MERCURY CHALLENGE HANDBOOK APPENDICES

MERCURY
CHALLENGE

Last update: 06/25/2018

TABLE OF CONTENTS

OVERVIEW

This collection of appendices is aimed at providing participants of the IARPA Mercury Challenge deeper context and understanding of the various procedures and processes used to design the challenge. If you have any questions about the content here, please send a message to mercury-challenge@iarpa.gov for assistance.

WARNING AND UPDATE SEQUENCES

UPDATE SEQUENCES

As more information is available, the Participant may choose to update a previously submitted warning. The set of updates to a single warning is called an update sequence. The individual elements of the update sequence are scored in the same way as for a single warning. The score for the update sequence is the average of the individual scores for each update sequence member. Note that only those warnings submitted at least one day in advance of the reported date are considered; others are dropped. Dropped warnings in the update sequence are not used to calculate the update sequence score.

Updates are limited to one every four hours for each individual warning.

PERMISSIBLE WARNING UPDATES

If an update to a predefined field is not explicitly permitted, then it is not allowed. Any update to non-predefined fields is allowed. For example, an update to participant-designed fields for debugging, tracking, or other purposes would be allowed. The following table lists the permissible warning updates.

TABLE 1. PERMISSIBLE WARNING UPDATES

Field/Class	CU	MA	Disease
Event_Date	No	Yes	No
State	No	Yes	No
City	NA	Yes	NA
Latitude	NA	Yes	NA
Longitude	NA	Yes	NA
Actor	NA	Yes	NA
Case_Count	Yes	NA	Yes
Subtype	NA	Yes	NA

BASE RATE MODEL

For the purposes of the Mercury Challenge, base rate models are models that only use information included in the history of observed events. These models are used for comparison purposes with participant results. It is expected that participant models will score better than the base rate models.

In this appendix we will describe the several base rate models to be used by the Government team in Mercury Challenge Test and Evaluation.

HISTORY SAMPLE MODEL

The history sample base rate model is used for the Military Activity event type. This model examines the most recent 12 months of event history for the country of interest. The event rate over the history determines how many events to generate for the month of interest. These events are distributed by sampling from the days of the forecast period using a random uniform distribution with replacement.   For example, if the 12-month historical event rate is 2 events per day and the month to be forecast has 30 days then the model will take 60 samples with replacement from the days in that month. The model independently samples the attributes of the event history with replacement to synthesize the warning details.

TIME SERIES MODELS

Time series models predict future values of case counts using the history of case counts. These models are used for Non-Violent Civil Unrest and Disease event types.

There are several types of time series models we will use in the Mercury Challenge:

AUTOREGRESSIVE MOVING AVERAGE MODEL AND EXPONENTIALLY WEIGHTED MOVING AVERAGE MODEL

The autoregressive moving average, or ARMA model combines the properties of autoregressive models and moving-average models. In a purely autoregressive model the future values of the series are determined by the tendency of the series to revert to the mean following one or more shocks. In a pure moving average model the future values of the series are predicted to be the weighted average of preceding values. The ARMA model combines aspects of these two models. The parameters for the ARMA model for each time series will be determined by training on the time period leading up to the start of the Challenge.

An exponentially weighted moving average (EWMA) model is a moving average model where the weight of the value from n days ago is $p^{(n-1)}$ as much as the weight of the value from 1 day ago, where p is strictly between 0 and 1. If an ARMA model cannot be trained for a time series an EWMA model will be used instead.

Either the ARMA or EWMA model will be used to generate baserate warnings for the Non-Violent Civil Unrest event types.

SEASONAL MODEL, ADJUSTED BY RECENT TREND

MERS case counts tend to follow a seasonal pattern. The Baserate model for this time series will use the average case count for the epidemiological week multiplied by the EWMA of the most recent weeks with a half-life of 1 week. This model will be used to generate warnings for the Disease event type.

GSR REFERENCE MATERIALS

MA ACTOR TABLES

MA events contain an Actor field; this is a restricted field in the GSR. Because the situation for MA events in the Mercury Challenge countries is fluid, the list of permissible Actor names may change over time.

The Actor and table is modified on an as-needed basis to add Actors that are approved by the Mercury Challenge PM. Any updates to the Actor table will be released monthly along with any updates the GSR and challenge participants will be notified by a challenge update or forum post. Furthermore, in order to allow time for challenge participants to train new actors will not be used until the 1st of the second month following the GSR release.

The Actor table as of the start of the Challenge is provided below:

TABLE 2. STATE ACTORS

Actor	Arabic	isCurrent
Abu Al-Fadl Al-Abbas Forces		
Abu Al-Fadhel Al-Abbas Brigade / Liwa Abu Al-Fadl Al-Abbas / Al-Abbas Brigade / Abu Al-Fadl Al-Abbas Forces	كتيبة لواء أبو الفضل العباس	no
Asa'Ib Ahl Al-Haq / AAH / Khazali Network	عصائب أهل الحق	
Asayish / Al-Asayish / Kurdish Security Forces	الأسايش	
Badr Organization	منظمة بدر	
Baghdad International Airport Security	أمن مطار بغداد الدولي	
Egyptian Military	القوات المسلحة المصرية	
Egyptian Police	الشرطة المصرية	
Hezbollah	حزب الله	

Iranian Military	القوات المسلحة الإيرانية	
Iranian Quds Force	فيلق القدس	
Iraqi Intelligence Service	المخابرات العراقية	
Iraqi Military	القوات المسلحة العراقية	
Iraqi Police	الشرطة العراقية	
Iraqi Security Forces	قوات الأمن العراقية	
Iraqi Special Forces	قوات العمليات الخاصة العراقية	
Jordanian Police	الأمن العام الأردني	
Jund Al-Imam	جند الإمام	
Kata'ib Al-Imam Ali	كتائب الإمام علي	
Kata'ib Hezbollah / KH / Hezbollah Battalions	كتائب حزب الله	
Lebanese General Security	الأمن العام اللبناني	
Lebanese Military	الجيش اللبناني	
National Defence Forces Syria / NDF / Quwat Ad-Difa Al-Watani / Popular Committees / Lijan Militias / Al-Lijan Al-Sha'Biyah / Peoples Committees	اللجان الشعبية / قوات الدفاع الوطني	
Palestine Liberation Army	جيش التحرير الفلسطيني	
People's Mobilization / National Mobilization / Popular Mobilization Forces / Units / Committee / Hashd Al-Sha'abi / Hashid Shaabi	الحشد الشعبي	
Peshmerga	قوات البشمركة الكردية	
Qatari Military	القوات المسلحة القطرية	
Royal Bahraini Military	الجيش البحريني الملكي	
Royal Bahraini Police / Public Security Forces / PSF	قوات الأمن العام	
Royal Jordanian Military	القوات البرية الملكية الأردنية	
Royal Saudi Military	القوات المسلحة السعودية	
Royal Saudi Police	الشرطة الملكية السعودية	
Russian Military	القوات المسلحة للاتحاد الروسي	
Syrian Arab Military	الجيش العربي السوري	
Turkish Military	القوات المسلحة التركية	
Unspecified	غير محدد	
Yemeni Military	القوات المسلحة اليمنية	

PERFORMANCE MEASURES

CHALLENGE PERFORMANCE MEASURES

The set of overall performance measures follows. The calculation of each measure is specific to each event type.

TABLE 3. OVERVIEW OF PERFORMANCE MEASURES

Performance Measure	Definition
Quality Score (QS)	Unique to event type.
Lead-Time (LT)	Difference between the warning-receipt time-stamp and ground truth earliest reported date, as an integer number of days.
Precision	Number of warnings with an Event Date during the scoring period that are matched to GSR events divided by the number of warnings submitted with an Event Date during the scoring period (#Warnings matched/#Warnings submitted)
Recall	Number of events with an Event Date during the scoring period that are matched to warnings divided by the number of events during the scoring period (#Actual Events matched/#Actual Events)
F-Score	The harmonic mean of the Precision and Recall. 2 times the product of Precision and Recall divided by the sum of Precision and Recall.
Mean F-Score	The F-Score metric, computed using all countries.
Mean LT	Mean LT for all W-E pairs, reported for each country for a scoring period.
Mean QS	Mean QS for all W-E pairs, reported for each country for a scoring period.
Mean Precision	The Precision metric, computed using all countries.
Mean Recall	The Recall metric, computed using all countries.

Not all performance measures are relevant to every event type. The following table summarizes the computed performance measures for all Mercury Challenge event classes. An "X" in a cell indicates that the measure is computed for the associated event types.

TABLE 4. SUMMARY OF COMPUTER PERFORMANCE MEASURES BY EVENT TYPE

Event Type	Quality Score	Lead Time	F-Score	Precision	Recall	Mean QS	Mean LT
Non-violent Civil Unrest	X	X				X	X
Disease	X	X					
Military Activity	X	X	X	X	X	X	X

PRECISION AND RECALL FOR MILITARY ACTIVITY EVENTS

Precision = # matched warnings/ # warnings

Example: 84 of 100 warnings matched to events => Precision = 0.84

Recall = # matched events/ # events

Example: 84 of 126 events matched by warnings => Recall = 0.67

F-Score = 2 x Precision x Recall/(Precision + Recall)

Example: Using Precision and Recall as computed above,

```
F = 2x.84x.67/(.84+.67) = 0.75
```

TABLE 5. CONFUSION MATRIX

Confusion Matrix		Events	
		Matched	Unmatched
Warnings	Matched	84	16
	Unmatched	42	NA

LEAD TIME (LT)

Difference in days between warning receipt and earliest reported date for the event.

> Example: Warning submitted 2016-04-10 is matched to an event first reported on 2016-04-17. Lead Time is 7 days.

Note: For some events the submission date may be after the event date but before the earliest reported date. This is often the case for disease events, where the public reporting lags the event significantly. This is still a valid warning and will have a positive lead time.

DETAILED GSR GUIDANCE

Note: This Appendix provides insight into decisions made in the process of teams creating the Mercury Challenge GSR, including some elements of the GSR that are not relevant to the Mercury Challenge.

MILITARY ACTIVITY

DETAILED COLLECTION GUIDANCE

This section provides guidelines to researchers responsible for the collection of articles related to the Military Activity (MA) portion of the Gold Standard Report (GSR). MA events will include event locations and/or state actor involvement from the Areas of Responsibility (AoR). The AoR for the MA portion of the Gold Standard Report (GSR) are Bahrain, Egypt, Iraq, Jordan, Lebanon, Saudi Arabia, Syria, and Qatar.

OVERVIEW OF THE ARTICLE COLLECTION PROCESS

What is article collection?
Article collection is the first phase of the event collection and encoding process. During this phase, the collection staff (also known as 'collectors') are responsible for identifying articles that contain events

carried out by State Actors (SA) within the eight countries of the study. To perform this task, the collection staff reviews articles found in pre-selected news sources; hereafter referred to as 'Gold Standard Sources' (GSS). Using Google Advanced Search (GAS) or Lexis Nexis, the Collector searches the GSS sources to identify articles using a standard set of keywords, which will be discussed in more detail in the "
Article Collection" section. The collector then examines the returned articles to identify and catalogue those which contain events. After the articles with events are identified, they are passed along to the event entry team so that the events occurring within the article are encoded into a relational database.

What is an event?
Military activity events are events by any state-sanctioned military, police, or security organization. If an article contains an event as described, it will be collected.

Events that are not MA include:
- International Coalition events
  - Variations of this include: Coalition, Air Alliance, Alliance Air, and Air International Alliance. (This is includes the United States, United Kingdom, Canada, Australia, and New Zealand as well as Israel, France, Germany, and Netherlands)
- Non-physical attacks (ex. Cyber-attacks).
- Events in the planning stages.
- Communication events (which include meetings, foreign or domestic).
- Events considered criminal activity (ex. Looting, theft, organized crime).
- Routine patrols.

The Collector is responsible for identifying articles that contain events and collecting them for the Entry Staff to encode into the database. If the article does not contain an event as described above, it should not be collected.

DAILY OPERATIONS

At the start of each shift, the Collection Staff will access the shared document titled "Master Collection Sheet". This document will serve as the task orders for the day unless the Collection Staff is informed otherwise by management.

Master Collection Sheet
The day to day activities of the Collection Staff are organized using the 'Master Collection Sheet'. This document serves as a continuous work order for collectors. In this document, each GSS source is listed along with every day in a given month. The Collection Staff is responsible for ensuring that every source is reviewed for every day of the month.

As the collection staff member begins their shift, they select the lowest row that has not been marked as 'In Progress' or 'Completed'. This row will represent the first task order of the day. The collector will enter their name into the Collector column, indicating they are responsible for collecting articles from that source for the date specified. They will then mark the status column for that row as 'In Progress'. At this

point, the collector is responsible for utilizing the search methodology detailed in the "Article Collection" section for the time period specified by the row selected.

The table below contains each column within the Master Sheet and how those columns are to be utilized.

TABLE 6. MASTER COLLECTION SHEET COLUMNS DEFINED

Column	Defined
GSS	This column has each of the 22 GSS listed out for each collection date
Date Published	Serves as the date filter utilized for the GAS or Lexis Nexis Search. By the end of the collection period, all dates within the month will be marked as collected for each GSS, ensuring 100% coverage of the GSS list.
Date Accessed	The date that the collector began collection for a specific source
Collector	The last name of the collection personnel that is assigned to a specific source and date
Status	Whether the collection is unassigned (blank), in progress, or complete
English Returned	The total number of Google results returned when English keywords are used in GAS. Accounts for all results in all languages excluding results in Arabic.
English Collected	The number of articles collected from GAS in all languages excluding Arabic
Arabic Returned	The total number of Google results returned when Arabic keywords are used in GAS. Excludes results in all other languages except Arabic.
Arabic Collected	The number of Arabic articles collected from GAS only
QA Reviewed	Utilized by Collection QA to indicate whether a collection is pending review by spot checking or has been reviewed. "No" means pending review and "Yes" means review is complete.

The English Returned, English Collected, Arabic Returned, and Arabic Collection fields will account for the total number of results that were returned and collected using the English and Arabic keywords.
- A collector must apply the keywords, domain name, and collection date before totaling the number of results that were returned/collected from GAS.
- If the collector goes into the site to get articles, and no results are returned on the GAS, the collector must mark these fields with a zero.
- If the collector collects any articles from the site itself, these will not be included by the collector in the "Collected" columns.

Daily Workflow
At the beginning of each shift, the Collection Staff will use the workflow below until all sources for a given month have been reviewed. At such time, they are responsible for moving on to the next month of encoding when appropriate or contacting their supervisor for additional tasking.

1. Open the Master Collection Sheet on Google Drive.
2. Select the tab that pertains to the month that collection is currently working on.
3. Scroll down to the date that other collectors are working on.
4. Collection personnel will go through all GSS for the region that the source applies to, as seen in the "GSS List" tab on the Master Sheet.

      a. For example, if a collector is looking for articles on the Jordan Times, they will look for events that took place in Jordan.

5. Collectors are to put their last name down for the next available GSS on the collection date that they are working on.
6. Mark the collection as "In Progress".
7. If the collectors cannot gain access to the GSS, they are to mark the collection as "Inaccessible" and collect from that source once the site becomes available again.
8. Collectors mark the "Date Accessed" field with the date that they began collection for that source.
9. When collection for a GSS is fully completed, collectors mark the collection as "Complete".

ARTICLE COLLECTION

Article collection is performed using a structured process that utilizes GAS, Lexis Nexis Advanced Search, or a combination of both. This section will detail which sources to use, which key terms to use, and how to perform the searches using the technologies available.

What are the sources?

Collection staff are responsible for collecting from the Gold Standard Sources (GSS). The Collection Staff is only responsible for identifying articles that have events involving the country specified for the source. For example, if Gulf News (a Bahrain source) has an article with an event in Egypt, the collector is not responsible for collecting the article from Gulf News. On the other hand, if Gulf News contains an event involving Bahrain, the article may merit collection.

Depending on the source, the methodologies for GAS and/or Lexis Nexis Search may be utilized.

WHAT ARE THE KEYWORDS?

Each source will be searched using the English and Arabic keywords identified in the tables below.

TABLE 7. ENGLISH KEYWORDS FOR MA

English Keywords
Aid Attack Attacks Bomb Bombed Campaign Civilians Coalition Forces Gunman Hostage Hostages Kidnapped Killed Maneuvers Military Patrol Raid Raids Rebel Rescue Seize Soldier Soldiers Surrender Terrorism Terrorist Terrorists

TABLE 8. ARABIC KEYWORDS FOR MA

Arabic Keywords
الإغاثة أغار مداهمة هجمات قنبلة قذيفة قصف حملة المدنيين
التحالف القوات مسلح رهينة الرهائن خطف ذبح مناورات عسكري دورية غارة غارات متمرد إنقاذ
استولى جندي جنود استسلام إرهاب إرهابي الإرهابيين

Google Advanced Search (GAS)

GAS is often the best method for collectors to discover events reported in online news media. Note that the search engines available on a news media site frequently miss relevant articles. The steps below provide an overview on how Google Advanced Searching of a given GSS should be performed.

1. Access GAS here: http://www.google.com/advanced_search.
2. Copy/paste ALL of the keywords into the "any of these words" box.
3. In the "site or domain" box, enter the web address of the GSS to search. Enter only the site itself – omit the www., http://, and the slash (/) after .com. (Ex. Iraqinews.com)
4. Click the "Advanced Search" button at the bottom of the page. (This will bring up the Google Search Results screen.)
5. Click on the "Search Tools" button underneath the search bar. (This will bring up another line below the top line.)
6. Open the drop-down menu by clicking "Any time" and select "Custom range…" (This is done to narrow down the results to the collection date.)
7. Enter in the date that corresponds with the current row being processed in the "Master Collection Sheet" are collecting for into the "From" and "To" fields. (The dates in both fields should be the same.)
8. Click "Go". (This will bring up the Google Search Results screen again, this time with results bounded by the collection date.)
9. Review ALL the returned headlines – skip opinion articles, blogs, any articles that are clearly about events in other countries, and events that do not merit entry.
10. If the GAS returns articles that are in another language besides English and Arabic, check those articles in addition to the English and Arabic to determine if they are collectible. (Ex. If a GAS is done using English keywords and French results are returned, check those results to see if they contain collectible events).If they do, the collector may collect these articles.
11. If the GAS date for an article does not match the date on the site itself for the same article, go with the earliest date
    a. Ex. The GAS returns an article that says the published date was Feb. 14th on Google but the site itself says Feb. 15th. Go with the 14th.
    b. This happens because the article was possibly posted on the 14th but then updated on the 15th.

Lexis-Nexis Search

Lexis-Nexis provides a search service for many "print" versions of the GSS's. For those GSS's available on Lexis-Nexis, select and download Nexis articles using the keywords from Table 4 and 5. Instructions for selecting and downloading Lexis-Nexis articles follow:

1. Go to the library home page.
2. Click the "Databases" option on the top-left side of the website.
3. Click "Alphabetical Listing of databases".
4. Scroll down and click "LexisNexis Academic".
    a. Can also press Control + F and search LexisNexis.
5. Input the keywords into the Search bar under "Academic Search".

       a. Make sure there is an "or" that separates each keyword.

6. Click the drop down arrow next to "Advanced Options".
7. Input the collection date into both the "From" and "To" boxes.
8. Type the source name into the box under "Source".
9. Choose the option with "(Arabic)" next to it when using Arabic keywords.
10. Click "Apply".
11. Click the "Search" button.
12. Review the returned articles to determine whether each article contains event, as described in the "What is an event" section.
13. If the article contains an event, collect the article using the methods described in the "Article Entry" section, making sure to search the database for the article of interest before entry, as it may have been collected in the GAS search
14. Once the article has been collected or determined to not merit collection, move on to the next result in the list.

The table below indicates which source within Lexis Nexis should be selected from the drop down menu for the corresponding GSS:

TABLE 9. LEXIS NEXIS SOURCE LIST

GSS	English	Arabic
Gulf News	Gulf News (United Arab Emirates)	X
Arab News	Arab News	X
Bahrain News Agency	Bahrain News Agency (BNA)	Bahrain News Agency (BNA) (Arabic)
Almasry Alyoum	X	Al-Masry Al-Youm (Arabic)
Egypt Independent	Egypt Independent	X
Daily News Egypt	Daily News Egypt	X
NINA News	National Iraqi News Agency (NINA)	National Iraqi News Agency (NINA) – Arabic
Jordan Times	Jordan Times, The	X
Al Bawaba	Albawaba.com	Albawaba.com (Arabic)
NNA	National News Agency Lebanon (NNA)	National News Agency Lebanon (NNA) (Arabic)
Al Riyadh	Al Riyadh	X
SANA News	Syrian Arab News Agency (SANA)	Syrian Arab News Agency (SANA) (Arabic)

IN-SITE SEARCHING

Collectors follow a step-by-step process to look for relevant articles on every GSS website. The steps provided cover both the English and Arabic versions (respectively) for each GSS. Collectors are to look for relevant articles on the GSS's website after the GAS collection is completed. This is to ensure that all collectible articles were gathered for that collection date, region, and source.

● Make sure to go through both the English and Arabic versions of the site (if both exist) and collect relevant articles.
● Focus on one language at a time.

ARTICLE ENTRY

Upon finding an article that has an event, the Collector is responsible for 'collecting' or encoding the article into the database. This process utilizes a relational database to capture key fields from the article and source before the article is handed over to the event entry staff.

Does it merit collection?

After performing an article collection, collectors will determine if the articles they reviewed will merit collection and entry. Articles that merit entry will be contain MA events that meet the criteria outlined in the section "What is an Event". After it has been determined that the article contains an enterable MA event, the collector will collect the source and enter key details of the article into the database.

Does it already exist in the database?

Upon finding an article that contains MA events, the Collection Staff is responsible for ensuring the article will not be entered into the database twice. Double entry may occur when two different collectors identify the same article from the same source despite running separate queries. In order to reduce duplicate entry, the collection staff should follow the procedure below:

1. If an article merits entry, first use the GSS Collection search screen in the database to ensure the article is not already in the database.
   a. This helps diminish the possibility of duplicates, and saves time for both collection and entry.
2. Copy and paste the original title of the article word-for-word into the "Title" field.
   a. Example: If the article was originally in Arabic, copy and paste the Arabic title into the "Title" field.
3. If nothing returns, copy and paste the first line of the original title into the "Title" field and search.
4. If nothing returns, use the first few words of the first line of the original title, paste into the "Title" field, and search.
5. If the article is already in the database, review it and make sure that every field is filled in properly with correct information.

How to Enter the Data

Once the Collection Staff confirms an article has not been entered, they must complete all required Source fields. The table below contains all Source fields within the database which are to be filled out by the Collection Staff upon collection of an article containing an MA event.

TABLE 10. SOURCE FIELDS DEFINED

Field Name	Definition
Created By	Name of Collector. This is automatically filled in when a new article is being entered.
Source Name	Enter the name of the source that the article was collected from.

Field Name	Definition
Date Accessed	The date that the Collector accessed the article on. This field will be automatically filled in.
Language	Select the *ORIGINAL* language of the article
Title	Enter the title of the article in its *ORIGINAL* language
Translated Title	Enter the title of the article translated into *English*
Article Text	Enter the article text translated into *English* (**NO original language text**)
Article Code	Leave this field blank
Sorters Note	Leave this blank unless leaving a note for a sorter about additional information
Reasons for Review	Leave this field blank
Number of Events in Article	Leave this field blank
View Translated Text	This button does not pertain to collection
Finish and Sign Sorters Note	Do not click this button. It does not pertain to collection.

****Please pay attention as to when to enter an original language or English****

Once all fields have been entered, the collector will proceed to the Master Collection Sheet and begin with the next row that is not marked as "In Progress" or "Complete". Collectors must finalize previous row by marking it as "Complete".

<div align="center">ADDITIONAL RESOURCES</div>

How to find Additional Information
If the event in the article is too vague for a collector to determine if it is collectible or not, additional research can be done using ANY of the GSS's that we use for collection or a reputable non-GSS.
- There are no region restrictions for the GSS's when looking for more information about an event.
  - Ex. Collectors may look at Jordan Times for an event that happened in Saudi Arabia.
- The additional information cannot be from the Prohibited Links list.
- If information is found from another article that updates or supplements the GSS event, the collector must put the article URL in the sorters note with their first initial and last name. Further, they must leave a note to the sorters stating what the additional link contained that supplemented the GSS.

Prohibited Links
Below are a list of websites that cannot be used to find additional information:
- Twitter
- Facebook
- Wikimapia
- Wikipedia
- WikiLeaks
- Google Maps
- Reddit

- Maplandia
- Google Search Pages
- GeoNames
- Falling Rain
- Instagram
- YouTube

SPECIAL CASES

1. If an article has an event that takes place in a border area that is between two regions and a specific geographic location is not included that points to what country the event took place in, the event is not collectible.
2. The Lebanese military conducts an operation against militants in the Qalamoun mountain range on the Syrian-Lebanese border but a specific location is not provided, the event is not collectible.
   a. The Qalamoun mountain range is in both Syria and Lebanon and it cannot be determined which country the event took place in.
3. If security forces surround/besiege a city, building, etc. where terrorists are located, it is collectible.
   a. The event is considered a Security Forces - Force Posture.
4. If an AOR countries' military forces arrive in a non-AOR country to carry out operations against NSAs, it is collectible.
   a. The event is considered a force posture.
   b. Ex. Saudi Arabian military aircraft arrive at Incirlik air base in Turkey to conduct operations against militants in Syria.
5. If an AoR State Actor seizes/captures a vehicle belonging to an NSA, it is collectible.
   a. Ex. Iraqi military forces seize an ISIS truck in Iraq.
6. If a rocket/missile/artillery piece/etc. is fired into Saudi Arabia from Yemen with no specific actor mentioned, it is collectible.
7. If the Saudi-led coalition/Saudi military forces bomb an area in Yemen and the target is not Yemeni military forces, the event is considered a Security Forces- Armed Conflict event and is collectible.
8. For articles where it states that security forces raised a flag over a building, bridge, etc., make sure the article states that the location was liberated from terrorists/militants/NSAs or that a collectible event leading up to the flag raising took place. If the article just says that security forces raised a flag over a location, do not collect.
   a. Ex. 1: Iraqi Army forces liberated Ramadi from ISIS militants and raised the flag over the local government building- collectible.
   b. Ex. 2: Iraqi federal police raised a flag over the bridge in the area of Anbar- do not collect.
9. If security forces evacuate civilians from an area, it is collectible.
   a. This event is considered a force posture.
10. If an AoR State Actor drops leaflets on an area inside or outside of the AoR (where the AoR country is the actor), the event is collectible.
    a. This event is collectible because it is considered a force posture.

   b. Ex. http://www.iraqinews.com/iraq-war/iraqi-aircrafts-throw-leaflets-on-heet-demanding-people-prepare-for-liberation/
11. If security forces come across an IED/bomb making plant, stash of IEDs/weapons, etc. with no connection to who the materials/weapons belong to and no mention of a defusal/dismantling/seizure, do additional research on who the materials belong to and if they were dismantled/defused/seized. If additional research turns up nothing, the event is DME.
   a. Ex. https://www.ninanews.com/News_Details.aspx?sv%252bdg52nPDighCE7X5K6UA%253d%253d

DETAILED ENCODING GUIDANCE

This section provides guidelines to researchers responsible for encoding events related to the MA portion of the Gold Standard Report (GSR).

HOW TO ENCODE MA EVENTS

Following collection, articles are provided to entry staff to be encoded into a relational database. During this process, articles are divided into a timeline of events, each described by a series of structured and unstructured fields. Table 12 provides an overview of these fields. The values for structured fields are provided in more detail in the following section.

EVENTS SEPARATED BY TIME AND SPACE

Complex events, such as coordinated attacks or a string of attacks are encoded as unique events when the attacks or actions differ in time or space. All events will have one date / time and one geographic location. Use the table below for additional guidance when entering complex attacks.

TABLE 11. EVENT SEPARATIONS

If the following....	Perform the following	Examples
If an event is separated by time	Encode multiple events	- The Syrian Arab Air Force conducted airstrikes in Raqqa on Monday and Tuesday. - Army units clashed with terrorists who attacked army checkpoints. Army units are targeting and pursuing the remaining terrorists who fled. - Minutes ago, the Syrian Arab Army … imposed full control over Abu Allaj village in the western countryside of Al-Raqqa. The Syrian Armed Forces originally liberated Abu Allaj on Thursday, but were forced to retreat after the Islamic State of Iraq and Al-Sham (ISIS) launched a large counter-offensive to recover the points they lost along the Salamiyah-Raqqa Highway.

If an event is separated by location	Encode multiple events	• For reports that attacks occurred at five different mosques, houses, headquarters, buildings, etc. within the same city, five separate events are encoded with the same city as the location. • For reports that attacks occurred in a specific location and other areas, the specific location and other areas will be separated. • For reports that clashes occurred in "Town A", SA established control of "Town A" while clashes continue west of "Town A", "Town A" and west of "Town A" will be separate events
If events occur simultaneously in one location	Encode as one event	• Militants attack a military base and the military repels/repulses/responds to the attack. • If an event states "regime forces conducted four airstrikes" the researcher will encode one event unless there is separation of time or space.
If event spans over a few days	Default to the first day of the event unless the article details that specific events occurred on a specific day in that range.	• Shelling continuously occurred in Aleppo over the past 4 days, but on Tuesday they caused material damage and X number of casualties. Tuesday will be its own event. • If an event states: "a man died of wounds three days after clashes in area X", the man's death will be added to the event entry for clashes that occurred three days ago.

GSR FIELDS

If the precipitating event action is a crime, the action is not an MA event.

TABLE 12. GSR FIELDS AND GUIDANCE

GSR Fields	Guidance
Event_ID	Unique identifier for each GSR event.
Event_Date	The date of the event. If the date cannot be determined, enter the earliest reported date. Encode in MM/DD/YYYY format.
Default Date	This field is used when the exact date of event is unknown.
Event_Type	This field contains a drop down list of the five event types that researchers can choose between. For detailed descriptions on the event types, see Table 14
Source ID	Unique identifier for each GSS article.
Source Name	Name of the GSS source.
Language	The language of the GSS article.
Date Published	This field contains the date the original article was published. Encode in MM/DD/YYYY.
URL	URL of the report that best describes the MA Event.
Title - Original Language	The article headline in the native language
Title - Translated	This article headline translated into English
Number of Events	This field contains a numerical figure entered by the entry staff.
Earliest Reported Link	This field contains the earliest reported link of the news source that first reported the event.
Earliest Reported Date	This field contains the earliest reported date of the news source that first reported the event.

GSR Fields	Guidance
Additional Links	This field contains any additional URL's that were collected during the collection phase that have information about the event entered. If there is more than one additional link the researcher should separate each link with a semi-colon ";".
Geography (Geo) ID	This field is automatically generated detailing the specific numerical ID of the selected geographical location.
Country	This field contains the name of the country in which the event occurred.
State	This field contains the name of the state/province/department/governorate in which the event occurred.
City/Town/Village of Event	This field contains the name of the city, village, or town in which the event occurred.
Alt City/Town/Village Names	This field contains any alternate names for the City/Town/Village.
Latitude	This field records the latitude (decimal degrees) of the city in which the event occurred.
Longitude	This field records the longitude (decimal degrees) of the city in which the event occurred.
Description	This field contains the feature designations of the locations in the gazetteer.
Approximate Location	This field is used when the exact location is not found in the gazetteer. The researcher has the option of selecting "Yes" or "No".
Actor ID	This field is automatically generated with an unique identifier for the actor.
Actor	This field records the name of the actor(s).
Actor Status	This field contains SA (State Actor).

DETAILED DESCRIPTIONS OF EVENT FIELDS

The following sections provide a detailed overview of each field utilized by the entry staff.

Event Date (Required)

The event date is determined by the information provided within the news article. Entry staff should be careful when entering the date, as sources in the Middle East use the DD/MM/YYYY format. In the database, the date is always recorded in the MM/DD/YYYY format. If the article mentions that the event took place today, then checking the hour of publication is critical—articles with a publish hour of 12:00am-5:00am are most likely referring to events that occurred the day before. If the article does not mention when the event took place and it is a source where there is no time stamp, encode the event date as the first date of the month.

Default Date (Required)

The default date is used to identify whether the event date is known or unknown. If the event date is known, then the default date is "No". If an event date is not known or remains unclear, then the default date is "Yes". An article may provide some hints, such as "yesterday," "three days ago," or "last week." In such cases, the researcher counts backward to determine the date. If no determination can be made, encode the event date as the date of the earliest reported article.

TABLE 13. DATE ENCODING

If the following…	Perform the following
If an article details the time of the events as "after yesterday's midnight" For example, "clashes between the "Islamic state" on one hand and the Islamic factions and Jabhat Fateh Alsham on the other hand, after yesterday's midnight, at the center of Jabal al Afa`i"	Encode the date as "today" with a default date of "No"
If an article details an event as "Update" or "Breaking News"	Encode the default date as "Yes". Reason: updates and breaking news indicate a time of discovery of an event but not the time the event took place.

Event Type (Required)

The Event Type is used to describe the event that occurred and is determined by the actor. Table 14 provides a detailed description of each event type. Excerpts from articles detailing each event type can be found below.

TABLE 14. EVENT TYPES AND DEFINITIONS

Event Types	Examples
Security Forces -Conflict Incident where a SA (police, military, or sanctioned other state/government security forces) initiates an encounter with NSA, or takes action in some way. Incident where SA performs an event together and violence occurs.	Airstrikes, engaging SAs in combat, hostage rescues involving force, drone strikes, raids (includes arrests). Cannot be conducted by the United States, United Kingdom, Canada, Australia, or New Zealand as well as Israel, France, Germany or Netherlands
Security Forces - Force Posture Incident where a SA (police, military, or other state/government security forces) **DO NOT** use force and/or are not met with resistance.	Defusing an IED, relocating personnel without taking other actions, supply drops, accepting peaceful surrenders, arrests made without force, non-routine and atypical security patrols through a city, etc.

Incident where a SA perform event together without using force and/or are not met with resistance.	

Excerpts:

Security Forces – Armed Conflict:

- "…while the regime forces renewed the shelling on areas in the city of Darayya in the Western Ghouta, no information about injuries."
- Russian warplanes carried out this morning several raids on areas in the township of Kinsabba and other areas at the border with Turkey in the northern countryside of Latakia, no information about casualties yet.
- It is mention that "the Turkish aircraft, a few days ago, began bombing the headquarters of the PKK in Qandil, Amidiah, Sidakan and Joakourk areas, causing significant damage to forests and villages of the Kurdistan region and property of citizens" *This is a common occurrence where multiple airstrikes are detailed at once. In this type of situation, four separate airstrikes would be entered.*
- "…while helicopters dropped barrel bombs on areas in the town of Ma'aarrat Hermeh and the outskirts of al-Hbit town in the southern countryside of Idlib" *While the article does not detail the actor, FVEY countries do not use helicopters during airstrikes in Syria.*
- "…followed by shelling by regime forces on the premises in the town of al-Taman'aa,"
- Unidentified warplanes conduct operations within the area of responsibility AOR. *After additional research yields no event information, the researcher should forgo entering the event. With that said, events involving helicopters should always be entered.*
- "…Violent clashes continue between Hezbollah and the regime forces backed by gunmen loyal to them from Arab and Asian nationalities against Al-Nusra Front (al-Qaeda in the Levant) and the rebel and Islamic Factions amid advancement by the Hezbollah and the regime forces."
- "…the army units, in cooperation with the popular defense groups, carried out intensive military operations in the north eastern countryside of Lattakia during which they restored security and stability to Ayn al-Ghazal...after eliminating the remaining gatherings of terrorist organizations there."

Security Forces – Force Posture:

- NNA - The State Security of Nabatieh branch on Tuesday managed to catch a Syrian individual (25 years) for belonging to terrorist groups and working in the field of weapons trading. Investigations with him are ongoing in coordination with the competent judicial authority.
- "Idlib province: The helicopters dropped food supply baskets on Kafria and al-Foua which are inhabited by Shia and besieged by Jabhat al-Nusra and Islamic battalions amid targeting the towns with shells by Islamic battalions."
- Spokesman for the Baghdad Operations Brigadier General Saad Maan told the National Iraqi News Agency / NINA / that the security forces today defused a car bomb aimed at the citizens in al- Husseiniya area east of the capital Baghdad. / End

- A source in the cell told the National Iraqi News Agency / NINA / "large military reinforcements, include new manufactured armored vehicles, arrived yesterday to the outskirts of Fallujah, in preparation for military operations and engage in the battles of Anbar province," ./ end
- "Police have arrested Essam Derbala, chair of renowned Islamist group Al-Jamaa Al-Islamiya's Shura Council in Qena governorate Wednesday, following an arrest warrant issued by State Security authorities."
- "The security forces today closed all roads leading to the entrances to the Green Zone in central Baghdad, including the bridges linking the Karkha district to Rasafa, under heavy security deployment in a number of areas of the capital since last night."

TABLE 15. SECURITY FORCES ARMED CONFLICT EVENT TYPE ENCODINGS

If the following…	Perform the following
If an event is initiated, but the goal of the action is not attained • Suicide bombers are shot and killed by security forces before they can detonate their explosives	Only encode the "Security Forces-Armed Conflict" event and not the initiating event since the explosives did not detonate.
If an article details clashes amid warplane bombings	If the event type is the same ("Security Forces - Armed Conflict"), then encode as one event. If the clashes are an "Armed Assault" and the warplane bombing is a "Security Forces - Armed Conflict" then encode as two events.
If an article details the destruction of weapons caches and other SA military equipment	The researcher should encode the event as a "Security Forces – Armed Conflict", since there was force used in the event.
If an article details coinciding aerial attacks in an area by warplanes and helicopters • For example, "Coincided with more than 80 intense airstrikes by 7 helicopters and warplanes targeting positions for the fighters of the factions"	Encode as one event - "Security Forces – Armed Conflict".
If an article details SAs carrying out a raid or ambush in an area • For example, "A joint security force seized a car bomb and arrested the driver in ambush northwest of Baquba in early hours on Monday."	Encode the event as "Security Forces – Armed Conflict".
If an article details an event where an SA aborts an active event • For example: *"In the suburbs north of Soueidaa, the army aborted an attempted infiltration of an armed terrorist group in Tal Bouthaina, killing and wounding a number*	Encode event as a "Security Forces – Armed Conflict" Although attempted events are not entered, the event is considered completed when the army kills and wounded terrorists and destroyed their weapons.

If the following...	Perform the following
among its elements and destroying their weapons."	
If an article details clashes amid advancements in an area by an SA	Encode as a "Security Forces – Armed Conflict".
Syria Specific - If an article details unidentified warplanes attacking alongside the regime	Encode as "Security Forces - Armed Conflict". FVEY countries do not work in coordination with the regime.
If an article details an event where a helicopter was used to launch attacks on an area	Encode as "Security Forces – Armed Conflict".
If an article details clashes amid shelling	If both the clashes and shellings are a "Security Forces – Armed Conflict", encode as one event. If the event type for the clashes and the shelling are different encode as two events.
If an article details mutual shelling amid clashes	Encode as one event. If the actor for the clashes is a SA, encode as "Security Forces – Armed Conflict".
If an article details SA liberating or freeing an area	Encode as a "Security Forces – Armed Conflict".
If an article details an event where SA kills non-state actor and dismantles IEDs in an operation	Encode as one "Security Forces – Armed Conflict".
If an article details an event where SA evacuates civilians during clashes with non-state actor	Encode as one "Security Forces – Armed Conflict".
If an article details an event where SA destroyed weapons or a device allegedly used by NSA • For example, "Iraqi security forces shot down two drones allegedly used by ISIS"	Encode as "Security Forces – Armed Conflict".
Egypt Specific – If an article details "arrests of members of the Dignity Party, April 6 Movement, and other groups/organizations that support/are in association with the Muslim Brotherhood"	Encode the event as "Security Forces – Force Posture" if there was no violence. If violence occurred during the arrest, then encode as "Security Forces – Armed Conflict".
Egypt Specific – If an article details a protest staged by the Muslim Brotherhood which results in SA arresting members of the Muslim Brotherhood	Encode the event as a "Security Forces - Force Posture" if no violence was present, or as a "Security Forces - Armed Conflict" if the arrest was accompanied with violence. ***Muslim Brotherhood is considered a terrorist group by the Egyptian government.***
If an article details an event with mixed reporting's where one reporting is made by a social media channel and the second is reported by the GSS	Encode based on the GSS.
If an article details an event where a GSS is reporting an event based on the report given by a civilian, activist, SA, etc. (by everyone except by a NSA)	Encode the event. Reporting made by the GSS cannot come from a NSA source.

If the following…	Perform the following
• For example, "The local residents told the … NINA … that "members of the resistance were able to kidnap one of Daash police (Husba) in Surj Khana market in central Mosul, and stabbed him to death and then threw his body"	
If an article details an event where clashes are occurring amidst one side's advancement For example, "Clashes took place between regime forces and rebels after an attack by rebels on the area leading to material damages, the regime advanced in the area amid aerial bombardment on Khan al-Shekh."	Encode the event based on the advancing side Regime advancement would make the event an "Armed Conflict"

TABLE 16. SECURITY FORCES FORCE POSTURE EVENT TYPE ENCODINGS

If the following…	Perform the following
If an article details a patrol following a bombing and it resulted in the security personnel apprehending an individual or locating an IED	Encode as "Security Forces - Force Posture".
If an article details that a SA provides medical or food aid…	The researcher should encode the event as "Security Forces – Force Posture", unless the actor is from a FVEY country.
If an article details an event where a helicopter dropped baskets, papers (leaflets), or non-violent items in an area	Encode as "Security Forces – Force Posture".
Egypt Specific – If an article details "arrests of members of the Dignity Party, April 6 Movement, and other groups/organizations that support/are in association with the Muslim Brotherhood"	Encode the event as "Security Forces – Force Posture" if there was no violence. If violence occurred during the arrest, then encode as "Security Forces – Armed Conflict".
Egypt Specific – If an article details a protest staged by the Muslim Brotherhood which results in SA arresting members of the Muslim Brotherhood	Encode the event as a "Security Forces - Force Posture" if no violence was present, or as a "Security Forces - Armed Conflict" if the arrest was accompanied with violence. ***Muslim Brotherhood is considered a terrorist group by the Egyptian government.***
If an article details an event where weapons and/or personnel from an outside source are entering an AOR to enhance an AOR's military and is not a sales transaction between countries	Encode event as "Security Forces – Force Posture"
If an article details an event where SA advance to a location but then retreats prior to taking an action.	Encode event as "Security Forces – Force Posture"
If an article details an event where SA evacuates civilians in an area prior to violent conflicts	Encode as "Security Forces – Force Posture".

If the following…	Perform the following
If an article details an event where SA arrests an unidentified gunmen/assailant for planting bombs	Encode as "Security Forces - Force Posture". Planting bombs push the individual outside the realm of crime and into extremism.
If an article details the discovery and collection of NSA weapons cache by a SA as a non-violent event • For example, if an article details a military raid of a warehouse, without physical NSA security or presence, that resulted in discovery and collection of weapons cache.	Encode as "Security Forces - Force Posture".
If an article details an event where SAs lay siege to a city without force and clashes • For Example, "The security forces began its siege on the town which cut off all supply lines excepting the road connect the road between the town and Ḥammām al ʿAlīl (حمام العليل) (30) km south-east of Mosul). "	Encode the event as "Security Forces - Force Posture".
If an article details an event where SAs destroy seized items, weapons, etc in a controlled setting For example, "While the army forces seized four IED's, the military expert worked to detonate them immediately in their places."	Encode the event as "Security Forces – Force Posture"
If an article details an event with mixed reporting where one reporting is made by a social media channel and the second is reported by the GSS	Encode based on the GSS.
If an article details an event where a GSS is reporting an event based on the report given by a civilian, activist, SA, etc. (by everyone except by a NSA) • For example, "The local residents told the … NINA … that "members of the resistance were able to kidnap one of Daash police (Husba) in Surj Khana market in central Mosul, and stabbed him to death and then threw his body"	Encode the event. Reporting made by the GSS cannot come from a NSA source.

Does Not Merit Entry (Grey Articles)

Events are encoded based on the criteria detailed above. If an event does not meet the criteria, the event will be designated as "Does Not Merit Entry", or "DME", and will not be encoded. Examples for events that do not merit entry are provided in Table 6 below.

TABLE 17. EVENT DME (DOES NOT MERIT ENTRY)

If the following…	Perform the following
If an article details an event with the actor as Red Cross, UN, Israel, United States, Canada, Great Britain, Australia or New Zealand.	Do not encode event. **Reason**: All events with FVEY actors or non-government organizations where FVEYs are active are DME.
If an article details an event that does not require a physical presence, i.e. Cyber attacks	Do not encode event. **Reason**: Locations are encoded for each event therefore a physical presence is required for each event encoding.
If an article details an event that is a routine patrol • *For example, "while The Syrian Democratic Forces finished inspecting the three villages they controlled in the past days near the banks of the Euphrates in the northeastern countryside of Aleppo, following clashes against the "Islamic state"*	Do not encode event. **Reason**: Only encode events that are non-routine as it is routine for patrols, reconnaissance of an area, etc.
If an article details a raid following an attack that returns zero results (i.e. no arrests or seizures) • *For example, "This morning, a bomb exploded on a police patrol while passing in Yusufiya area… A security force rush to the area and… carried out a raid to search for the perpetrators of the bombing."*	Do not encode event. **Reason**: A raid in the aftermath of a bombing or clashes is considered routine. The event will only be entered should it yield an arrest or seizure related to extremism/terrorism.
If an article details operations that are occurring amid strikes by International Coalition Forces • *For example, "the Iraqi air force and the international coalition aircraft launched strikes on the Daash organization"*	Do not encode event. **Reason**: The extent of FVEY activity is unknown and since they are considered a co-actor, the event would not merit entry.
If an article details "cessation of hostilities" • *"For example, "Palestinian forces in Ain El Hilwe have succeeded on Sunday in established a ceasefire"*	If the article details that a ceasefire was established, do not encode. However, if terrorists breached the ceasefire using violence, then encode the event.
If an article details "unspecified warplanes" or "warplanes" or "aerial bombardment" or "air force" with no reference to a country or a specific actor	Do not encode event. **Reason**: FVEY actors are involved in the area. If it cannot be confirmed by GSS that they were uninvolved in the event, then the event cannot be entered.
If an article details an event where an SA attacks another SA • *For example, "Saudi fighter jets also targeted Hamzah military base in the Central Ibb Province seven times."*	Do not encode event. **Reason**: SA vs SA events are not included.
If an article details an event where SA is continuing to hold an arrested individual • *For example, "The Egyptian State Security Prosecution held in remand four individuals*	Do not encode event. **Reason**: The only event that can be encoded for this situation is the arrest itself but anything detailing an extension of detaining the arrestee will not merit entry.

affiliated with "Sinai Province" on Monday for 15 days pending investigations, according to local media reports."	
If an article details an extension to an active state of emergency in an area • *For example, "Egypt said on Saturday it had extended by three months a state of emergency imposed on parts of Northern Sinai"*	Do not encode event. **Reason**: Until SAs are involved in a conflict or seize items or individuals linked to terrorism/extremism, there is no event.
If an article details an event where an unspecified warplane drops baskets, paper (leaflets), or non-violent items in an area • *For example, "while a cargo planes dropped for the second time today parachutes carrying aid, they fell on areas controlled by the regime forces on the mountain overlooking the city of Deir Ezzor."*	Do not encode event. **Reason**: FVEY actors as well as FVEY supported non-government organizations are involved in the area. If it cannot be confirmed by GSS that they were uninvolved in the event, then the event cannot be entered.
If an article details an event where SA arrest an individual or individuals without ties to extremism/terrorism • *For example, "State Security Patrol arresting a 'dissident' a Syrian Army officer"*	Do not encode event. **Reason**: A SA arresting an individual without ties to extremism/terrorism is considered to be normal criminal activity. Do not encode for normal criminal activity.
If an article details an event where weapons are being bought by/sold to an AOR • *For example, "On Wednesday Russian newspaper revealed that Iraq signed a contract worth $1 billion USD with Russia for the supply of weapons, noting that the contract was concluded during the visit of the Defense Minister, Saadoun al-Dulaimi to Moscow last week."*	Do not encode event. **Reason**: Transactions will not be encoded. It will only be encoded if an AOR is receiving or giving support through manpower and weaponry to a specific location and not as a general transaction for the AOR.
If an article details an event where either an SA or NSA dies from injuries sustained but with no indication as to the cause of the injuries • *For example, "fighter passed away due to wounds sustained earlier"*	Do not encode event. **Reason**: Cannot determine the cause of injuries and death therefore it cannot be entered.
If an article details an event where a body is found dead with no information as to cause of death • *For example, "At least 250 civilians were killed in Aleppo in the past nine days, according to the Britain-based Syrian Observatory for Human Rights."*	Do not encode event. **Reason**: Only encode the event if the source indicates how the civilians were killed.
If an article details SA capturing a human trafficking network without ties to extremism/terrorism • *For example, "The statement noted that a force of the brigade (54) managed to capture a network of trafficking children within the Mansour district"*	Do not encode event. **Reason**: Without ties to extremism or terrorism, human trafficking is considered criminal activity and will not be encoded.

If an article details an event where an explosion is heard but the cause is not given or is not related to extremism/terrorism • *For example, "An explosion was heard in the outskirts of the city caused by a motorcycle exploding"*	Do not encode event. **Reason**: Only encode if the cause of the explosion is given and it is related to terrorism/extremism, such as bombing.
If an article details an event where SA conducted a raid and seized weapons with no ties to terrorism/extremism • *For example, "The security operation carried out ... orders of the judicial arrest ... resulted in the arrest / 36 / ... he said he was also find ... mortar rounds and 82 mm fillings launcher with 106 mm cannon shells, and a quantity of ammunition and light"*	Do not encode event. **Reason**: Without ties to extremism and terrorism, the seizure is considered a seizure from normal criminal activity.
If an article details an event that is a summary of an operation over several months • *For example, "The dismantling of explosive devices and booby-trapped houses in the city of Ramadi that started two months ago, resulted in the killing of 15 members of the army and police."*	Do not encode event. **Reason**: Without concise event information, the time frame is too broad to research and enter exact date information. Summaries over several months will not be encoded as concise events cannot be determined.
If an article details an event where a non-government organization provides humanitarian aid in an AoR For example, "Al-Mawadda charity association started distribution of 500 food packages….in Hasaka"	Do not encode event. **Reason:** Non-government organizations are NSAs. NSAs cannot be an actor for Force Posture events and there is no encoding for non-violent NSA-led events.
If an article details an event where State Actors are providing discounted aid For example, "Army in cooperation with government organizations distributed 2 million food boxes at a discounted price in the poor villages and areas selling it for 25 Egyptian pounds instead of 87 Egyptian pounds."	Do not encode the event. **Reason:** Treat discounted aid as sales of items and not as aid provided.
If an article details an event where the actor is detailed as an individual who works for or on behalf of the government but is not detailed to be a State Actor (military, police, security forces, etc) • For example, "Federal Ministry of Electricity, on Sunday, he had sent engineering teams to the city of Fallujah in Anbar province in order to conduct the initial detection of the damage caused to the power stations … due to the very large terrorist acts"	Do not encode the event. Only encode State Actor events when a State Actor is the actor of the event.
If an article details condolences or condemnations of an event	Do not encode the event. These are second-hand discussions and not reportings of an event.

Event Summary (Required)

27

This field is used to summarize the "Who, What, When, Where, Why and How" of each event within an article. The event summary cannot surpass five lines and must be directly copied and pasted from the article, thus leaving any misspellings or grammatical errors. If there are certain parts of the text that is not pertinent to the event, researchers must utilize an ellipses (…) to note any deleted text. If there is additional information that the researcher adds to the event summary, brackets ([ ]) must be utilized. If the researcher adds additional information from a source in the additional links, the following format must be used: "Update from (Source Name) [unique URL identifier]: added source information".

Items of Note

The Items of Note (ION) is used to capture detailed information that cannot be encoded in the GSR Complete Entry fields. ***ION should not contain any information that can be found in the event summary, with the exception of Exact Geo.***

TABLE 18. EXAMPLES OF ION ENTRIES

Type of Information	When to Use	Examples
Exact Geo:	Used when the location is provided, but is not in the gazetteer. *This will always be included in Items of Note.*	Exact Location: Al-Mansoura
Actor:	Used when the actor provided is more specific than the actors provided in the database.	Actor: Islamic battalions (coded as Unspecified NSA)
Casualties Split Between Event IDs:	Used when the casualties are split among several events.	Target casualties split between events ####, ####, and ####
Non-Target Casualties:	Used when the article reports deaths/injuries from the actor.	Actor deaths: #
Noteworthy Person Killed/Arrested:	Used when the article specifies the name of a person either killed or arrested.	The terrorists killed/arrested were Yusuf Ramadan and Samer Abu Hashish.

Number of Events (Required)

This field records the number of enterable events the researcher discovers within the article. This field should be a numeric figure. This information is delivered to the client as a quality assurance measure as it is compared between the Entry, Sorting, and QA team who each provide a value for this field separately. The entry staff must encode the number of events within a given article, regardless as to whether or not those events have already been entered into the database under a different GSS source.

Earliest Reported URL (Required)

This field contains the earliest reported link that first reported the event, and the link does not have to come from a GSS source. If only the GSS link exists for the event, it may also be used as the Earliest Reported URL.

Earliest Reported Date (ERD) (Required)

The ERD is the publication date of the article that first reported the event. The date must be entered in MM/DD/YYYY format. If the GSS publication date is the earliest reported date, it may be used as the ERD.

Additional Links

The additional links are the links of other articles that report the event - if applicable. The additional links should not contain any URLs from the Prohibited Links list detailed in that section. Examples of additional links include updates to an event from a separate article, articles that provide additional context not available in the initial GSS source (ex location details), or other GSS articles that are identified, even if no new information is present. Researchers are encouraged to use IHS Janes as an additional link for ambiguous events. An event must be reported as enterable by a GSS, therefore, additional links cannot be used to make an event enterable.

Location

Each event has one location which may be coded to the Country, State (also known as Province, Department, Governorate, etc.), and City (also known as Town and Village) level of detail depending on the information provided in the article. Always encode the lowest level of detail given in the article for each event described.

Country (Required)

This field identifies the country where the event occurred.

State (Required)

This field identifies the state/province/department/governorate where the event occurred. If the researcher cannot locate the state in the gazetteer, then they must conduct additional research. If the state still cannot be discovered, then researchers should default to "NA" and complete a "Location Request Form".

City/Town/Village (Required)

This field identifies the city, town, or village in which the event occurred. The researcher must enter the city, town or village of the event in the "City/Town/Village" field. If the researcher is able to locate the

location, but it is not in the gazetteer, the researcher must identify the nearest, largest city, town, or village within 50 kilometers (km) of the reported area.

A city may refer to a city-name or another descriptor such as "section of populated place (neighborhood)", "strait", "plain", "mountain", "road", or "mosque".

TABLE 19. LOCATION ENCODINGS

If the following…	Perform the following
If the researcher is unable to locate the neighborhood in the gazetteer	Up-aggregate to the city, town, or village that the neighborhood is located in.
If the article details a specific "strait", "plain", "mountain", "road", or "mosque" that an event occurred in, but the location is not in the gazetteer	Follow the 50 km rule and flag the event as an approximate location. The researcher should include the name of the feature in the Items of Note and submit form for the feature to be added to the gazetteer.
If the article details a specific "strait", "plain", "mountain", "road", "mosque", etc. that an event occurred in and it is located within the gazetteer	Encode the strait, plain, mountain, etc. as the city.
If the article does not specify the city where the event occurred, but states that the event took place in front of a particular building or street	Use Google to verify the location of the building or street and encode the location accordingly. (If the researcher cannot find the location after five minutes of searching, encode the field as "NA" and detail in Item of Note.
If after additional research and the city is still unknown	Encode the city as "NA".
If the article states that an event occurred "around city x" or "in the vicinity of city x"	Encode the city as "X" or the most detailed location.
If the article states that an event occurred in the countryside of a province	Encode the city as "NA".
If the article states that an event occurred in the countryside of a city	Encode the city as the city.
If an event occurs on the road between cities	Encode the city with the larger location (larger is determined first by the population size, then by square footage).
If an event details a source that states "in the vicinity" before a list of geographies For example, "while clashes continued between the regime forces and militiamen loyal to them and Hezbollah against the rebel and Islamic Factions, in the	Encode each geography as a separate location and apply "in the vicinity" to each geography.

If the following…	Perform the following
vicinity of al-Rihan town and Tal Kurdi in Eastern Ghouta"	

If the event location cannot be found in the gazetteer, the researcher is to submit a "Location Request Form". Specific details about this form are located in Section V of the Appendix.

Alternate City/Town/Village of Event

This field is used as an additional search field that contains several alternate spellings, in both English and Arabic, of cities, towns, and villages in the gazetteer.

Latitude/Longitude

These fields are automatically populated with the latitude and longitude once the researcher adds a location to the event.

Approximate Location

The approximate location should only be encoded as "Yes" if the level of detail on the location is unclear.

TABLE 20. APPROXIMATE LOCATION ENCODINGS

If the following…	Perform the following
If the event location was found in the gazetteer	Encode the approximate location as "No".
If the event location is within 50 km of the exact location	Encode the approximate location as "Yes".
If the event location is a neighborhood and not in the gazetteer	Encode the location as the city and encode the approximate location as "No".
If an event location occurred on a mountain or another non-populated place and it is listed in the gazetteer	Encode the the non-populated place as a City and encode the approximate location as "No".
If the event location is unknown and coded as "NA"	Encode the approximate location as "No".
If an event details a source that states "in the vicinity" for a geography For example, "while clashes continued between the regime forces and militiamen loyal to them and	Encode the approximate location as "Yes".

| Hezbollah against the rebel and Islamic Factions, in the vicinity of al-Rihan town" | |

Actor (Required)

Each event must contain at least one actor, although multiple actors are possible. The actor is the group, organization, or person who initiates the MA event.

General Actor Encoding Guidelines:

- Encode this field in detail with a group name from the approved list of Actor names. Ensure that the spelling and diacritic marks conform to the approved list.
- If the Actor name is not on the approved list of Actors, propose it for submission.
- Do not encode individual names; determine the group with which the individual is associated. If that cannot be determined, enter "Unspecified" into the field and write a detailed comment. (Individual names are encoded in the Item of Note).
- Actors may emerge, merge, or disappear throughout the Challenge; therefore, the Actor dictionary will evolve over time.
- Note some Actors share the same names but have different identities based on location. In this case, encode as Actor-Location in conformance with the approved Actor list.

TABLE 21. ACTOR ENCODINGS

If the following...	Perform the following
If the actors are International coalition forces, Red Cross or UN	Do not encode the event.
If an article reports that an actor was "believed to be responsible"	Encode the actor.
If an article details the "Army's Air Force"	Encode the actor as "Military".
If an article details an event where the actor involved is a defected individual	Encode the target as the group they were previously identified by.
If an article details an event where the actor involved switched allegiances between groups	Encode the actor based on its last allegiance.
If an article details an event where the actor is detailed as Iraqi counter terrorism force and/or Iraqi anti-terrorism forces	Encode the actor as Iraqi Special Forces.
Syria Specific - If an article details an event where the attack is a barrel bombing	Encode the actor as Syrian Arab Military. The Syrian Arab Military are the only actors that utilize barrel bombs in the AOR.

For example, "the helicopters dropped 12 barrel bombs on areas in the city of Darayya"	
If an article details ministries as the actors For example, Russian Ministry of Defense, Ministry of Interior, etc	Encode the actor as Unspecified State Actor.
If an article details "anti-terrorism forces" as the actor with no country designation or affiliation.	Encode the actor as Unspecified State Actor.
If an article reports an event from January 1, 2017 to current where the actor is: Abu Al-Fadl Al-Abbas Forces; Al-Qassas Army; Ansar Allah Al-Awfiyya; Asa'Ib Ahl Al-Haq / AAH / Khazali Network; Badr Organization; Faylaq Al-Wa'Ad Al-Sadiq / The True Promise Corps; Harakat Al-Abdal; Harakat Al-Nujaba / Harakat Hezbollah Al-Nujaba; Jund Al-Imam; Kata'ib Al-Fateh Al-Mubiin; Kata'ib Al-Imam Ali; Kata'ib Al-Zahra; Kata'ib Dir' Al-Wilaya; Kata'ib Hezbollah / KH / Hezbollah Battalions; Kata'ib Malik Al-Ashtar / Malik Ashtar Brigades; Kata'ib Sayyid Al-Shuhada / Sayyid of Martyrs Battalions; Kata'ib Tayyar Al-Rasuli / Kata'ib Al-Risali; Liwa Ali Al-Akbar; Liwa Assad Allah Al-Ghalib; Martyr Sadr Forces / Quwet Shahid Al-Sadr; People's Mobilization / National Mobilization / Popular Mobilization Forces / Units / Committee / Hashd Al-Sha'abi / Hashid Shaabi; Saraya Al-Jihad; Saraya Al-Khorasani / Khorasani Brigades; Saraya Al-Salam / The Peace Brigades; Saraya Ansar Al-Aqeeda; Saraya Ashura; Kata'b Imad Mughniyeh, Iraqi Hezbollah-Kata'bs; Kata'ib alshaheed al-Awal Dawa Party Iraq Organization-Hashim al-Musawi; Kata'b of elite and Ghaith Alhaidara Dawa Party Home Organization Abdelkarim Anzi; Liwa al-Shabab al-Risali; Kata'b Ansar al-Marja'yya; Jaysh al-Mukhtar; Kata'ib Ansar al-Hijja; Kata'ib Qamar Bani Hashim; Hezbollah Tha'irun; Brigade of Qasim Al-jabareen; Liwa al-Imam al-Qa'im; Kata'ib A'imat al-Baqi; Liwa al-Munta'ar; Kata'ib tharullah; Kata'b Ashbal Al-Sadr; Kata'ib Thaer al-Hussein; Kata'ib Al-Dem Al-Zakeya; Liwa Dhu al-Fiqar; Kata'ib Muslim bin Aqeel; Liwa al-Imam al-Mahdi; Liwa al-Mu'ammal; Kata'ib al-Adalah (justice); Islamic Army in Iraq; Kata'ib Al-Ataba Al-Husayniya; Kata'ib al-Aqila Zainab; Liwa al-Tuff; Kata'ib Al-Imam al-Galib; Kata'ib Al-Imam Al-Hussein; Kata'ib Al-Husseini Al-qiam; Kataib Al-Qariah; Kata'ib Yed Allah (Hand of God); Kata'ib Bqiyat Allah; Kata'ib al-Shabab al-Islami; Kata'ib Ahl Al-Bayt; Saraya PDF; Kataib alttifl Al-radie; Kata'ib Mukhtar Althagafi; Saraya Al-Sajad; Kata'ib waad ullah; Kata'ib al-Ghawth al-Adham; Kata'ib of Babylon	Encode the actor as Iraqi Military. **Reason**: The Iraqi government formally recognized Hashd Al-Sha'abi and its 67 affiliates as part of the Iraqi Military on December 19, 2016. Per encoding change agreement, changes are to be made for whole months therefore events that occurred on January 1, 2017 to present will be changed to Iraqi Military. Anything Prior to January 1, 2017, will be encoded as the reported group.

If an article details an event with the actor as Saudi-led Arab Coalition operating outside of the AOR	Encode the actor as Royal Saudi Military.
If an article details the actor as Fifth Legion	Encode the actor as Syrian Arab Military

PROHIBITED LINKS

Below is a list of websites that cannot be included in the Additional Links text box.

- Twitter
- Facebook
- Wikimapia
- Wikipedia
- WikiLeaks
- Google Maps
- Reddit
- Maplandia
- Google Search Pages
- GeoNames
- Falling Rain
- Instagram
- YouTube

   *Facebook collection for Syrian Observatory of Human Rights is permitted as it is used for article collection when the main site for Syrian Observatory of Human Rights is inaccessible.

LOCATING GEOGRAPHIES

The following techniques must be used by researchers when searching for a location. To add or find a location in the gazetteer researchers should follow the steps provided below:

1) Enter and search the country, state and city into the appropriate text fields in the gazetteer.
   a. If this does not yield any results, enter only the city and state.
2) Search for the city using a fuzzy search (Enter a tilde [~] directly prior to the city name).
3) Researchers should use the following online resources to locate alternate spellings of the city name:
   a. www.google.com
   b. http://www.getty.edu/research/tools/vocabularies/tgn/index.html
   c. http://www.fallingrain.com/world/index.html
   d. http://geonames.nga.mil/namesviewer/
   e. http://wikimapia.org/#lang=en&lat=41.273000&lon=-72.957400&z=12&m=b
   f. http://www.geonames.org/
   g. http://www.findlatitudeandlongitude.com/
      http://freedomfalconofsyria.blogspot.com/2013/01/binggoogle-translations-of-towns.html
4) Search any new names of cities through alternate city names within the gazetteer.
5) Conduct a fuzzy search of alternate names within the gazetteer.

6) Search alternate city names with the Arabic spelling of locations within in the gazetteer.
7) Fuzzy search alternate city names with the Arabic spelling of locations within the gazetteer.

UPDATING EVENT INFORMATION

As new event information is discovered, researchers will continue to update the original event encoding. To update the event information, researchers will adjust the entry fields beginning with the event summary. Event summary additions will be encoded in the following format:
"Update from *Source Name* [*URL Snippet*]: *Updated information entered here*". For example, an update can be written as "Update from SOHR [1234]: Shelling by regime forces results in 2 ISIS members wounded." Following the updated event summary, change can then be made to reflect the updated event summary in the remaining entry fields.

V&V CHECKS

The following list provides the end-user with the checks that are currently performed by the V&V.

1) All field headers are present, and match the required names/format.
2) All coded actor, target, and GSS names match the standardized lists.
3) Listed Published Date, Accessed Date, and ERD are not before Event Date.
4)
4) Military Activity events have a SA as listed Actor.
5) There are no violent events coded outside the AoR.
6) All listed geonameid's exist in the standard gazetteer.
7) All location triplets (Country, State, City) exist in the gazetteer.
8) There are no blank cells in fields that should have no blanks.
9) Identifies potential duplicates
a. event type/date, geonameid, actor all match, event summaries >60% the same (based on unique words used in each).
10) Identifies potential duplicates based on actors/targets involved (grouped into a single list), event date, and event proximity (currently within a 10-mile radius).
11) Searches for potential coalition strikes in Iraq and Syria.
a. Armed Conflict event in Iraq and Syria, Unspecified SA as the actor, "coalition" in title or body of article.
12) Comparison between GSS and listed URL, making sure they match.
13) Checks additional links field for prohibited sites (Facebook, YouTube, Twitter, Reddit, etc.).
14) Checks for state-on-state events (SA as both actor and target).
15) Checks for anything not formatted as a URL in the additional links field.
16) Checks for the word "exact" in the IoN of any event where the city = NA, to generate a list of events for further review.

NON-VIOLENT CIVIL UNREST

This section presents guidance for the non-violent Civil Unrest (CU) analyst. CU analysts collect CU events in GSS articles, analyze relevant articles, identify all pertinent information, and enter that information into the CU Database. Because the GSR is the definitive record of CU in a region, the CU Database must be complete and accurate. Therefore, CU analysts must be methodical and thorough in searching for CU events and, take great care to ensure that events are encoded correctly.

<div align="center">DETAILED COLLECTION GUIDANCE FOR NON-VIOLENT CIVIL UNREST</div>

The process of finding relevant articles for CU is called "collection." Events are derived from pre-selected news sources; these sources are called Gold Standard Sources (GSS). Generally, both online and printed newspapers are used as GSS. Sources that are not typically GSR-appropriate include:

- Blogs, unless pre-approved by IARPA.
- Editorial articles.
- The comments sections on news articles.
- Anything from Facebook, Twitter, or other social media sites.

<div align="center">KEYWORDS</div>

Civil Unrest event collection involves the use of pre-selected search terms or "keywords." The following table provides the Arabic keywords approved for use in CU event collection. A tailored list of keywords specific to a region or country may be provided to the CU analyst as a research tool.

<div align="center">TABLE 22. ARABIC KEYWORDS</div>

Arabic Keywords
احتشاد حشود مشاغبات شغب ناشط نشطاء اعتراض احتجاجا احتجوا احتج معتصمين محتجين محتج احتجاجات الاحتجاجات احتجاج انتفاضة مضربين إضرابات إضراب اعتصام مسيرات مسيرة تمرد تظاهر متظاهرين متظاهر مظاهرات مظاهرة تجمهر

<div align="center">PERFORMING SEARCHES</div>

Google Advanced Search

Google Advanced Search is often the best method for analysts to collect events that are reported in online news media. Note that the search engines available on a news media site frequently miss relevant articles.

To perform event collection using Google Advanced Search, follow the steps listed below and use the set of Arabic keywords provided in Table 22 and the GSS listed below

Country	News Source	Website (for Google Advanced Search)	Nexis
Egypt	Al-Masry Al-Youm	almasryalyoum.com	X
	Al-Ahram	ahram.org.eg	X
	Al Wafd	alwafd.org	
	Youm7*	youm7.com	
Jordan	Ad-Dustour	addustour.com	X

	Al Rai	alrai.com	X
	Al Ghad	alghad.com	

*Youm7 is used as a GSS for CU in Egypt beginning September 2016.

Lexis-Nexis Search

Lexis-Nexis provides a search service for many "print" versions of the GSS's. For additional details on the Collection workflow, see section 0, Daily Operations.

DETAILED ENCODING GUIDANCE FOR NON-VIOLENT CIVIL UNREST

Analysts open the CU relational database to encode all CU events. Table 23 provides the name and description of all fields included in the database. Detailed guidance and examples are provided in the following sections, in the order that they appear in the table.

Detailed guidance and examples are provided in the following sections, in the order that they appear in the table.

TABLE 23. CU COMPLETE ENTRY FIELDS AND GUIDANCE

GSR Complete Entry Fields	Guidance
Country	Country where the event occurred.
State	State, province, region, department, or the next largest administrative division where the event occurred.
City	City where the event occurred.
Date of Event	Enter date the event occurred in Day, Month, Year format; e.g., 11 October 2012.
Earliest Reported Date	Publication date of first (or only) article about event. If there is no publication date associated with the article, the article may be not be used as the primary source for event encoding; however, it may be used for additional information.
GSS News Source	Name of the news source for the article with most complete event information.
Headline (Original Language)	Copy/Paste headline of article that reported the event.
Headline (English)	Translate headline of the article that reported the event.
Event Description	Brief extract from the article that best describes the event. Permitted to Copy/Paste only 2-3 sentences from the article. Leave the description in the original source language.
First Reported Link	URL link to article that first reported the event.
GSS Link	URL link to article with most complete event information. GSS Link and First Reported Link may be the same, but it is not a requirement.
Other_Links	URL links from other GSS articles that reported the event as available. Separate using semi-colons with a space on either side (" ; ").
Encoding_Comment	Optional. Can be used for notes for the QC analyst or GSR manager that clarify encoding actions or provide details about the event that are deemed important.

Detailed guidelines and clarifying examples to encode each Collection field are presented in the following sections.

LOCATION

Location is determined in three levels: Country, State (also known as Province, Department, etc.), and City. To ensure the consistency and standardization of place or locations names, only use the location information that is provided in a gazetteer. Always Copy/Paste location names directly from the gazetteer.

Be aware of the following situations or special cases when encoding event locations:

- An article may refer to a place or location by an abbreviated name or using a colloquial form; for example, "Los Angeles" is commonly referred to as "LA".

- Spelling of place or location names may vary.

- If the article does not specify the city where an event occurs, but states that the protest took place in front of a particular building or street, the analyst should use Google to verify the location of the building or street and encode Location accordingly. Note: If the location cannot be determined within five minutes, the field is left blank and the analyst makes a note in the *Encoding Comment* field so it can be reviewed by another analyst or during the quality control process.

- If the location is not in the gazetteer, encode the next administrative level up and leave a note in the *Encoding Comments*.

- The same protest may take place in more than one location, e.g., bakers go on strike in two different cities on the same day. Encode this as two discrete events.

- If a protest moves from one city to another, it is encoded as one event. The location is the city where the protest began.

- Embassy grounds are treated as being internal to the country in which they are located.

VIOLENCE

For the Mercury Challenge, violent civil unrest events will not be documented in the GSR. Violence is a physical confrontation between people or damage done to property. For example, a clash with police or a clash between two opposing groups that results in injuries or significant destruction of property (e.g., burning cars, looting) is encoded as Violent. Injuries not directly connected to the event, such as heat exhaustion or heart attack during the protest, is considered Non-Violent. The following table contains examples and guidance for Violence indicators:

TABLE 24. VIOLENCE INDICATORS

Violent	Non-Violent
• Clashes with police, e.g., police using teargas or high pressure water hoses to disrupt a protest • Clashes between opposing groups resulting in injuries • Pushing or being forcibly removed from an area • Hostage taking	• Threats of violence, e.g., yelling, cursing • Police arresting protestors
• Self-inflicted wounds, e.g., protesters sewing their mouths shut	• Hunger strikes
• Throwing hard objects which could cause injury or damage • Hitting with clubs, machetes, etc.	• Throwing things that would not cause harm, e.g., eggs • Brandishing clubs, machetes, etc.
• Burning tires, burning barricades, burning cars or buses, burning buildings	• Fireworks • Blockades of streets
• Looting shops (where the shop is damaged) • Defacement of property, e.g., painting graffiti on buildings	• Theft without damage

DATE OF EVENT

Articles typically indicate the exact date for most Civil Unrest event occurrences, but they must be read thoroughly to determine the correct date. Enter the date when the event occurred, in the *Date of Event* field.

Other considerations for encoding the event date include:

- If an online article mentions that the event took place that same day, it is important to check the hour of publication. Articles with a publish hour of 12:00 am - 5:00 am are most likely referring to events that occurred the day before.

- If the article does not mention when the event took place and there is no time stamp, encode the event date as the date of publication.

- In the case where an article does not specify an event date, but provides a hint (e.g., "yesterday" "three days ago" or "last week"): Count backward from the published date of the article to determine the event date.

- If a news report is undated and no date information is provided, leave the *Date of Event* field blank and enter a note in the *Encoding Comment* field for review by the Quality Assurance team.

- Most of the media sources used for Civil Unrest are online. This enables a reporter to post an article on the same day as the event occurred. For print media sources such as Nexis, a reported event may have occurred prior to the date of publication.

EARLIEST REPORTED DATE

Encode the publication date for the first article that reported the CU event. Dates are entered using the Day, Month, Year format such as 11 October 2012. Note, if the article does not include a publication date, it may not be used as the primary news source for event encoding.

GSS NEWS SOURCE

Record the name of the GSS news source used to collect the CU event. If there are multiple sources that reported the same event, use the news source with the most complete event information.

HEADLINES

In the CU GSR Spreadsheet, there are two fields for entering article headlines: *Headline (Original Language)* and *Headline (English)*.

- For Headline (Original Language) field: Copy/Paste the exact headline in the original language of the article.

- For Headline (English) field: Enter the English translation of the article headline. Make sure to spell check the translation.

EVENT DESCRIPTION

Copy/Paste information from the news article that helps to explain the CU event. Typically, this information can be extracted from the first or second paragraph of the article. For example: "A group of beauticians marched Thursday to protest the high cost of electricity and frequent power cuts. The protest became violent after one of the beauticians attacked a police officer."

Enter the information in the *Event Description* field in the original language of the source article.

FIRST REPORTED LINK

Copy/Paste the URL link from the GSS news source used to collect the CU event, into the *First Reported Link* field in the GSR spreadsheet. Other considerations include:

- If the article is a print article with no URL (e.g., Nexis), conduct a Google search using the Headline to find an online version of the article and encode the URL link.

- If a URL link cannot be found, leave the column blank and enter a comment in the *Encoding Comment* field.

GSS LINK

Copy/Paste the URL link from the GSS news source that provides the most complete event information, into the *GSS Link* field in the GSR spreadsheet. Other considerations include:

- The *GSS Link* may or may not be the same as the *First Report Link*. Typically, these links are the same if the event is reported the same day the event occurred.

- If the article is a print article with no URL (e.g., Nexis), conduct a Google search using the Headline to find an online version of the article and encode the URL link.

- If a URL link cannot be found, leave the column blank and enter a comment in the *Encoding Comment* field.

OTHER LINKS

The *Other Links* field is used to document URLs from other collected articles that report the same CU event. Copy/Paste URL links directly from the article into the GSR spreadsheet. Separate the links with a semicolon and include a space on either side of the semicolon. For example: www.otherlink1.com ; www.otherlink2.com ; www.otherlink3.com.

EVENT BOUNDARIES AND CONTINUOUS EVENTS

WHAT CONSTITUTES THE BOUNDARIES OF AN EVENT?

The temporal granularity of an event is one day. Most events take place on a particular day. Generally, people taking part in a protest march or standing on the street holding protest signs go home at the end of the day, possibly to return the next. In contrast, if a building were "occupied" for several days, it would be unlikely that the occupation would end one day and resume the next.

- An event starts stops and starts again within the same day; it is still the same event.

- An event that moves from one location to another within the same day would still be the same event.

- If an event evolves over the course of the day, such as a march leading to a building that is then seized by the protestors, it is still considered one event.

- An event changes from non-violent to violent the same day, it is one event and should not be encoded since it was violent.

- An event that is continuous over the course of multiple days is encoded as the date the protest began.

- Note that unless the article specifically mentions that the protest continued through the night or that the protesters set up camp, assume that they stopped at night and came back the next day. In this case, the analyst encodes one event for each day of protest.

DISTINGUISHING CONTINUOUS EVENTS FROM MULTIPLE EVENTS

An event that continues for several days can be encoded as a single event or multiple events, according to the following criteria:

- An event that spans multiple days, is continuous, and pertains to the same event reason (e.g., a strike) is encoded as a single event. In this case, protestors will typically occupy a space or location. They do not stop protest activities at the end of the day, and then resume the next.

- An event that stops and resumes again on another day, is encoded as a separate event.

- A new population group joins the continuous protest; this is a new event for the new population where the event date is the day the new population group joined the continuous protest.

- An event that occurs with no violence, then becomes violent after the first day. This event is really two events; the first for the period of days the protest was non-violent, and a second event is not recorded since it became violent.

- If a counter protest begins after the first day of an event, encode the counter protest as a separate event.

  For example: There is a student protest that is ongoing for several days. The protesters are blocking the entrance to their university and classes are suspended. On the 4th day of the protest another group of students (wishing to resume classes) start to "counter protest" against the original group blocking the school entrance. Since the counter protest started on a different day than the original protest, a separate event is created. Note: if the counter protest starts on the same day as the original event, a separate event is not encoded.

- A continuous event is punctuated by specific acts of protest. For example, workers occupy the town square for an extended period in protest for better wages. This is encoded as one event where the event date is the day the occupation began. However, during the occupation, workers stage a march to block traffic. The march is a specific act of protest that is distinct from the occupation and is encoded separately.

CU EXAMPLES OF NON-EVENTS

The following examples are "grey-areas" where there was disagreement as to whether the incident in question would be considered an Event under Mercury Challenge guidelines. These examples are grouped into categories of incidents that are *not* considered events.

1. State Funerals where there are no organized protest events.

Translation:	Original:
Chief of the Armed Forces, His Majesty Advisor for Tribal Affairs Sharif Fawaz Zabin Abdullah participated in the presence of Defense Minister Dr. Abdullah eagles Prime Minister, the King's Advisor for Military Affairs Chairman of the Joint Chiefs of Staff Gen. Gen. Mashal Mohammed Zabin on Saturday, at the funeral of the former chairman of the Joint Chiefs of Staff, Field Marshal Gen. Abdel Hafez Ka'abneh Merhi, who moved to the mercy of God, on Friday, March 25, 2016.	عمان ـ بترا ـ مندوباً عن جلالة الملك عبدالله الثاني، القائد الأعلى للقوات المسلحة، شارك مستشار جلالته لشؤون العشائر الشريف فواز زبن عبدالله، بحضور رئيس الوزراء وزير الدفاع الدكتور عبدالله النسور، ومستشار جلالة الملك للشؤون العسكرية رئيس هيئة الأركان المشتركة الفريق أول الركن مشعل محمد الزبن امس السبت، في تشييع جثمان رئيس هيئة الأركان المشتركة الأسبق المشير الركن عبدالحافظ مرعي الكعابنة، الذي انتقل إلى رحمة الله تعالى، يوم الجمعة 25 آذار 2016.

DISEASE

The Disease analyst is responsible for collecting, encoding, and conducting quality assurance (QC) review of Disease ground truth data for the Gold Standard Report (GSR). The Disease analyst compiles encoded and QC'd events into a Disease GSR Spreadsheet. The disease analyst must be methodical in researching and encoding events. Attention to detail and accuracy is a priority. Appendix D contains references to detailed Disease training materials.

DISEASE COLLECTION

The analyst performs weekly event searches within select Gold Standard Sources (GSS) for all disease types in the eight Mercury countries. Searches are restricted to GSS sites to ensure continuity and consistency of the encoded data. Disease events are encoded as either Rare Disease Events or Weekly Case Count Events. Each category requires a different encoding strategy and utilizes the GSS resources in a prescribed manner.

All diseased events must be confirmed by an authoritative source such as:

- A news article citing confirmation of a positive test result from a Health Ministry or hospital official.
- Announcement of a confirmed case from a country's health website.
- Report of a confirmed case on an official website; e.g., EMPRES-i, HealthMap

Event collection is complete when the analyst has searched all appropriate sources and has thoroughly read all pertinent articles.

DISEASE GOLD STANDARD SOURCES (GSS)

Pre-selected news sources called Gold Standard Sources (GSS) are used to collect and research Disease events. The primary sources include EMPRES-i (the Global Animal Disease Information System), and HealthMap. Google's Advanced Search tool is also used as a back-up for further research of events. The following table lists Disease GSS.

TABLE 25. GSS FOR DISEASE

GSS for Disease
EMPRES-I Global Animal Disease Information System URL: http://empres-i.fao.org/eipws3g/#h=1 Maintained by the United Nations' Food and Agriculture Organization (FAO)
HealthMap URL: http://www.healthmap.org/en/index.php Information acquired from ProMed, WHO, GeoSentinel, OIE: World Organization for Animal Health, FAO, and news articles.
Google Advanced Search URL: http://www.google.com/advanced_search

EMPRES-I

EMPRES-i is a web-based application maintained by the Food and Agriculture Organization of the United Nations. It is designed to provide timely and reliable disease information that provides early warning transboundary and high impact animal diseases. Detailed searches may be conducted for diseases of interest; searches may be limited to "human" disease cases.

EMPRES-i—Global Animal Disease Information System is the preferred primary news source for collecting Avian Influenza and MERS disease events in Saudi Arabia. Detailed instructions are presented in the detailed Disease analyst training material, see Appendix D for the reference.

HEALTHMAP

HealthMap is a website that presents a unified view of the current state of infectious diseases worldwide, through the aggregation of publically available information from: ProMED, World Health Organization, GeoSentinel, the Food and Agriculture Organization (FAO) of the United Nations, and other news resources.

HealthMap is a preferred primary news source for collecting MERS in non-Saudi Arabia countries as well as Dengue in all countries. Healthmap is the preferred back-up source for Avian Influenza, and MERS in Saudi Arabia. Detailed instructions are presented in the detailed Disease analyst training material, see Appendix D for the reference.

GOOGLE ADVANCED SEARCH

Google Advanced Search is a robust search engine that supports very detailed search requests of publically available information. It is the preferred back-up research tool for all Disease events. Detailed instructions are presented in the detailed Disease analyst training material, see Appendix D for the reference.

The following table summarizes all selected news sources for Disease.

TABLE 26. GSS SUMMARY BY DISEASE TYPE

Disease	Primary Sources	Backup Sources
MERS in Saudi Arabia	EMPRES-i / HealthMap	Google Advanced Search

DISEASE ENCODING

WEEKLY CASE COUNT

As information is available, weekly case counts are assembled for each EW. Weekly case count data is available for the following:

- MERS in Saudi Arabia

The Mercury Challenge conforms to the CDC's definition of epidemiological week (EW). Each EW begins on a Sunday and ends on the following Saturday.

GSR COMPLETE ENTRY SPREADSHEET FOR DISEASE

The GSR complete entry spreadsheet for Disease is used for entering encoded Disease events. The following table provides a list of all fields included in the spreadsheet, including a brief explanation of how each field is completed. Detailed guidance pertaining to each field is provided in the following sections, appearing in the order listed below.

TABLE 27. DISEASE GSR FIELDS

GSR Complete Entry Fields	Optional or Required	Description
Event_ID	Required	Unique identifier for each GSR event.

GSR Complete Entry Fields	Optional or Required	Description
Event_Type	Required	Rare Disease or Weekly Case Count
Country	Required	Country where the disease was confirmed. Location information must conform to a gazetteer value.
State	Rare Disease only	State where the disease was confirmed. Location information must conform to a gazetteer value.
City	Rare Disease only	City where the disease was confirmed. Location information must conform to a gazetteer value.
Approximate_Location	Optional	Mark location as Approximate if: 1) the location is not in a gazetteer value and the closest gazetteer City location is used, or 2) if the location is described in relation to a city in a gazetteer value.
Latitude	Optional	Latitude for the City is provided in a gazetteer value.
Longitude	Optional	Latitude for the City is provided in a gazetteer value.
Disease	Required	The name of the disease.
Case_Count	Weekly Case Count only	This field contains the number of confirmed Weekly Case Count cases.
Event_Date	Required	The date of disease onset for Rare Disease. The EW for Weekly Case Count events.
Earliest_Reported_Date	Required	The publication date of the first official confirmation of the Rare Disease case; encode in yyyy-mm-dd format.
News_Source	Required	Name of primary news source that reported the disease event.
First_Reported_Link	Required	URL of the GSS that first reported the Disease case confirmation.
GSS_Link	Required	URL of the GSS that best describes the Disease event.
Other_Links	Optional	URL from other supporting news sources, if available. Separate multiple URLs with a semi-colon.
Revision_Date	Required	Date of most recent update to the Event ID.
Encoding_Comment	Optional	May contain information that clarifies the encoded event.

Event ID

This is a unique number assigned to individual Disease events listed in the GSR.

Location Fields

For Rare Disease, encode locations based on where the disease was confirmed to have occurred, not where it was believed to have been contracted. If the news source for the event does not specify location information, the analyst should attempt to determine the location by conducting a search using a backup news source. Rare Disease events are encoded at the City-level. Spelling and diacritic marks must conform to gazetteer entries. Copy/Paste location information directly from the gazetteer into the spreadsheet to ensure consistency and accuracy.

Weekly Case Count Disease is encoded at the country-level only.

Case Count

For events that fall in the Weekly Case Count category, enter the number of confirmed cases that occurred within the EW. If no cases occurred, enter "0" in the Population field.

Event Date

For Weekly Case Count Disease the event date is the first day of the reported EW.

All event dates are encoded in year-month-day format, using a 4-digit year.

Earliest Reported Date (ERD)

For Weekly Case Count Disease, the ERD is the publication date of the first official confirmation of the reported EW cases. If the ERD cannot be determined, encode the first day of the associated EW.

All event dates are encoded in year-month-day format, using a 4-digit year.

News Source

Enter the name of the GSS news source used to collect the Disease event. If there are multiple sources that report the same event, use the news source with the most complete event information. Add additional event information from the other news sources if it adds to the completeness of the encoded event.

First Reported Link

Copy/Paste the URL link from the GSS article that first reports the Disease event.

GSS Link

Copy/Paste the URL link from the GSS news source that provides the most complete description of the Disease event.

Other Links

The Other Links field is used to document URLs from other news sources that provide additional information regarding the Disease event. If more than one link is added, separate the links with a semicolon.

Encoding Comment

Comments are used to explain encoding decisions and point out encoding omissions. Detailed comments are encouraged.

ACRONYMS AND KEY DEFINITIONS

CU	Non-violent Civil Unrest
RD	Earliest Reporting Date
EW	Epidemiological Week or "epiweek"
Event class	Broad category of a societal event; e.g., Military Activity or Disease.
Event subtype	Specific event category within an event type; e.g., NSA – Bombing
Gazetteer	A geographical dictionary or directory.
GDB	GeoNames database. GeoNames.org maintains a geographical database or gazetteer that is available for download.
GSR	Gold Standard Reference (report containing an encoded-event data set).
GSS	Gold Standard Source
IARPA	Intelligence Advanced Research Projects Activity
Lat-Long	Latitude and Longitude coordinates
Location triplet	Related Country/State/City names. Used when identifying an event location.
LS	Location Score
LT	Lead Time
MA	Military Activity
MENA	Middle East and North Africa
MERS	Middle East Respiratory Syndrome
MoH	Ministry of Health
Nexis	Authoritative news source product provided by the LexisNexis corporation
NGA	National Geospatial-Intelligence Agency
NSA	Non-State Actor
QA	Quality Assurance
QS	Quality Score
Participant	Research team responsible for developing automated predictive models for societal events.
ProMED	Program for Monitoring Emerging Diseases
RD	Rare disease
SME	Subject Matter Expert
T&E	Test and Evaluation
V&V	Verification and Validation
W-E	Warning-event
Warning	Participant event forecast

WHO   World Health Organization