org . https://www.openstreetmap.org.

Tanmay Parekh, Anh Mac, Jiarui Yu, Yuxuan Dong, Syed Shahriar, Bonnie Liu, Eric Yang, Kuan-Hao Huang, Wei Wang, Nanyun Peng, and Kai-Wei Chang. 2024. Event detection from social media for epidemic prediction. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 5758–5783, Mexico City, Mexico. Association for Computational Linguistics.

Kevin Pei, Ishan Jindal, and Kevin Chang. 2023. Abstractive open information extraction. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 6146–6158, Singapore. Association for Computational Linguistics.

Amir Pouran Ben Veyseh, Javid Ebrahimi, Franck Dernoncourt, and Thien Nguyen. 2022. MEE: A novel multilingual event extraction dataset. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 9603–9613, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Suhan Prabhu, Pranav Goel, Alok Debnath, and Manish Shrivastava. 2019. Incorporating sub-word level information in language invariant neural event detection. In *Proceedings of the 16th International Conference on Natural Language Processing*, pages 36–44, International Institute of Information Technology, Hyderabad, India. NLP Association of India.

Maxim Rabinovich, Mitchell Stern, and Dan Klein. 2017. Abstract syntax networks for code generation and semantic parsing. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1139–1149, Vancouver, Canada. Association for Computational Linguistics.

Dragomir R. Radev, Eduard Hovy, and Kathleen McKeown. 2002. Introduction to the special issue on summarization. *Computational Linguistics*, 28(4):399–408.

Benjamin J. Radford. 2020. Multitask models for supervised protests detection in texts. *Preprint*, arXiv:2005.02954.

Clionadh Raleigh, Rew Linke, Håvard Hegre, and Joakim Karlsen. 2010. Introducing acled: An armed conflict location and event dataset. *Journal of peace research*, 47(5):651–660.

Jeff Rasley, Samyam Rajbhandari, Olatunji Ruwase, and Yuxiong He. 2020. Deepspeed: System optimizations enable training deep learning models with over 100 billion parameters. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '20, page 3505–3506, New York, NY, USA. Association for Computing Machinery.

Revanth Gangi Reddy, Yi R Fung, Qi Zeng, Manling Li, Ziqi Wang, Paul Sullivan, and Heng Ji. 2023. Smartbook: Ai-assisted situation report generation. *arXiv preprint arXiv:2303.14337*.

Yubing Ren, Yanan Cao, Hao Li, Yingjie Li, Zixuan ZM Ma, Fang Fang, Ping Guo, and Wei Ma. 2024. DEIE: Benchmarking document-level event information extraction with a large-scale Chinese news dataset. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 4592–4604, Torino, Italia. ELRA and ICCL.

Baptiste Rozière, Jonas Gehring, Fabian Gloeckle, Sten Sootla, Itai Gat, Xiaoqing Ellen Tan, Yossi Adi, Jingyu Liu, Romain Sauvestre, Tal Remez, Jérémy Rapin, Artyom Kozhevnikov, Ivan Evtimov, Joanna Bitton, Manish Bhatt, Cristian Canton Ferrer, Aaron Grattafiori, Wenhan Xiong, Alexandre Défossez, Jade Copet, Faisal Azhar, Hugo Touvron, Louis Martin, Nicolas Usunier, Thomas Scialom, and Gabriel Synnaeve. 2023. Code llama: Open foundation models for code. *Preprint*, arXiv:2308.12950.

Chanatip Saetia, Areeya Thonglong, Thanpitcha Amornchaiteera, Tawunrat Chalothorn, Supawat Taerungruang, and Pakpoom Buabthong. 2024. Streamlining event extraction with a simplified annotation framework. *Frontiers in Artificial Intelligence*, 7:1361483. ECollection 2024.

Oscar Sainz, Iker García-Ferrero, Rodrigo Agerri, Oier Lopez de Lacalle, German Rigau, and Eneko Agirre. 2024. GoLLIE: Annotation guidelines improve zero-shot information-extraction. In *The Twelfth International Conference on Learning Representations*.

Sam Jones. 2022. New expansion brings ACLED to full global coverage. Accessed: 2024-09-28.

Taneeya Satyapanich, Francis Ferraro, and Tim Finin. 2020. Casie: Extracting cybersecurity event information from text. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(05):8749–8757.

MeiHan Tong, Bin Xu, Shuai Wang, Meihuan Han, Yixin Cao, Jiangqi Zhu, Siyu Chen, Lei Hou, and Juanzi Li. 2022. DocEE: A large-scale and fine-grained benchmark for document-level event extraction. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 3970–3982, Seattle, United States. Association for Computational Linguistics.

Lewis Tunstall, Edward Beeching, Nathan Lambert, Nazneen Rajani, Kashif Rasul, Younes Belkada, Shengyi Huang, Leandro von Werra, Clémentine Fourrier, Nathan Habib, Nathan Sarrazin, Omar Sanseviero, Alexander M. Rush, and Thomas Wolf. 2023. Zephyr: Direct distillation of lm alignment. *Preprint*, arXiv:2310.16944.

Christopher Walker, Stephanie Strassel, Julie Medero, and Kazuaki Maeda. 2006. Ace 2005 multilingual training corpus ldc2006t06.

Xiaozhi Wang, Ziqi Wang, Xu Han, Wangyi Jiang, Rong Han, Zhiyuan Liu, Juanzi Li, Peng Li, Yankai Lin, and Jie Zhou. 2020. MAVEN: A Massive General Domain Event Detection Dataset. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1652–1671, Online. Association for Computational Linguistics.

Xingyao Wang, Sha Li, and Heng Ji. 2023. Code4Struct: Code generation for few-shot event structure prediction. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 3640–3663, Toronto, Canada. Association for Computational Linguistics.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, and Denny Zhou. 2023. Chain-of-thought prompting elicits reasoning in large language models. *Preprint*, arXiv:2201.11903.

Haoyang Wen, Ying Lin, Tuan Lai, Xiaoman Pan, Sha Li, Xudong Lin, Ben Zhou, Manling Li, Haoyu Wang, Hongming Zhang, Xiaodong Yu, Alexander Dong, Zhenhailong Wang, Yi Fung, Piyush Mishra, Qing Lyu, Dídac Surís, Brian Chen, Susan Windisch Brown, Martha Palmer, Chris Callison-Burch, Carl Vondrick, Jiawei Han, Dan Roth, Shih-Fu Chang, and Heng Ji. 2021. RESIN: A dockerized schema-guided cross-document cross-lingual cross-media information extraction and event tracking system. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Demonstrations*, pages 133–143, Online. Association for Computational Linguistics.

Brandon T. Willard and Rémi Louf. 2023. Efficient guided generation for large language models. *Preprint*, arXiv:2307.09702.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander M. Rush. 2019. Huggingface's transformers: State-of-the-art natural language processing. *Preprint*, arXiv:1910.03771.

Zhenran Xu, Yulin Chen, Baotian Hu, and Min Zhang. 2023. A read-and-select framework for zero-shot entity linking. *arXiv preprint arXiv:2310.12450*.

Vanni Zavarella, Dilek Küçük, Hristo Tanev, and Ali Hürriyetoğlu. 2014. Event extraction for Balkan languages. In *Proceedings of the Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics*, pages 65–68, Gothenburg, Sweden. Association for Computational Linguistics.

Vanni Zavarella, Hristo Tanev, Ali Hürriyetoğlu, Peratham Wiriyathammabhum, and Bertrand De Longueville. 2022. Tracking COVID-19 protest events in the United States. shared task 2: Event database replication, CASE 2022. In *Proceedings of the 5th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE)*, pages 209–216, Abu Dhabi, United Arab Emirates (Hybrid). Association for Computational Linguistics.

Xin Zhang, Yanzhao Zhang, Dingkun Long, Wen Xie, Ziqi Dai, Jialong Tang, Huan Lin, Baosong Yang, Pengjun Xie, Fei Huang, Meishan Zhang, Wenjie Li, and Min Zhang. 2024. mgte: Generalized long-context text representation and reranking models for multilingual text retrieval. *Preprint*, arXiv:2407.19669.

Mengna Zhu, Zijie Xu, Kaisheng Zeng, Kaiming Xiao, Mao Wang, Wenjun Ke, and Hongbin Huang. 2024. CMNEE:a large-scale document-level event extraction dataset based on open-source Chinese military news. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 3367–3379, Torino, Italia. ELRA and ICCL.

Yuri M Zhukov, Christian Davenport, and Nadiya Kostyuk. 2019. Introducing xsub: A new portal for cross-national data on subnational violence. *Journal of Peace Research*, 56(4):604–614.

## A Details of Creation of ACLED-DS

In this section, we describe the steps we took in creating ACLED-DS, including cleaning and re-annotation of certain event arguments. The general process involved a combination of spot-checks conducted by domain experts, and several rounds of improvements from two authors of this paper and help from prompting a strong large language model (GPT-4o) for simple but cumbersome steps.

**Data Filtering and Cleaning** We obtain all events from the first 11 months of 2024. Overall, this includes 285,777 events, each paired with one or more URLs pointing to documents about that event, and an annotated event. After analyzing the data, we realized that many social media posts in the data are accompanied by an image (e.g. protest fliers), and the text alone is not enough to annotate the event. Therefore, we exclude social media posts. We also remove the 1% longest and shortest writings, because very short ones (often from local partner organizations' reports) do not include enough context for annotating the event, and very long texts are often concatenation of multiple articles, included by error. We prompt GPT-4o to detect the language of each document. We obtain the full text of writings from the provided URLs, and clean them by removing advertisements etc. using an LLM prompt.

**Balancing the Dataset** Given the ACLED data attempts to mirror the real world, in each language, the number of events with each event type is heavily influenced by the political stability of countries with that language. For example, most events in English and Korean are Peaceful Protest events, while most Burmese events (from Myanmar) are Armed Clash. This imbalance can affect the performance of AI systems, therefore, we downsample the most frequent events in each language. This results in 114,743 events. We further limit the data to languages that have at least 500 events, to be able to have enough events for each language for model evaluations.

**Converting ACLED to a document-level dataset** As mentioned above, each event in ACLED data is paired with one or more documents. To facilitate document-level EE research, we convert this to one event per document by pairing each document with the event it is associated with. However, it is possible that the document only contains partial information from the event, for example, only the "attacker" in a Mob Violence event, while the "victim" is mentioned in a different document.

We process each event-document pair independently in the following steps, ensuring each event annotation only contains information from its paired document. In the end, we deduplicate these pairs by only keeping the most complete event and its paired document.

**Location** For reannotation, we use the original ACLED location annotations to consult the OpenStreetMap geographic database (OpenStreetMap contributors, 2017) to find the full hierarchy of location above the neighborhood level for each event. We then start from the lowest location level and remove the items that are not supported by the writing, until we reach one that is. We then keep that location and all levels above that. A carefully designed LLM prompt was used for this last stage. The final location arguments were spot-checked by the authors, and 97% of them were correct according to the above definition.

**Schematization** ACLED uses the same event argument roles for all event types, resulting in some argument roles being always empty for some event types or the argument names being too generic. We define separate event argument roles for each event type. For example, we remove "fatalities" argument from "Peaceful Protest" and rename "actor 1" to "Abductor" for the "Abduction or forced disappearance" event type. We also provide a short description for each event type, and expert descriptions for each event argument, to facilitate the development of zero-shot models.

Following the recent trend in event extraction, we use Python code to represent annotations. This has been shown to improve the performance of various supervised (Sainz et al., 2024) and few-shot (Wang et al., 2023) models because it makes the labels closer to the code data many language models have been pre-trained on. Furthermore, this enables the use of constrained decoding (Rabinovich et al., 2017; Willard and Louf, 2023) algorithms to eliminate malformed outputs. Appendix H presents the full schema for ACLED-DS.

| | ACE05 | DocEE | ACLED-DS | WikiEvents | RAMS | Maven-Arg |
|---|---|---|---|---|---|---|
| Languages | English, Chinese, Arabic | English, Chinese* | 22 Languages | English | English | English |
| Event Argument Types | string | string | string, numerical, categorical, boolean | string | string | string |
| Entity Database | Wikipedia* | - | Domain Expert Curated | - | - | - |
| Avg. Doc len | 2,410 | 2052.1 | 4,477 | 3,919 | 591 | 1,589 |
| Annotators | LDC Annotation Group | Crowdworkers | ACLED Domain Experts | Graduate Students | Crowdworkers | Crowdworkers |
| Num. Docs | 1,635 | 64,214* | 45,426 | 246 | 9,124 | 4,480 |
| Num. Events | 8,878 | 64,214* | 45,426 | 3,951 | 9,124 | 98,591 |
| Source | News | Wikipedia | News | Wikipedia | News | Wikipedia |

Table 4: Comparison of ACLED-DS with several other document-level event extraction datasets. ACLED-DS covers the most languages, and has the longest documents on average. Because it is an AEE dataset, it also covers more types of event arguments. '-' means the dataset does not have an entity linking subtask.
*: Includes aggregate statistic of multiple datasets. Entity linking was added to ACE05 by Bentivogli et al. (2010). Liu et al. (2024a) created another dataset with the same ontology as DocEE in Chinese.

# B Comparison of ACLED-DS and Several Other Document-Level EE Datasets

# C ACLED-DS Statistics

Table 5 shows the number of events per language in each split of ACLED-DS.

Tables 6 and Figure 2 show the distribution of event types and country-level locations of events in ACLED-DS respectively.

We show another example from ACLED-DS with abstractive event annotation in Figure 3

# D Entity description examples

Tables 7, 8, 9, 10 and 11 contain several examples of ACLED-DS entities and their description.

# E Prompts use in the experiments

Here we provide the prompts used in various baselines, including ZEST. The syntax used is the Jinja2 template language, which supports Python-like loops (`{% for %}{% endfor %}`), conditions (`{% if %}{% endif %}`), variables (`{{ var }}`) and comments (#).

After variables are substituted in the prompt, the string under `# instruction` and `# input` are sent to the LLM as the system prompt and user message respectively.

# F Hyperparameters

All fine-tuned models were fine-tuned with batch size 64 for 3 epochs. The final model checkpoint was selected for evaluation. Learning rate of $2 \times 10^{-5}$, cosine learning rate scheduler and AdamW (Loshchilov and Hutter, 2017) optimizer are used.

Training is done on a machine with 4 NVIDIA A100 GPU with 80GB, using DeepSpeed (Rasley et al., 2020) and the Transformers (Wolf et al., 2019) library. In total, fine-tuning took around 3 hours.

For access to GPT-4o model, we used the OpenAI API. We report the aggregate usage of this API, which costed around $2,000. We access LLaMA-3.1-70B-instruct model via Azure's serverless hosting. The total cost for querying this model was about $1500 for all experiments.

For access to OpenStreetMap, we used the publicly hosted version via Nominatim `https://nominatim.openstreetmap.org/`

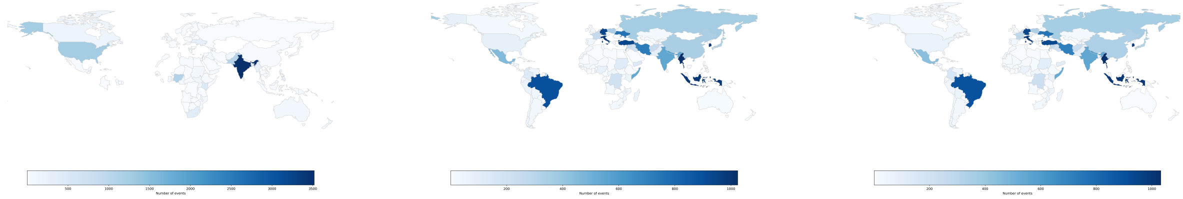
# G Baseline Details

We use GoLLIE-7B model which is based on CodeLLaMA (Rozière et al., 2023). For zero-shot experiments, we provide it with event descriptions in the same format of the data it has been instruction-tuned with, a Python class definition for each event type where the type's description is provided in the docstring and typical trigger words are included in the class comments. For the supervised experiments, we fine-tune it to given the input text, directly output event type, arguments and entities in one go.

OneNet is based on the 7B-parameter Zephyr (Tunstall et al., 2023) which is an instruction-tuned version of Mistral (Jiang et al., 2023). In our preliminary experiments, OneNet

| Language (language code) | Train | Dev | Test |
|---|---|---|---|
| English (en) | 12000 | 1000 | 1000 |
| Arabic (ar) | - | 1000 | 1000 |
| Spanish (es) | - | 1000 | 1000 |
| Portuguese (pt) | - | 1000 | 1000 |
| Burmese (my) | - | 1000 | 1000 |
| Korean (ko) | - | 1000 | 1000 |
| French (fr) | - | 1000 | 1000 |
| Turkish (tr) | - | 1000 | 1000 |
| German (de) | - | 1000 | 1000 |
| Indonesian (id) | - | 1000 | 1000 |
| Italian (it) | - | 1000 | 1000 |
| Persian/Farsi (fa) | - | 956 | 956 |
| Ukrainian (uk) | - | 717 | 717 |
| Nepali (ne) | - | 585 | 585 |
| Chinese (zh) | - | 563 | 563 |
| Somali (so) | - | 555 | 555 |
| Hebrew (he) | - | 518 | 518 |
| Russian (ru) | - | 455 | 455 |
| Dutch (nl) | - | 380 | 380 |
| Polish (pl) | - | 342 | 342 |
| Japanese (ja) | - | 330 | 330 |
| Hindi (hi) | - | 312 | 312 |
| Total | 12000 | 16713 | 16713 |

Table 5: ACLED-DS statistics per language. Language names and the abbreviation used throughout the paper is also specified.



(a) Train set (b) Validation set (c) Test set

Figure 2: Distribution of event locations in ACLED-DS. Note that the dataset includes more specific locations, but here we only plot the country level. In addition to being linguistically diverse, ACLED-DS is also geographically diverse. The distribution of the train set is skewed towards India, because it only contains English events.

Table 6: The number of event types in all splits of ACLED-DS. While imbalanced, the distribution of event types tracks that of the real world. For example, by far the most common among these events are peaceful protests.

| **Event Type** | **Count** |
|---|---|
| GovernmentRegainsTerritory | 48 |
| NonStateActorOvertakesTerritory | 142 |
| ArmedClash | 4068 |
| ExcessiveForceAgainstProtestors | 68 |
| ProtestWithIntervention | 1480 |
| PeacefulProtest | 21035 |
| ViolentDemonstration | 1560 |
| MobViolence | 2336 |
| AirOrDroneStrike | 1873 |
| SuicideBomb | 15 |
| ShellingOrArtilleryOrMissileAttack | 2073 |
| RemoteExplosiveOrLandmineOrIED | 872 |
| Grenade | 197 |
| SexualViolence | 120 |
| Attack | 3713 |
| AbductionOrForcedDisappearance | 788 |
| Agreement | 120 |
| Arrest | 964 |
| ChangeToArmedGroup | 648 |
| DisruptedWeaponsUse | 977 |
| BaseEstablished | 15 |
| LootingOrPropertyDestruction | 1369 |
| NonViolentTransferOfTerritory | 70 |
| OtherStrategicDevelopment | 875 |
| Total | 45426 |

performed poorly on ACLED-DS, therefore, we replace Zephyr with a stronger LLM, and call the resulting entity linking model OneNet (GPT-4o). Since OneNet expects entity spans as input, we first conduct EAE using GoLLIE to obtain entity argument spans. Following the OneNet setting, we retrieve 64 entities for each entity span, and run them through OneNet. We use GPT-4o instead of Zephyr, due to its higher performance.

Both OneNet and ZEST have an entity retrieval component. We use gte-multilingual-base model from Zhang et al. (2024) to create dense embeddings for entities from their name and description.

For ED, GoLLIE predicts both the event type and its trigger, which we discard. For EAE, GoLLIE expands its prediction to include the event type, trigger, and associated arguments. We implement a two-stage Event Detection approach using GoLLIE. Initially, we predict from six general event categories: Battle, Protest, Riot, ExplosionOrRemoteViolence, ViolenceAgainstCivilians, and StrategicDevelopment. After predicting the general event type, we further use GoLLIE to predict the corresponding subtype. The Event Argument Extraction (EAE) is subsequently performed based on this predicted subtype.

To adapt our data to GoLLIE's format, we provide detailed descriptions for each event type, several typical trigger words, and short descriptions for each argument type.

Following (Logeswaran et al., 2019a), we perform an initial entity reduction by identifying the top 64 most relevant entities for a given argument mention. We then employ LLM to evaluate whether each mention matches one of 64 entities in a pointwise manner given the contextual information. The process yields a refined and filtered set of potential entities. In the dual-perspective entity linking stage, we leverage the LLM to perform entity linking through two complementary approaches: con-

| Entity Name | Entity Description |
| --- | --- |
| Women | Women are adult human females who can play diverse roles in society, ranging from caregivers and economic participants to political and social activists. They may be involved in a variety of social, economic, and political events, sometimes facing unique challenges such as discrimination or violence. Women's roles and their societal impact can be profound, as seen in their involvement in protests, advocacy for rights, and even in conflict situations where they may be victims or participants. Globally, women continue to strive for gender equality and empowerment, often organizing and mobilizing to address issues affecting their communities and themselves. |
| Men | Men are adult human males who may be involved in a variety of societal roles and activities. As an entity, men can be participants in diverse events ranging from everyday community interactions to more extreme scenarios such as protests, violence against civilians, and riots. Men, as a group, can be both perpetrators and victims of violence, including sexual violence, as evidenced in various global incidents. Their involvement in these events can be influenced by cultural, social, and political contexts. This entity operates globally across all countries and societies. |
| Police Forces of the United States | The Police Forces of the United States are a collective entity composed of various local, state, and federal law enforcement agencies tasked with maintaining public order, enforcing laws, and ensuring public safety across the nation. These forces include municipal police departments, sheriff's offices, and specialized agencies such as the Federal Bureau of Investigation. They are recognized for their involvement in a wide range of activities, from managing public protests and investigating crimes to ensuring security during emergencies. While they play a crucial role in law enforcement, they have also faced scrutiny and legal challenges related to incidents of misconduct and use of force. Their operations are governed by state and federal laws, and they are accountable to governmental oversight bodies. They have been active since at least 1993 and continue to operate across the United States. |
| Military Forces of Russia | The Military Forces of Russia are the armed forces of the Russian Federation, responsible for national defense and military operations both within and outside Russia. Established in 2000, they operate under the command of the President of Russia, who is the supreme commander-in-chief. These forces are composed of various branches, including the Ground Forces, Navy, and Air Force, along with strategic missile troops and airborne troops. They have been involved in international military operations, peacekeeping missions, and domestic security tasks. The Russian military is recognized for its significant involvement in various conflicts, including actions in Ukraine, Syria, and other regions, often collaborating with or opposing other nations' forces. The Military Forces of Russia are known for their extensive use of armored vehicles, aerial support, and advanced military technology. |

Table 7: 20 entities in ACLED-DS entity database – Part 1

| Entity Name | Entity Description |
|---|---|
| Protestors | Protestors are individuals or groups who actively participate in demonstrations to express opposition or demand action on specific issues. They can be found globally and may engage in peaceful protests or civil disobedience to draw attention to their causes. Protestors often advocate for political, social, or economic changes and can be associated with various movements, including anti-corruption, electoral fairness, and human rights. While they primarily aim for peaceful expression, their activities can sometimes lead to confrontations with authorities or opposing groups. Protestors play a crucial role in civil society by challenging perceived injustices and influencing public discourse and policy. |
| Civilians | Civilians are unarmed, non-combatant individuals who are often vulnerable to violence and conflict, particularly in areas of political or social unrest. They can be affected by or involved in a wide range of events, including riots, protests, and violence, as seen in various global contexts. Civilians may participate in social movements, such as protests at educational institutions, or be subject to violence and negotiation processes in conflict zones, like settlements in Syria or mass violence in Colombia. Their involvement can manifest in active participation in civic actions or as victims of political and criminal violence, highlighting their diverse roles and the threats they face in unstable environments. |
| Rioters | Rioters are loosely assembled groups or mobs that engage in violent and disruptive behavior during demonstrations or spontaneously, often in response to perceived injustices or grievances. They may be civilians acting without inherent organization, and their actions typically involve confrontations with law enforcement or other entities. Rioters can be motivated by various social, political, or economic factors and are known to participate in actions such as vandalism, clashes, and other forms of violence. Their activities can occur in any country and are often part of broader social movements or tensions. |
| Students | Students are individuals enrolled in educational institutions, ranging from primary schools to universities, and are often involved in various social and political activities. They can be a diverse and dynamic group that participates in protests, movements against discrimination, and other forms of activism, sometimes leading to confrontations with law enforcement or political opposition. Students can also be impacted by external conflicts, such as gang violence, which may directly affect their safety and educational environment. While they are typically associated with learning and academic pursuits, students have historically played significant roles in advocating for change and challenging established systems, sometimes at the risk of becoming involved in violent or controversial situations. |

Table 8: 20 entities in ACLED-DS entity database – Part 2

| Entity Name | Entity Description |
|---|---|
| Farmers | Farmers are individuals or communities engaged in agriculture, responsible for cultivating crops and raising livestock. They operate globally and can be involved in various socio-political and economic events such as land disputes, protests, and negotiations affecting their livelihoods. Farmers often face challenges like resource competition, violence from armed groups, and policy changes impacting their work conditions and income. Their role is crucial in food production and sustainability, and they frequently interact with governments, organizations, and other agricultural stakeholders to address issues like land rights, security, and agricultural policies. |
| Labor Group | A “Labor Group" is a collective of workers united to advocate for their rights and interests in various sectors of the economy. These groups often engage in activities such as protests, strikes, and negotiations to address issues related to working conditions, wages, and employment security. They may also become involved in larger civil unrest, participating in events like riots or demonstrations to exert pressure on employers or authorities. While not typically associated with violent activities, labor groups can sometimes be connected to broader social movements that may encounter conflicts with law enforcement or political entities. Labor groups operate globally, often organized at local, national, or industry levels, and play a crucial role in labor relations and policy advocacy. |
| Guatemalan Group | The “Guatemalan Group" refers to a collective of Guatemalan immigrants and workers who engage in activism, particularly around labor rights, in countries like the United States and Australia. This group is involved in protests and rallies advocating for fair labor conditions and justice for immigrant workers. Their activism is exemplified through participation in events such as May Day rallies, where representatives like Eder Juarez highlight issues such as wage theft and lack of employee rights, making them a voice for immigrant labor struggles. |
| JI: Jamiat-e-Islami | Jamiat-e-Islami (JI) is a significant political and military organization in Afghanistan, primarily composed of ethnic Tajiks. Established in the 1970s, it played a crucial role in the resistance against the Soviet invasion and later in the Afghan civil war. Historically aligned with prominent leaders such as Ahmad Shah Masoud, JI has maintained influence in Afghan politics, often representing non-Pashtun interests. Despite the Taliban’s dominance, JI continues to be active, reflecting ongoing ethnic and political tensions within the country. Its members, including prominent diplomats and officials, have been involved in key governmental roles and resistance efforts against various regimes. |

Table 9: 20 entities in ACLED-DS entity database – Part 3

| Entity Name | Entity Description |
|---|---|
| Chang Tribal Group | The Chang Tribal Group is an indigenous community in India, primarily located in the state of Nagaland. Represented by the Chang Wedoshi Setshang (CWS), the group is known for advocating for their rights and addressing local grievances, particularly in the educational sector. They have been involved in protests to demand better resources and support from the government, as seen in their actions to secure transportation for Sao Chang College. The group's activities underscore their active role in seeking improved living and educational conditions for their community. |
| SD: Solidarity Party | The SD: Solidarity Party, also known as Solidariedade, is a political organization based in Brazil that is categorized as a political militia. It is known for its involvement in violent actions against civilians, often linked to political motives. The party has been associated with political figures in vulnerable positions, such as José Erlânio Firmiano, a city councilor who was assassinated in Alagoas. The Solidarity Party remains active in Brazilian politics, highlighting ongoing challenges related to political violence in the region. |
| Los Motonetos Gang | The Los Motonetos Gang is a political militia group operating primarily in Mexico, known for using violence to further their political aims. The gang gained notoriety for its involvement in riots and for the use of high-caliber weapons, which has resulted in significant unrest and necessitated interventions by local and national security forces, including the Municipal Police, State Preventive Police, and the Mexican Army. The group's influence and operational capacity were highlighted following the assassination of their presumed leader, Juan Hernández López, also known as El Fayo, which led to armed protests and heightened security measures in the region of San Cristóbal de las Casas, Chiapas. |
| Mebri Tribal Group | The Mebri Tribal Group is an indigenous community in Indonesia, specifically located in Papua. They are actively involved in advocating for the recognition and protection of their ancestral land rights. The group is known for organizing protests to demand fair compensation for the use of their land by government projects, such as healthcare infrastructure. They emphasize negotiation and dialogue with government authorities to resolve land disputes, as exemplified by their demands directed at the Indonesian Ministry of Health regarding land claims in areas under development. |

Table 10: 20 entities in ACLED-DS entity database – Part 4

| Entity Name | Entity Description |
|---|---|
| Ara Communal Group | The Ara Communal Group is a communal entity based in the town of Ara, located in the western countryside of As-Suwayda, Syria. Formed in 2024, the group is involved in regional socio-political activism and has participated in significant anti-Hayat Tahrir al-Sham protests across Idlib and Aleppo. These protests have called for political changes including the resignation of the group's leader "al-Jolani," the release of detainees, and the dismantling of the General Security Apparatus. The group has also been linked to incidents of remote violence, such as assassination attempts using explosive devices, amidst a backdrop of security instability and weak law enforcement in areas controlled by regime forces. The Ara Communal Group remains active and continues to influence political dynamics in the region. |
| Back the Blue | "Back the Blue" is a slogan and movement within the United States that expresses support for law enforcement officers. It is often used by individuals and groups, including political supporters, during protests and public demonstrations to show solidarity with police forces. The phrase is commonly associated with conservative and pro-law enforcement sentiments, frequently appearing in contexts where participants oppose policies perceived as critical of the police or supportive of police reform. Ëack the Bluečan also signify a broader political stance that emphasizes law and order. |
| Nalia Communal Group | The Nalia Communal Group is a factional community group based in Nalia village, located in the Lohagara Upazila of Narail, India. It is characterized by internal conflict, with power struggles between different factions, notably those led by Shaukat Khan and Ravi Khan. The group has been involved in violent clashes, often requiring police intervention to restore order. These conflicts are primarily driven by issues of dominance within the community, and the group remains active in its region. |
| ZPR: For Justice and Order | ZPR: For Justice and Order, also known as Za Pravdu i Red, is a political militia operating in Bosnia and Herzegovina since 2020. It is involved in political activities and protests, aiming to address issues of governance and electoral integrity. The group is led by Nebojša Vukanović and has been active in organizing demonstrations against political corruption and foreign exploitation of natural resources. ZPR is also linked to political candidates in regional elections, such as Slaviša Pavlović, whose affiliation with the group highlights its engagement in local politics. |

Table 11: 20 entities in ACLED-DS entity database – Part 5

```
# instruction
You are tasked with determining the best matching Event types for a given news article. You will be
     provided with annotation guidelines and a news article to analyze. Your goal is to identify
     the most relevant event types and rank them in order of their match to the article content.

# input
Here is the news article you need to analyze:
{{ article }}

Now, carefully review the annotation guidelines for various event types:

{% for ed in event_definitions.items() %}
[{{ loop.index }}] "{{ ed[0] }}": {{ ed[1] }}

{% endfor %}


1. For each event type, determine how well it matches the article content. Consider the following
     factors:
  - How closely the event description aligns with the main focus of the article
  - The presence of key actors or entities mentioned in the event type description
  - The occurrence of specific actions or outcomes associated with the event type

2. Rank the event types based on their relevance to the article content. Only include event types
     that have a meaningful connection to the article.

3. Output your results using the following format:
   - List the relevant event types in descending order of match quality
   - Use the ">" symbol to separate the event types

Your output should look like this:

[Explain your reasoning for the event types you decide to include, and their order]

event_type_1 > event_type_2 > ...

Provide only the ranked list of event types in your final answer.
```

Table 12: Prompt for event type detection (ED) using GPT-4o and LLaMA-3.1-70B.

```
You will be given a news article about an event. Your task is to identify all potential Entities
    who are directly or indirectly involved in the event. Then, write a very short Wikipedia
    paragraph describing each entity in the general sense.

An Entity is defined as an individual, group, collective, or organization involved in an event.
    This includes:
  * Organized armed groups with political purposes (e.g. "Hezbollah", "ISIS")
  * Organizations, governments, and political parties (e.g. "BJP: Bharatiya Janata Party", "
    Government of India", "Democratic Party of U.S.")
  * Ethnic, religious, social or occupational groups (e.g. "Jewish Group", "Muslim Group", "Women",
     "Students", "Farmers", "Journalists", "Teachers", "Lawyers")
  * General terms describing people involved (e.g., "Rioters", "Protestors", "Civilians", "Labor
    Group")

When identifying Entities, follow these guidelines:

1. Be as thorough as possible. Think about what groups are implicitly or indirectly involved in the
     event. Ask yourself:
  - Can the identity group (religion, gender, occupation etc.) of the victims or perpetrators be
    inferred? If so, you should create an entity for that group.
  - Does the event involve workers or unions, or is it a labor issue? If so, you should add "Labor
    Group" as an entity.
  - Does the event in any way involve students, school or university? If so, you should add "
    Students" as an entity.
  - Does the event involve women in any way? If so, you should add "Women" as an entity.
  - Does the event involve civilians? If so, you should add "Civilians" as an entity.
  - Is the event a protest or a riot? If so, you should add "Protestors" or "Rioters" as an entity.
  - Does the event involve an unknown or unspecified group? If so, you should add one of "
    Unidentified Armed Group", "Unidentified Gang", "Unidentified Communal Group" etc. as an
    entity.
  - Given the country the event is taking place in, what are the major political parties, religious
     groups, armed groups, or social movements that could be involved? Consider cultural context
    of the region, like common religions, ethnicities etc.
  - And the like.

2. Include alternative names or spellings of each entity if mentioned in the article

3. For individuals, infer their role, affiliation, or social group as explained above.

4. For each entity you identify, think about its affiliated, parent or member groups. For example,
    if a politician is mentioned, think about their political party or any other group they are
    associated with. If a union is mentioned, think about the workers or labor groups it
    represents.

Use a scratchpad to think through your process:

<scratchpad>
[Your thought process here, including your answer to the above questions]
</scratchpad>

Then, present your output in the following JSON format. Output as many entities as you can possibly
     think of.
<entity_list>
{
    "entity 1": "Wikipedia paragraph 1",
    "entity 2": "Wikipedia paragraph 2",
    ...
}
</entity_list>

# input
Article: {{ article }}
```

Table 13: Prompt for the first stage ZEST, to generate queries.

```
You will be given a news article about an event and potential Entities who are directly or
    indirectly involved in the event. Your task is to find supporting evidence for each of the
    specified entities in the given article.

An Entity is defined as an individual, group, collective, or organization involved in an event.
    This includes:
  * Organized armed groups with political purposes (e.g. "Hezbollah", "ISIS")
  * Organizations, governments, and political parties (e.g. "BJP: Bharatiya Janata Party", "
    Government of India", "Democratic Party of U.S.")
  * Ethnic, religious, social or occupational groups (e.g. "Jewish Group", "Muslim Group", "Women",
     "Students", "Farmers", "Journalists", "Teachers", "Lawyers")
  * General terms describing people involved (e.g., "Rioters", "Protestors", "Civilians", "Labor
    Group")

Follow these steps carefully:

1. First, you will be provided with the full text of the news article. Read the article carefully
    to understand the context of the event.

2. Next, you will be given a list of entities involved with the event.

3. Identify all supporting evidence of each given entity. Each evidence should be a short span from
     the article that has one of the following:
   - Contains the entity name, abbreviation or variations of its name
   - Implies the entity indirectly. For example "Madrasa" could be an evidence for "Muslim Group".
   - Mentions an affiliated group or organization of the entity.

4. If there are multiple evidence for the involvement of an entity, output one of them. If no
    evidence is found for an entity, respond with a mostly empty `EntitySpan` and only fill the `
    explanation` field.

5. For each evidence you find for an entity, provide your answer in the provided JSON format.
    Include the original entity name in the `entity_name` field to denote which entities the
    evidence is for.

6. If unsure, err on the side of including the span as evidence.


# input
<article>
Country of event: {{ country }}
{{ article }}
</article>

<entities>
{% for e in entities %}
  - {{ e.name }}
  {{ e.description }}

{% endfor %}
</entities>
```

Table 14: Prompt for the second stage of Zest.

```
You will be given a news article, and structured information about a {event_type} event.
    A {event_type} {schema.event_name_to_description(event_type)}.
Given a list of Entities that are involved in the event, your task is to assign each entity to the
     correct field.

An Entity is defined as an individual, group, collective, or organization involved in an event.
     This includes:
* Organized armed groups with political purposes (e.g. "Hezbollah", "ISIS")
* Organizations, governments, and political parties (e.g. "BJP: Bharatiya Janata Party", "
     Government of India", "Democratic Party of U.S.")
* Ethnic, religious, social or occupational groups (e.g. "Jewish Group", "Muslim Group", "Women", "
     Students", "Farmers", "Journalists", "Teachers", "Lawyers")
* General terms describing people involved (e.g., "Rioters", "Protestors", "Civilians", "Labor
     Group")

Possible fields are:
{possible_fields_string}

To complete this task, follow these steps:

1. Analyze the news article and the {event_type} event carefully.

2. For each entity in the provided list, determine their appropriate field based on the information
      in the news article.

3. Assign each entity to the most appropriate field. Try to assign all entities, event if their
     involvement in the event is very indirect. For example, "Government of India" is still an
     actor if the Indian congress is involved in the event.

4. If a field doesn't have a corresponding entity, leave it as an empty list.

Output the assignment of entities to fields in the following JSON format. Note that you should
     always include the full name of the entities without change.

{
    "field_name 1" : ["entity 1", "entity 2", ...],
    "field_name 2" : ["entity 3", "entity 4", ...],
}

# input
<news_article>
    {{ article }}
</news_article>

<event>
    {{ event_with_empty_entities }}
</event>

Here is the list of entities and their definitions.
<entities>
    {% for e in linked_entities %}
        - {{ e.name }}: {{ e.description }}

    {% endfor %}
</entities>
```

Table 15: Prompt used in the third stage of ZEST for assigning entities to their correct event argument. A Pydantic schema is also passed to the model to follow.

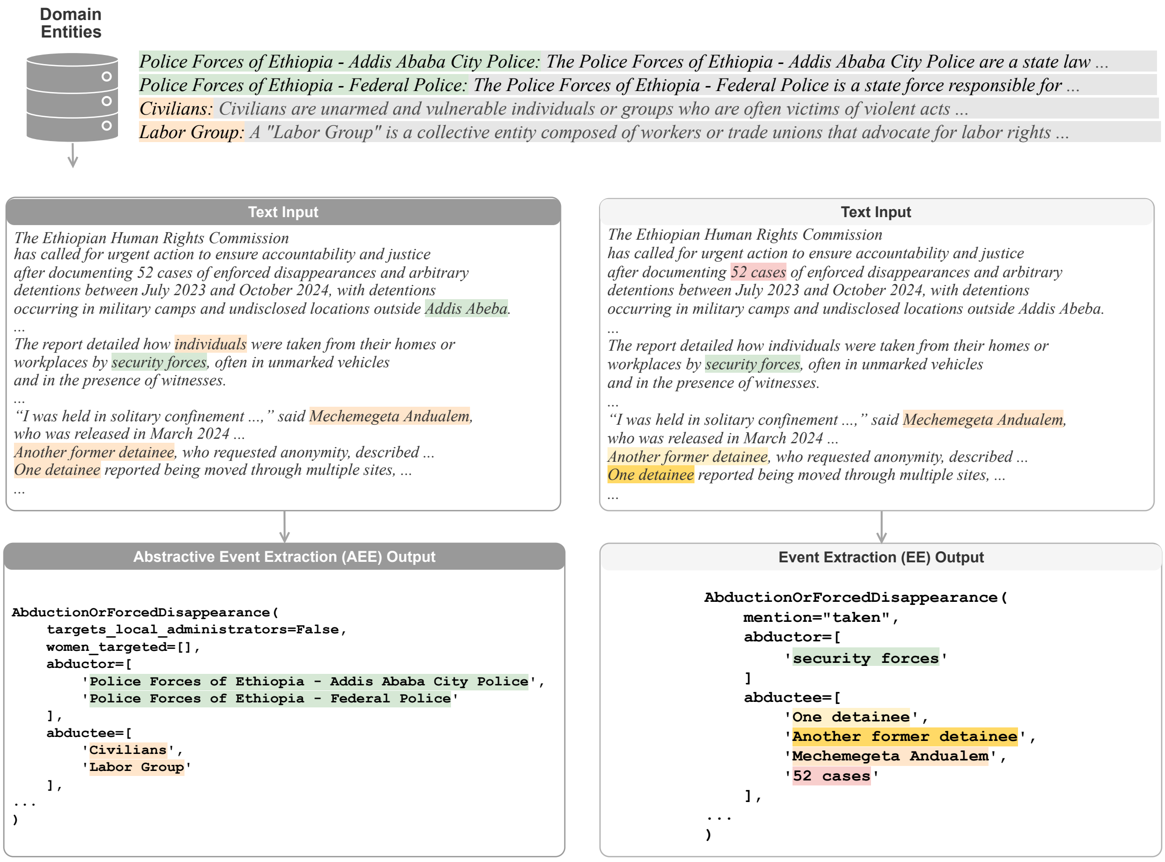



Figure 3: Another example from ACLED-DS with its abstractive event annotation. The input text and annotations are summarized for this figure. A hypothetical extractive annotation for the same event is also provided for comparison. The fact that the abductors are police forces of Ethiopia needs to be inferred from the event location (Adis Ababa) and the context.

textual analysis and prior knowledge. For each perspective, the LLM needs to select the most appropriate entity from the previously filtered set. In the contextual linking approach, the LLM is provided with both context and argument mention to enable context-aware prediction. Conversely, for the prior knowledge approach, the LLM receives only the argument mention, allowing it to draw upon its inherent understanding. The final merging stage involves using the LLM to select the final entity from the two candidates identified in the previous stage.

After EAE, we adopt a similar framework to OneNet for linking arguments to entities in the database. OneNet introduces an innovative approach using a fixed LLM to perform entity linking through few-shot inputs. The original framework comprises three distinct stages: entity reduction, dual-perspective entity linking, and merging linked entities. We closely follow this three-stage method with some nuanced modifications. During the entity reduction stage, we first generate summaries for each entity description.

## H Full Schema of ACLED-DS

The following is the full schema of ACLED-DS, after conversion to Python code, in Pydantic (Colvin et al., 2024) format. Abstract classes (denoted by `ABC` are only meant to group event types together and store common event arguments, are not counted as an event type, and are not used by ZEST. Docstrings are modified from the ACLED codebook (ACLED, 2023). `WomenTargetedCategory` and `Location` are two event types.

```
class Battle(ACLEDEvent, ABC):
    """
    A "Battle" event is defined as a violent interaction
      between two organized armed groups at a particular
      time and location. "Battle" can occur between armed
      and organized state, non-state, and external groups,
      and in any combination therein. There is no fatality
      minimum necessary for inclusion. Civilians can be
      harmed in the course of larger "Battle" events if they
       are caught in the crossfire, for example, or affected
       by strikes on military targets, which is commonly
      referred to as "collateral damage" (for more, see
      Indirect Killing of Civilians). When civilians are
      harmed in a "Battle" event, they are not recorded as
      an "Actor", nor is a separate civilian-specific event
```

```
You are an AI assistant tasked with extracting event arguments from a given news article. You will
    be provided with annotation guidelines for an event type and a news article to analyze.
Extract the arguments of the main event in the article, which is of type {event_type}.

{event_type}: {schema.event_name_to_description(event_type)}

When extracting event arguments, only pay attention to the main event in the article. Do not
    include any background information or other previous events that may be mentioned in the
    article.

# input
{{ article }}
```

Table 16: Prompt used for Abstract Code4Struct with GPT-4o and LLaMA-3.1-70B-instruct.

```
      recorded. If any civilian fatalities are reported as
      part of a battle, they are aggregated in the "
      Fatalities" field for the "Battle" event.
    The specific elements of the definition of a "Battle"
      event are as follows:
    Violent interaction: the exchange of armed force, or the
       use of armed force at close distance, between armed
      groups capable of inflicting harm upon the opposing
      side.
    Organized armed groups: collective actors assumed to be
      operating cohesively around an agenda, identity, or
      political purpose, using weapons to inflict harm.
      These groups frequently have a designated name and
      stated agenda.
    The "Battle" event type may include: ground clashes
      between different armed groups, ground clashes between
       armed groups supported by artillery fire or
      airstrikes, ambushes of on-duty soldiers or armed
      militants, exchanges of artillery fire, ground attacks
       against military or militant positions, air attacks
      where ground forces are able to effectively fire on
      the aircraft, and air-to-air combat.
    Cases where territory is regained or overtaken without
      resistance or armed interaction are not recorded as "
      Battle" events. Instead, they are recorded as "
      NonStateActorOvertakesTerritory" under the "
      StrategicDevelopment" event type
    "Battle" event type has the following subtypes:
    - GovernmentRegainsTerritory: Government forces or their
       affiliates regain control of a location from
      competing state forces or non-state groups through
      armed interaction.
    - NonStateActorOvertakesTerritory: A non-state actor or
      foreign state actor captures territory from an
      opposing government or non-state actor through armed
      interaction, establishing a monopoly of force within
      that territory.
    - ArmedClash: Armed, organized groups engage in a battle
       without significant changes in territorial control.

    """

    location: Location = Field(..., description="Location
      where the event takes place")
    fatalities: Optional[int] = Field(
        ...,
        description="Total number of fatalities, if known",
    )


class GovernmentRegainsTerritory(Battle):
    """
    Is a type of "Battle" event. This event type is used
      when government forces or their affiliates that are
      fighting against competing state forces or against a
      non-state group regain control of a location through
      armed interaction. This event type is only recorded
      for the re-establishment of government control and not
       for cases where competing non-state actors exchange
      control. Short-lived and/or small-scale territorial
      exchanges that do not last for more than one day are
      recorded as "ArmedClash".
    """
```

```
    # Possible "Interaction" codes include: 11, 12, 13, 14,
      and 18.

    government_force: List[str] = Field(
        ...,
        description="The government forces or their
      affiliates that regain control of the territory",
        is_entity_field=True,
    )
    adversary: List[str] = Field(
        ...,
        description="The competing state forces or non-state
       group that lose control of the territory. Can be
      State Forces, Rebel Groups, Political Militias,
      Identity Militias or External Forces",
        is_entity_field=True,
    )


class NonStateActorOvertakesTerritory(Battle):
    """
    Is a type of "Battle" event. This event type is used
      when a non-state actor (excluding those operating
      directly on behalf of the government) or a foreign
      state actor, through armed interaction, captures
      territory from an opposing government or non-state
      actor; as a result, they are regarded as having a
      monopoly of force within that territory. Short-lived
      and/or small-scale territorial exchanges that do not
      last for more than one day are recorded as "ArmedClash
      " events. In cases where non-state forces fight with
      opposing actors in a location many times before
      gaining control, only the final territorial
      acquisition is recorded as "Non-state actor overtakes
      territory". All other battles in that location are
      recorded as "ArmedClash".
    """

    # Possible "Interaction" codes include: 12, 13, 14, 18,
      22, 23, 24, 28, 33, 34, 38, 44, 48, and 88.

    non_state_actor: List[str] = Field(
        ...,
        description="The non-state actor overtaking
      territory. Can be Rebel Groups, Political Militias,
      Identity Militias or External Forces",
        is_entity_field=True,
    )
    adversary: List[str] = Field(
        ...,
        description="The opposing government or non-state
      actor from whom the territory was taken. Can be State
      Forces, Rebel Groups, Political Militias, Identity
      Militias or External Forces",
        is_entity_field=True,
    )


class ArmedClash(Battle):
    """
    Is a type of "Battle" event. This event type is used
      when two organized groups like State Forces, Rebel
      Groups, Political Militias, Identity Militias or
      External Forces engage in a battle, and no reports
```

```
      indicate a significant change in territorial control.
    `side_1` and `side_2` denote the two sides of the armed
      clash.
    Excludes demonstrations that turn violent, riots, and
      other forms of violence that are not organized armed
      clashes.
    """

    # Possible "Interaction" codes include: 11, 12, 13, 14,
      18, 22, 23, 24, 28, 33, 34, 38, 44, 48, and 88.

    side_1: List[str] = Field(
        ...,
        description="Groups involved in the clash. Can be
      State Forces, Rebel Groups, Political Militias,
      Identity Militias or External Forces",
        is_entity_field=True,
    )
    side_2: List[str] = Field(
        ...,
        description="Groups involved in the clash. Can be
      State Forces, Rebel Groups, Political Militias,
      Identity Militias or External Forces",
        is_entity_field=True,
    )
    targets_local_administrators: bool = Field(
        ...,
        description="Whether this violent event is affecting
       current local government officials and administrators
       - including governors, mayors, councilors, and other
      civil servants.",
    )
    women_targeted: List[WomenTargetedCategory] = Field(
        ...,
        description="The category of violence against women,
       if any. If this violence is not targeting women, this
       should be an empty list.",
    )



class Protest(ACLEDEvent, ABC):
    """
    A "Protest" event is defined as an in-person public
      demonstration of three or more participants in which
      the participants do not engage in violence, though
      violence may be used against them. Events include
      individuals and groups who peacefully demonstrate
      against a political entity, government institution,
      policy, group, tradition, business, or other private
      institution. The following are not recorded as "
      Protest" events: symbolic public acts such as displays
       of flags or public prayers (unless they are
      accompanied by a demonstration); legislative protests,
       such as parliamentary walkouts or members of
      parliaments staying silent; strikes (unless they are
      accompanied by a demonstration); and individual acts
      such as self-harm actions like individual immolations
      or hunger strikes.
    Protestor are noted by generic actor name "Protestor".
      If they are representing a group, the name of that
      group is also recorded in the field.
    "Protest" event type has the following subtypes:
    - ExcessiveForceAgainstProtestors: Peaceful protestor
      are targeted with lethal violence or violence
      resulting in serious injuries by state or non-state
      actors.
    - ProtestWithIntervention: A peaceful protest is
      physically dispersed or suppressed without serious
      injuries, or protestor interact with armed groups or
      rioters without serious harm, or protestors are
      arrested.
    - PeacefulProtest: Demonstrators gather for a protest
      without engaging in violence or rioting and are not
      met with force or intervention.

    """

    location: Location = Field(..., description="Location
      where the event takes place")
    crowd_size: Optional[str] = Field(
        ...,
        description="Estimated size of the crowd. It can be
      an exact number, a range, or a qualitative description
       like 'small'.",
    )
    protestors: List[str] = Field(
        ...,
        description="List of protestor groups or individuals
       involved in the protest",
        is_entity_field=True,
    )


class ExcessiveForceAgainstProtestors(Protest):
    """
    Is a type of "Protest" event (Protest events include
      individuals and groups who peacefully demonstrate
      against a political entity, government institution,
      policy, group, tradition, business, or other private
      institution.) This event type is used when individuals
       are engaged in a peaceful protest and are targeted
      with lethal violence or violence resulting in serious
      injuries (e.g. requiring hospitalization). This
      includes situations where remote explosives, such as
      improvised explosive devices, are used to target
      protestors, as well as situations where non-state
      actors, such as rebel groups, target protestors.
    """

    # Possible "Interaction" codes include: 16, 26, 36, 46,
      56, and 68.

    perpetrators: List[str] = Field(
        ...,
        description="Entities perpetrating the violence. Can
       be State Forces, Rebel Groups, Political Militias,
      Identity Militias, External Forces",
        is_entity_field=True,
    )
    targets_civilians: bool = Field(
        ...,
        description="Indicates if the '
      ExcessiveForceAgainstProtestors' event is mainly or
      only targeting civilians. E.g. state forces using
      lethal force to disperse peaceful protestors.",
    )

    fatalities: Optional[int] = Field(
        ...,
        description="Total number of fatalities, if known",
    )


class ProtestWithIntervention(Protest):
    """
    Is a type of "Protest" event. This event type is used
      when individuals are engaged in a peaceful protest
      during which there is a physically violent attempt to
      disperse or suppress the protest, which resulted in
      arrests, or minor injuries . If there is intervention,
       but not violent, the event is recorded as "
      PeacefulProtest" event type.
    """

    perpetrators: List[str] = Field(
        ...,
        description="Group(s) or entities attempting to
      disperse or suppress the protest",
        is_entity_field=True,
    )
    fatalities: Optional[int] = Field(
        ...,
        description="Total number of fatalities, if known",
    )


class PeacefulProtest(Protest):
    """
    Is a type of "Protest" event (Protest events include
      individuals and groups who peacefully demonstrate
      against a political entity, government institution,
      policy, group, tradition, business, or other private
      institution.) This event type is used when
      demonstrators gather for a protest and do not engage
      in violence or other forms of rioting activity, such
      as property destruction, and are not met with any sort
       of violent intervention.
    """

    # Possible "Interaction" codes include: 60, 66, and 67.

    counter_protestors: List[str] = Field(
        ...,
```

```python
        description="Groups or entities engaged in counter
    protest, if any",
        is_entity_field=True,
    )


class Riot(ACLEDEvent, ABC):
    """
    "Riot" are violent events where demonstrators or mobs of
     three or more engage in violent or destructive acts,
    including but not limited to physical fights, rock
    throwing, property destruction, etc. They may engage
    individuals, property, businesses, other rioting
    groups, or armed actors. Rioters are noted by generic
    actor name "Rioters". If rioters are affiliated with a
     specific group - which may or may not be armed - or
    identity group, that group is recorded in the
    respective "Actor" field. Riots may begin as peaceful
    protests, or a mob may have the intention to engage in
     violence from the outset.
    "Riot" event type has the following subtypes:
    - ViolentDemonstration: Demonstrators engage in violence
     or destructive activities, such as physical clashes,
    vandalism, or road-blocking, regardless of who
    initiated the violence.
    - MobViolence: Rioters violently interact with other
    rioters, civilians, property, or armed groups outside
    of demonstration contexts, often involving disorderly
    crowds with the intention to cause harm or disruption.

    """

    location: Location = Field(..., description="Location
    where the event takes place")
    crowd_size: Optional[str] = Field(
        ...,
        description="Estimated size of the crowd. It can be
    an exact number, a range, or a qualitative description
     like 'small'.",
    )
    fatalities: Optional[int] = Field(
        ...,
        description="Total number of fatalities, if known",
    )
    targets_civilians: bool = Field(
        ...,
        description="Indicates if the 'Riot' event is mainly
     or only targeting civilians. E.g. a village mob
    assaulting another villager over a land dispute.",
    )
    group_1: List[str] = Field(
        ...,
        description="Group or individual involved in the
    violence",
        is_entity_field=True,
    )
    group_2: List[str] = Field(
        ...,
        description="The other group or individual involved
    in the violence, if any",
        is_entity_field=True,
    )
    targets_local_administrators: bool = Field(
        ...,
        description="Whether this violent event is affecting
     current local government officials and administrators
     - including governors, mayors, councilors, and other
    civil servants.",
    )
    women_targeted: List[WomenTargetedCategory] = Field(
        ...,
        description="The category of violence against women,
     if any. If this violence is not targeting women, this
     should be an empty list.",
    )


class ViolentDemonstration(Riot):
    """
    Is a type of "Riot" event. This event type is used when
    demonstrators engage in violence and/or destructive
    activity. Examples include physical clashes with other
     demonstrators or government forces; vandalism; and
    road-blocking using barricades, burning tires, or
    other material. The coding of an event as a "Violent
    demonstration" does not necessarily indicate that
    demonstrators initiated the violence and/or
    destructive actions.
    Excludes events where a weapon is drawn but not used, or
     when the situation is de-escalated before violence
    occurs.
    """


class MobViolence(Riot):
    """
    Is a type of "Riot" event. A mob is considered a crowd
    of people that is disorderly and has the intention to
    cause harm or disruption through violence or property
    destruction. Note that this type of violence can also
    include spontaneous vigilante mobs clashing with other
     armed groups or attacking civilians. While a "Mob
    violence" event often involves unarmed or crudely
    armed rioters, on rare occasions, it can involve
    violence by people associated with organized groups
    and/or using more sophisticated weapons, such as
    firearms.
    """


class ExplosionOrRemoteViolence(ACLEDEvent, ABC):
    """
    "ExplosionOrRemoteViolence" is defined as events as
    incidents in which one side uses weapon types that, by
     their nature, are at range and widely destructive.
    The weapons used in "ExplosionOrRemoteViolence" events
     are explosive devices, including but not limited to:
    bombs, grenades, improvised explosive devices (IEDs),
    artillery fire or shelling, missile attacks, air or
    drone strikes, and other widely destructive heavy
    weapons or chemical weapons. Suicide attacks using
    explosives also fall under this category. When an "
    ExplosionOrRemoteViolence" event is reported in the
    context of an ongoing battle, it is merged and
    recorded as a single "Battles" event. "
    ExplosionOrRemoteViolence" can be used against armed
    agents as well as civilians.
    "ExplosionOrRemoteViolence" event type has the following
     subtypes:
    - ChemicalWeapon: The use of chemical weapons in warfare
     without any other engagement.
    - AirOrDroneStrike: Air or drone strikes occurring
    without any other engagement, including attacks by
    helicopters.
    - SuicideBomb: A suicide bombing or suicide vehicle-
    borne improvised explosive device (SVBIED) attack
    without an armed clash.
    - ShellingOrArtilleryOrMissileAttack: The use of long-
    range artillery, missile systems, or other heavy
    weapons platforms without any other engagement.
    - RemoteExplosiveOrLandmineOrIED: Detonation of remotely-
     or victim-activated devices, including landmines and
    IEDs, without any other engagement.
    - Grenade: The use of a grenade or similar hand-thrown
    explosive without any other engagement.
    """

    location: Location = Field(..., description="Location
    where the event takes place")
    targets_civilians: bool = Field(
        ...,
        description="Indicates if the '
    ExplosionOrRemoteViolence' event is mainly or only
    targeting civilians. E.g. a landmine killing a farmer.
    ",
    )
    fatalities: Optional[int] = Field(
        ...,
        description="Total number of fatalities, if known",
    )
    attackers: List[str] = Field(
        ...,
        description="Entities conducting the violence",
        is_entity_field=True,
    )
    targeted_entities: List[str] = Field(
        ...,
        description="Entities or actors being targeted",
        is_entity_field=True,
    )
```

```python
    targets_local_administrators: bool = Field(
        ...,
        description="Whether this violent event is affecting
       current local government officials and administrators
       - including governors, mayors, councilors, and other
      civil servants.",
    )
    women_targeted: List[WomenTargetedCategory] = Field(
        ...,
        description="The category of violence against women,
       if any. If this violence is not targeting women, this
       should be an empty list.",
    )


class ChemicalWeapon(ExplosionOrRemoteViolence):
    """
    Is a type of "ExplosionOrRemoteViolence" event. This
      event type captures the use of chemical weapons in
      warfare in the absence of any other engagement. ACLED
      considers chemical weapons as all substances listed as
       Schedule 1 of the Chemical Weapons Convention,
      including sarin gas, mustard gas, chlorine gas, and
      anthrax. Napalm and white phosphorus, as well as less-
      lethal crowd control substances - such as tear gas -
      are not considered chemical weapons within this event
      type.
    """


class AirOrDroneStrike(ExplosionOrRemoteViolence):
    """
    Is a type of "ExplosionOrRemoteViolence" event. This
      event type is used when air or drone strikes take
      place in the absence of any other engagement. Please
      note that any air-to-ground attacks fall under this
      event type, including attacks by helicopters that do
      not involve exchanges of fire with forces on the
      ground.
    """


class SuicideBomb(ExplosionOrRemoteViolence):
    """
    Is a type of "ExplosionOrRemoteViolence" event. This
      event type is used when a suicide bombing occurs in
      the absence of an armed clash, such as an exchange of
      small arms fire with other armed groups. It also
      includes suicide vehicle-borne improvised explosive
      device (SVBIED) attacks. Note that the suicide bomber
      is included in the total number of reported fatalities
       coded for such events.
    """


class ShellingOrArtilleryOrMissileAttack(
      ExplosionOrRemoteViolence):
    """
    Is a type of "ExplosionOrRemoteViolence" event. This
      event type captures the use of long-range artillery,
      missile systems, or other heavy weapons platforms in
      the absence of any other engagement. When two armed
      groups exchange long-range fire, it is recorded as an
      "ArmedClash". "ShellingOrArtilleryOrMissileAttack"
      events include attacks described as shelling, the use
      of artillery and cannons, mortars, guided missiles,
      rockets, grenade launchers, and other heavy weapons
      platforms. Crewed aircraft shot down by long-range
      systems fall under this event type.  Uncrewed armed
      drones that are shot down, however, are recorded as
      interceptions under "DisruptedWeaponsUse" because
      people are not targeted (see below). Similarly, an
      interception of a missile strike itself (such as by
      the Iron Dome in Israel) is also recorded as "
      DisruptedWeaponsUse".
    """


class RemoteExplosiveOrLandmineOrIED(
      ExplosionOrRemoteViolence):
    """
    Is a type of "ExplosionOrRemoteViolence" event. This
      event type is used when remotely- or victim-activated
      devices are detonated in the absence of any other
      engagement. Examples include landmines, IEDs - whether
       alone or attached to a vehicle, or any other sort of
      remotely detonated or triggered explosive. Unexploded
      ordnances (UXO) also fall under this category.
    SVBIEDs are recorded as "Suicide bomb" events, while the
       safe defusal of an explosive or its accidental
      detonation by the actor who planted it (with no other
      casualties reported) is recorded under "
      DisruptedWeaponsUse".
    """


class Grenade(ExplosionOrRemoteViolence):
    """
    Is a type of "ExplosionOrRemoteViolence" event. This
      event type captures the use of a grenade or any other
      similarly hand-thrown explosive, such as an IED that
      is thrown, in the absence of any other engagement.
      Events involving so-called "crude bombs" (such as
      Molotov cocktails, firecrackers, cherry bombs, petrol
      bombs, etc.) as well as "stun grenades" are not
      recorded in this category, but are included under
      either "Riot" or "StrategicDevelopment" depending on
      the context in which they occurred.
    """



class ViolenceAgainstCivilians(ACLEDEvent, ABC):
    """
    ACLED defines "ViolenceAgainstCivilians" as violent
      events where an organized armed group inflicts
      violence upon unarmed non-combatants. By definition,
      civilians are unarmed and cannot engage in political
      violence. Therefore, the violence is understood to be
      asymmetric as the perpetrator is assumed to be the
      only actor capable of using violence in the event. The
       perpetrators of such acts include state forces and
      their affiliates, rebels, militias, and external/other
       forces.
    In cases where the identity and actions of the targets
      are in question (e.g. the target may be employed as a
      police officer), ACLED determines that if a person is
      harmed or killed while unarmed and unable to either
      act defensively or counter-attack, this is an act of "
      ViolenceAgainstCivilians". This includes extrajudicial
       killings of detained combatants or unarmed prisoners
      of war.
    "ViolenceAgainstCivilians" also includes attempts at
      inflicting harm (e.g. beating, shooting, torture, rape,
       mutilation, etc.) or forcibly disappearing (e.g.
      kidnapping and disappearances) civilian actors. Note
      that the "ViolenceAgainstCivilians" event type
      exclusively captures violence targeting civilians that
       does not occur concurrently with other forms of
      violence - such as rioting - that are coded higher in
      the ACLED event type hierarchy. To get a full list of
      events in the ACLED dataset where civilians were the
      main or only target of violence, users can filter on
      the "Civilian targeting" field.
    "ViolenceAgainstCivilians" event type has the following
      subtypes:
    - SexualViolence: Any event where an individual is
      targeted with sexual violence, including but not
      limited to rape, public stripping, and sexual torture,
       with the gender identities of victims recorded when
      reported.
    - Attack: An event where civilians are targeted with
      violence by an organized armed actor outside the
      context of other forms of violence, including severe
      government overreach by law enforcement.
    - AbductionOrForcedDisappearance: An event involving the
       abduction or forced disappearance of civilians
      without reports of further violence, including arrests
       by non-state groups and extrajudicial detentions by
      state forces, but excluding standard judicial arrests
      by state forces.
    """

    location: Location = Field(..., description="Location
      where the event takes place")
    targets_local_administrators: bool = Field(
        ...,
        description="Whether this violent event is affecting
       current local government officials and administrators
       - including governors, mayors, councilors, and other
      civil servants.",
    )
    women_targeted: List[WomenTargetedCategory] = Field(
        ...,
```

```python
        description="The category of violence against women,
      if any. If this violence is not targeting women, this
      should be an empty list.",
    )


class SexualViolence(ViolenceAgainstCivilians):
    """
    Is a type of "ViolenceAgainstCivilians" event. This
      event type is used when any individual is targeted
      with sexual violence. SexualViolence is defined
      largely as an action that inflicts harm of a sexual
      nature. This means that it is not limited to solely
      penetrative rape, but also includes actions like
      public stripping, sexual torture, etc. Given the
      gendered nature of sexual violence, the gender
      identities of the victims - i.e. "Women", "Men", and "
      LGBTQ+", or a combination thereof - are recorded in
      the "Associated Actor" field for these events when
      reported. Note that it is possible for sexual violence
       to occur within other event types such as "Battle"
      and "Riot".
    """

    fatalities: Optional[int] = Field(
        ...,
        description="Total number of fatalities, if known",
    )  # Is very very rare, only 7 events in English for
      2024
    perpetrators: List[str] = Field(
        ...,
        description="The attacker(s) entity or actor",
        is_entity_field=True,
    )
    victims: List[str] = Field(
        ...,
        description="The entity or actor(s) that is the
      target or victim of the SexualViolence event",
        is_entity_field=True,
    )


class Attack(ViolenceAgainstCivilians):
    """
    Is a type of "ViolenceAgainstCivilians" event. This
      event type is used when civilians are targeted with
      violence by an organized armed actor outside the
      context of other forms of violence like ArmedClash,
      Protests, Riots, or ExplosionOrRemoteViolence.
      Violence by law enforcement that constitutes severe
      government overreach is also recorded as an "Attack"
      event.
    Attacks of a sexual nature are recorded as
      SexualViolence.
    If only property is attacked and not people, the event
      should be recorded as LootingOrPropertyDestruction
      event type.
    Excludes discovery of mass graves, which are recorded as
       "OtherStrategicDevelopment" events.
    """

    fatalities: Optional[int] = Field(
        ...,
        description="Total number of fatalities, if known",
    )
    attackers: List[str] = Field(
        ...,
        description="The attacker entity or actor(s)",
        is_entity_field=True,
    )
    targeted_entities: List[str] = Field(
        ...,
        description="The entity or actor(s) that is the
      target of the attack",
        is_entity_field=True,
    )


class AbductionOrForcedDisappearance(
      ViolenceAgainstCivilians):
    """
    Is a type of "ViolenceAgainstCivilians" event. This
      event type is used when an actor engages in the
      abduction or forced disappearance of civilians,
      without reports of further violence. If fatalities or
      serious injuries are reported during the abduction or
      forced disappearance, the event is recorded as an "
      Attack" event instead. If such violence is reported in
       later periods during captivity, this is recorded as
      an additional "Attack" event. Note that multiple
      people can be abducted in a single "Abduction/forced
      disappearance" event.
    Arrests by non-state groups and extrajudicial detentions
       by state forces are considered "Abduction/forced
      disappearance". Arrests conducted by state forces
      within the standard judicial process are, however,
      considered "Arrest".
    """

    abductor: List[str] = Field(
        ...,
        description="The abductor person or group(s)",
        is_entity_field=True,
    )
    abductee: List[str] = Field(
        ...,
        description="People or group(s) that were abducted
      or disappeared. Note that multiple people can be
      abducted in a single AbductionOrForcedDisappearance
      event",
        is_entity_field=True,
    )



class StrategicDevelopment(ACLEDEvent, ABC):
    """
    This event type captures contextually important
      information regarding incidents and activities of
      groups that are not recorded as "Political violence"
      or "Demonstration" events, yet may trigger future
      events or contribute to political dynamics within and
      across states. The inclusion of such events is limited,
       as their purpose is to capture pivotal events within
      the broader political landscape. They typically
      include a disparate range of events, such as
      recruitment drives, looting, and incursions, as well
      as the location and date of peace talks and the
      arrests of high-ranking officials or large groups.
      While it is rare for fatalities to be reported as a
      result of such events, they can occur in certain cases
       - e.g. the suspicious death of a high-ranking
      official, the accidental detonation of a bomb
      resulting in the bomber being killed, etc.
    Due to their context-specific nature, "
      StrategicDevelopment" are not collected and recorded
      in the same cross-comparable fashion as "Political
      violence" and "Demonstration" events. As such, the "
      StrategicDevelopment" event type is primarily a tool
      for understanding particular contexts.
    "StrategicDevelopment" event type has the following
      subtypes:
    - Agreement: Records any agreement between different
      actors, such as peace talks, ceasefires, or prisoner
      exchanges.
    - Arrest: Used when state forces or controlling actors
      detain a significant individual or conduct politically
       important mass arrests.
    - ChangeToArmedGroup: Records significant changes in the
       activity or structure of armed groups, including
      creation, recruitment, movement, or absorption of
      forces.
    - DisruptedWeaponsUse: Captures instances where an
      explosion or remote violence event is prevented, or
      when significant weapons caches are seized.
    - BaseEstablished: Used when an organized armed group
      establishes a permanent or semi-permanent base or
      headquarters.
    - LootingOrPropertyDestruction: Records incidents of
      looting or seizing goods/property outside the context
      of other forms of violence or destruction.
    - NonViolentTransferOfTerritory: Used when actors
      acquire control of a location without engaging in
      violent interaction with another group.
    - OtherStrategicDevelopment: Covers significant
      developments that don't fall into other Strategic
      Development event types, such as coups or population
      displacements.
    """

    location: Location = Field(..., description="Location
      where the event takes place")
```

```python
class Agreement(StrategicDevelopment):
    """
    Is a type of "StrategicDevelopment" event. This event
      type is used to record any sort of agreement between
      different armed actors (such as governments and rebel
      groups). Examples include peace agreements/talks,
      ceasefires, evacuation deals, prisoner exchanges,
      negotiated territorial transfers, prisoner releases,
      surrenders, repatriations, etc.
    Excludes agreements between political parties, trade
      unions, or other non-armed actors like protestors.
    """

    group_1: List[str] = Field(
        ...,
        description="Group or individual involved in the
      agreement",
        is_entity_field=True,
    )
    group_2: List[str] = Field(
        ...,
        description="The other group or individual involved
      in the agreement",
        is_entity_field=True,
    )


class Arrest(StrategicDevelopment):
    """
    Is a type of "StrategicDevelopment" event. This event
      type is used when state forces or other actors
      exercising de facto control over a territory either
      detain a particularly significant individual or engage
       in politically significant mass arrests. This
      excludes arrests of individuals for common crimes,
      such as theft or assault, unless the individual is a
      high-ranking official or the arrest is politically
      significant.
    """

    detainers: List[str] = Field(
        ...,
        description="The person or group(s) who detains or
      jails the detainee(s)",
        is_entity_field=True,
    )
    detainees: List[str] = Field(
        ...,
        description="The person or group(s) being detained
      or jailed",
        is_entity_field=True,
    )


class ChangeToArmedGroup(StrategicDevelopment):
    """
    Is a type of "StrategicDevelopment" event. This event
      type is used to record significant changes in the
      activity or structure of armed groups. It can cover
      anything from the creation of a new rebel group or a
      paramilitary wing of the security forces, "voluntary"
      recruitment drives, movement of forces, or any other
      non-violent security measures enacted by armed actors.
       This event type can also be used if one armed group
      is absorbed into a different armed group or to track
      large-scale defections.
    """

    armed_group: List[str] = Field(
        ...,
        description="The name of armed group that underwent
      change",
        is_entity_field=True,
    )
    other_actors: List[str] = Field(
        ...,
        description="Other actors or groups involved. E.g.
      the government that ordered a change to its army.",
        is_entity_field=True,
    )


class DisruptedWeaponsUse(StrategicDevelopment):
    """
    Is a type of "StrategicDevelopment" event. This event
      type is used to capture all instances in which an
      event of "ExplosionOrRemoteViolence" is prevented from
       occurring, or when armed actors seize significant
      caches of weapons. It includes the safe defusal of an
      explosive, the accidental detonation of explosives by
      those allegedly responsible for planting it, the
      interception of explosives in the air, as well as the
      seizure of weapons or weapons platforms such as jets,
      helicopters, tanks, etc. Note that in cases where a
      group other than the one that planted an explosive is
      attempting to render an explosive harmless and it goes
       off, this is recorded under the "
      ExplosionOrRemoteViolence" event type, as the
      explosive has harmed an actor other than the one that
      planted it.
    """

    attackers: List[str] = Field(
        ...,
        description="The entity or actor(s) responsible for
      the remote violence",
        is_entity_field=True,
    )
    disruptors: List[str] = Field(
        ...,
        description="The entity or actor(s) disrupting the
      explosion or remote violence",
        is_entity_field=True,
    )
    targets_local_administrators: bool = Field(
        ...,
        description="Whether this violent event is affecting
       current local government officials and administrators
       - including governors, mayors, councilors, and other
      civil servants.",
    )
    women_targeted: List[WomenTargetedCategory] = Field(
        ...,
        description="The category of violence against women,
       if any. If this violence is not targeting women, this
       should be an empty list.",
    )


class BaseEstablished(StrategicDevelopment):
    """
    Is a type of "StrategicDevelopment" event. This event
      type is used when an organized armed group establishes
       a permanent or semi-permanent base or headquarters.
      There are few cases where opposition groups other than
       rebels can also establish a headquarters or base (e.g.
       AMISOM forces in Somalia).
    """

    group: List[str] = Field(
        ...,
        description="Entity or group(s) establishing the
      base",
        is_entity_field=True,
    )


class LootingOrPropertyDestruction(StrategicDevelopment):
    """
    Is a type of "StrategicDevelopment" event. This event
      type is used when actors engage in looting or seizing
      goods or property outside the context of other forms
      of violence or destruction, such as rioting or armed
      clashes. This excludes the seizure or destruction of
      weapons or weapons systems, which are captured under
      the "DisruptedWeaponsUse" event type. This can occur
      during raiding or after the capture of villages or
      other populated places by armed groups that occur
      without reported violence.
    """

    perpetrators: List[str] = Field(
        ...,
        description="The group or entity that does the
      looting or seizure",
        is_entity_field=True,
    )
    victims: List[str] = Field(
        ...,
        description="The group or entity that was the target
       of looting or seizure",
        is_entity_field=True,
    )
    targets_local_administrators: bool = Field(
```

```
        ...,
        description="Whether this violent event is affecting
      current local government officials and administrators
      - including governors, mayors, councilors, and other
     civil servants.",
    )
    women_targeted: List[WomenTargetedCategory] = Field(
        ...,
        description="The category of violence against women,
      if any. If this violence is not targeting women, this
      should be an empty list.",
    )


class NonViolentTransferOfTerritory(StrategicDevelopment):
    """
    Is a type of "StrategicDevelopment" event. This event
      type is used in situations in which rebels,
      governments, or their affiliates acquire control of a
      location without engaging in a violent interaction
      with another group. Rebels establishing control of a
      location without any resistance is an example of this
      event.
    """

    actors_taking_over: List[str] = Field(
        ...,
        description="The entity or actor(s) establishing
      control.",
        is_entity_field=True,
    )
    actors_giving_up: List[str] = Field(
        ...,
        description="The entity or actor(s) giving up
      territory, if known.",
        is_entity_field=True,
    )


class OtherStrategicDevelopment(StrategicDevelopment):
    """
    Is a type of "StrategicDevelopment" event. This event
      type is used to cover any significant development that
       does not fall into any of the other "
      StrategicDevelopment" event types. Includes the
      occurrence of a coup, the displacement of a civilian
      population as a result of fighting, and the discovery
      of mass graves.
    """

    group_1: List[str] = Field(
        ...,
        description="Group or individual involved in the
      StrategicDevelopment",
        is_entity_field=True,
    )
    group_2: List[str] = Field(
        ...,
        description="The other group or individual involved
      in the violence, if any",
        is_entity_field=True,
    )




class WomenTargetedCategory(str, Enum):
    CANDIDATES_FOR_OFFICE = "Women who are running in an
      election to hold a publicly elected government
      position"
    POLITICIANS = "Women who currently serve in an elected
      position in government"
    POLITICAL_PARTY_SUPPORTERS = "political party supporters
      "
    VOTERS = "Women who are registering to vote or are
      casting a ballot in an election"
    GOVERNMENT_OFFICIALS = "Women who work for the local,
      regional, or national government in a non-partisan
      capacity"
    ACTIVISTS_HRD_SOCIAL_LEADERS = (
        "Women who are activists/human rights defenders/
      social leaders"
    )
    RELATIVES_OF_TARGETED_GROUPS = "Women who are subject to
       violence as a result of who they are married to, the
      daughter of, related to, or are otherwise personally
      connected to (e.g. candidates, politicians, social
      leaders, armed actors, voters, party supporters, etc.)
      "
    ACCUSED_OF_WITCHCRAFT = "Women accused of witchcraft or
      sorcery, or other mystical or spiritual practices that
       are typically considered taboo or dangerous within
      some societies (excluding women who serve as religious
       leaders in religious structures that are typically
      not viewed as taboo or dangerous, such as nuns, female
       priests, or shamans)"
    GIRLS = "Girls who are under the age of 18; they may be
      specifically referred to by age or explicitly referred
       to as a child/girl"




class Location(BaseModel):
    """
    The most specific location for an event. Locations can
      be named populated places, geostrategic locations,
      natural locations, or neighborhoods of larger cities.
    In selected large cities with activity dispersed over
      many neighborhoods, locations are further specified to
       predefined subsections within a city. In such cases,
      City Name - District name (e.g. Mosul - Old City) is
      recorded in "specific_location". If information about
      the specific neighborhood/district is not known, the
      location is recorded at the city level (e.g. Mosul).
    """

    country: str = Field(
        ...,
        description="Name of the country in English. Example:
       United States",
    )
    address: str = Field(
        ...,
        description="Comma-separated address in order from
      neighborhood level to village/city, district, county,
      province, region, and country, if available. Excludes
      street names, buildings, and other specific landmarks.
       Example: Mosul, Old City, Nineveh, Nineveh, Iraq",
    )
```