# IARPA MERCURY CHALLENGE HANDBOOK



Last update:  06/25/2018

[Type here]

# TABLE OF CONTENTS

[Type here]

# LIST OF TABLES

# INTRODUCTION

Surprise events such as the fall of the Berlin Wall, Iraq's invasion of Kuwait, the civil unrest that gave rise to the Arab Spring, and Russian incursions into Ukraine, forced the U.S. government to respond rapidly, often in an absence of data related to the underlying causes of these events.

In an effort to provide early warning capabilities, the Department of Defense's Integrated Crisis Early Warning System (ICEWS) and IARPA's Open Source Indicators (OSI) programs leveraged novel statistical and machine learning techniques using publicly available data sources to forecast societal events such as civil unrest and disease outbreaks with a high degree of accuracy. The IARPA Mercury Challenge is looking for novel and advanced methods to provide early warning for the U.S. Government of such events.

The Mercury Challenge seeks innovative solutions and methods for the automated generation of event forecasts in the Middle East.

The specific event classes of interest are:

- Military Activity (MA) in Egypt, Saudi Arabia, Iraq, Syria, Qatar, Lebanon, Jordan, and Bahrain:
    - Conflict – Incident where police, military, or other state/government security forces take action in some way; and
    - Force posture – A newsworthy action of police, military, or other state/government security forces that does not involve the use of deadly force.

- Non-violent Civil Unrest (CU) in Egypt and Jordan, such as demonstrations, marches and protests:
    - Daily count of non-violent civil unrest events in Egypt;
    - Weekly count of non-violent civil unrest events in Jordan;
    - Weekly count of non-violent civil unrest events in a 75km radius of Tahrir Square, Egypt;
    - Monthly count of non-violent civil unrest events in the Jordanian provinces of Amman, Madaba, and Irbid.

- Infectious disease in Saudi Arabia: Weekly Middle East Respitory Syndrome (MERS) count.

Please note any references to an Appendix can be found in the supporting document "Mercury Challenge Handbook Appendices."

## JUDGING CRITERIA

The Mercury Challenge will compare participant submissions against a "base rate" model. Base Rate models (See Appendix) are models that only use information included in the history of observed events. It is expected that Participant models will score better than the base rate models. Participants are encouraged to develop and test innovative forecasting methods that ingest and process publicly available data sources to predict events in each of the preceding classes.

## MILITARY ACTIVITY (MA) JUDGING METRICS

For MA forecasts, the effectiveness of each participant's methods will be judged using the following metrics:

- *Lead Time (LT)*:  Number of days between the date the forecast was produced and the date the actual event was reported;
- *F-Score*: The harmonic mean of Precision and Recall

$$F \equiv \frac{2 \times Precision \times Recall}{Precision + Recall}$$

  - *Recall*:  The number of forecasts that matched actual events divided by the total number of actual events;
  - *Precision*:  The number of forecasts that matched actual events divided by the total number of forecasts issued;
- *Quality Score (QS)*:  The similarity of warning details to event details in terms of the distance between the warning location and the event location, the number of days between the warning Event Date and the actual Event Date, and agreement between warning and event actor and event subtype.  QS is measured on a scale of 0.0 to 1.0.  For more details see the Appendix.

To be eligible for a prize, participants must achieve the following 3 thresholds:

- Participant's system Mean Lead Time  >= 3 days

- Participant's system F-Score >= base rate's mean F-Score + 0.5 standard deviation of base rate F-Score, where the mean and standard deviations are computed using 100 simulations of the Base Rate model.  See the Appendix for more details.

- Participant's system Quality Score >= base rate's mean Quality Score + 0.5 standard deviation of base rate Quality Score

Ranking of participants who achieve these 3 thresholds will be done using a combined metric known as the Mercury Score:

$$Mercury\ Score = (F + QS)/2$$

Mercury Score will be carried to three significant digits.  In the event of a tie. additional significant digits will be used to determine the final winners.

## COUNT FORECAST (CU EVENTS AND DISEASE) JUDGING METRICS

For count forecasts (CU events and disease), the effectiveness of each participant's methods will be judged using:

- *Lead Time*:  Average number of days between the date the forecasted count was submitted and the effective Event Date, defined as:
  - Egypt Daily CU:  The date being forecasted.
  - Weekly Counts:  The week is defined as the International Organization for Standardization (ISO) week, which starts on Monday and ends on Sunday.  Weekly CU counts use Wednesday as the reference day for determining the Event Date.  Disease events follow Epidemiological Week (EW) which uses Sunday as the reference day.  For

example, consider the week that starts on 28 May 2018 and ends on 3 June 2018. The Event Date for CU counts is 30 May 2018 and the Event Date for Disease events is 3 June 2018.

- o Monthly Counts: The effective Event Date is the 15th of the month.
- *Quality Score*: Average quality score of each valid forecast (ranges from 0 to 1), which is based on the difference between the forecast count and the actual count.

To be eligible for a prize, participants must achieve the following 2 thresholds:

- Timely submission of warnings, defined as:

  - o For Egypt Daily CU, Warning Lead Time >= 3 days.

  - o For Weekly CU Counts, the warning must be submitted before 21:00:00 UTC on the Sunday before the start of the week being forecast.

  - o For Weekly Disease Counts, the warning must be submitted before 21:00:00 UTC on the Saturday before the start of the EW being forecast.

  - o For Monthly Counts, the warning must be submitted before 21:00:00 UTC on the 7th day of the month being forecast.

- Participant's system Quality Score > base rate's Quality Score

- Participants are expected to offer a case count for each applicable time period (day, week or month) within the challenge window. If a participant does not submit a Case Count value for a given time period, the scoring system will assign:

  - o The participant's submitted value from the preceding time period if the participant has ever submitted on this Country/Event Type

  - o A count of "0" if if the participant has never submitted on this Country/Event Type

Ranking of participants who achieve these 2 thresholds will be done using Quality Score carried to three significant digits. In the event of a tie additional significant digits will be used to determine the final winners.

# GOLD STANDARD REPORTS (GSR)

IARPA is currently generating a large volume of data in the form of Gold Standard Reports (GSRs). These data sets contain details on more than 120,000 significant events in the region of interest for the prior 24 months. The GSR is updated every two weeks with events from the preceeding weeks. These data will be made available to participants and used for training their algorithms and testing their forecasts.

To conclude this introduction, as you read further you'll learn more about the following: 1) the development of the GSR for the Mercury Challenge, 2) the submission and evaluation of Participant warnings, and 3) supplementary materials used in the Mercury Challenge.

## MERCURY CHALLENGE EVENT CLASSIFICATION

The following event classes are included in the Mercury Challenge:

TABLE 1. MERCURY CHALLENGE EVENT CLASSES

| Event Class | Event Types | Geographic Focus |
|---|---|---|
| Military Activity | Conflict | Egypt, Saudi Arabia, Iraq, Syria, Qatar, Lebanon, Jordan, and Bahrain |
| | Force Posture | |
| Non-violent Civil Unrest | Daily Count | Egypt |
| | Weekly Count | Jordan and the region within a 75km radius of Tahrir Square, Egypt |
| | Monthly Count | Jordanian provinces of Amman, Madaba, and Irbid |
| Disease | Weekly Count of MERS | Saudi Arabia |

The challenge team will provide a detailed set of information for each event class, which will allow participants to correctly encode and score each event class.

1. Definition—definition and general description of the event class.
2. Geographic focus—the associated geographic area of interest. Note that different event classes have different geographic focus.
3. Gold Standard Sources (GSS)—a discussion of the news sources used to develop GSR events.
4. GSR Fields—a definition for each of the fields in a GSR event record.
5. Warning Format—a definition of all fields in a warning record.

Please note that the handbook presents this information in the same order and format in each section as the information is specific to each event class. In addition, there are two processes common to each event class: Warning Ingest and the management of Warning Updates.

## SUBMITTING A WARNING

Participants submit warnings to the Mercury Challenge ingest system. All warnings will be submitted in JSON format and must include a unique Participant ID for the warning.  The top level fields for the JSON submission are "*participant_id*" and "*payload*", where the payload is a list of individual warnings.  This participant ID must be consistent with your Topcoder user ID name. All participants must have a Topcoder ID in order to authenticate into the ingest system. Upon submission, the ingest system will perform validity checks, timestamp the warning, increment the update number (if applicable) and return a JSON-formatted acknowledgement message to the participant system.

Warnings that do not conform to the format described above are rejected by the warning ingest process. Participants may resubmit rejected warnings, but resubmissions will receive a new date/time-stamp.

Warnings must be submitted at least one day before the earliest reported date (ERD); otherwise, the warning is considered late and will not be scored. The time zone of all GSR ERDs will be in Coordinated Universal Time UTC+3. The challenge team defines "one day before" as at least 24 hours or more before the ERD in the local time zone of the event.

For example, if the ERD of an event occurs on April 25, 2018 (UTC +3), only warnings submitted on or before 23:59 (UTC+3) on April 24, 2018 will be considered on time.

Participants may not submit more than 200 Military Activity warnings per country with the same Event Date.  Warnings in excess of this limit will be rejected.

## UPDATING A SUBMITTED WARNING

As more information becomes available, participants may choose to update a previously submitted warning. The set of updates to a single warning is called an update sequence. Participants may update their warnings by sending an updated warning with the same Participant Warning ID.  Participants should provide all required elements for the warning. The last version of each warning will be used in the matching process.  Metrics for a warning sequence will be computed as the average of the metrics of each member of the sequence. Not all updates are permitted, see the Appendix for a listing of which updates are permitted.

# MILITARY ACTIVITY EVENTS

## MA DEFINITION

Military Activity (MA) events are actions by any state-sanctioned military, police, or security organization.

MA events are characterized by the following:

- They are "local" events and occur in a (Country, State, City) location triplet.  For these events, "city" is a third-level location descriptor as defined in the on-line gazetteer at http://geonames.nga.mil/gns/html/index.html.  A "city" may be a populated place, country region, body of water, or other location.
- They are instigated by an Actor.
- They may result in casualties.

MA events must be reported by a GSS to merit inclusion. This includes but is not limited to GSS reports in written text, in text describing a photo, and in text describing a video. The release of photography, videography, and supporting text must be reported by a GSS, and include pertinent event information in the text as solely the text will be captured.

MA events are further broken in to subtypes as defined in the following table.

TABLE 2. SUMMARY OF MILITARY ACTIVITY EVENT DESCRIPTIONS

| Event Type/Subtype | Event Subtype Description |
|---|---|
| Conflict | Incident where police, military, or other state/government security forces take action in some way. |
| Force Posture | A newsworthy action of police, military, or other state/government security forces that does not involve the use of deadly force. |

## MA GEOGRAPHIC FOCUS

MA events are encoded for Egypt, Saudi Arabia, Iraq, Syria, Qatar, Lebanon, Jordan, and Bahrain. Events initiated by a Military Actor in a country outside this list are also encoded. For example, a military action initiated by Saudi Arabia inside of Yemen would be encoded with Yemen as the Country.

## MA GOLD STANDARD SOURCES (GSS)

All MA events are collected from local and regional news sources.  Event collection techniques include Google Advanced Search (limited to the newspaper website), Nexis queries, and IHS Janes. Google Advanced Search is used to collect events in online media. Nexis and IHS Janes are used to collect events in print media.

The following table lists GSS by country.

TABLE 3. MA GOLD STANDARD SOURCES

| Country | Source Name | URL |
|---|---|---|
| Bahrain | Gulf News | gulfnews.com |
| Bahrain | Arab News | arabnews.com |
| Bahrain | Bahrain News Agency | bna.bh |
| Egypt | Almasry Alyoum | almasryalyoum.com |
| Egypt | Egypt Independent | egyptindependent.com |
| Egypt | Daily News Egypt | dailynewsegypt.com |
| Egypt | Shorouk News | shorouknews.com |
| Iraq | NINA News | ninanews.com |
| Iraq | Bas News | basnews.com |
| Iraq | Iraqi News | Iraqinews.com |
| Jordan | Jordan Times | jordantimes.com |
| Jordan | Albalad News | albaladnews.net |
| Jordan | Al Bawaba | albawaba.com |
| Lebanon | NNA | nna-leb.gov.lb |
| Lebanon | El Nashra | elnashra.com |
| Lebanon | Al Bawaba | albawaba.com |
| Qatar | Qatar Daily Star | qatardailystar.com*** |
| Qatar | The Peninsula Qatar | thepeninsulaqatar.com |
| Qatar | Al Arab | alarab.qa |
| Qatar | Doha News | dohanews.co**** |
| Saudi Arabia | El Nashra | elnashra.com |
| Saudi Arabia | Al Bawaba | albawaba.com |
| Saudi Arabia | Shia Post | shiapost.com |
| Saudi Arabia | Al Riyadh | alriyadh.com |
| Syria | SANA | sana.sy |
| Syria | Syrian Observatory for Human Rights | syriahr.com* |
| Syria | Al-Masdar | almasdarnews.com** |
| Syria | LLC | lcc-sy.com** |

*During denials of service on the syriahr.com site, articles may be collected from the corresponding Facebook site, facebook.com/syriahroe, beginning June 2016.

**Due to frequent denials of service on syriahr.com, collection of almasdarnews.com articles began in June 2016. Collection of LLC began August 2016.

***Qatar Daily Star was discontinued as of September 2016.

****Doha News was added as of July 2017.

## MA GSR FIELDS

GSR fields for MA are defined as follows:

TABLE 4. MA GSR FIELDS

| GSR Field | Optional or Required | Description |
|---|---|---|
| Event_ID | Required | All GSR records are assigned a Unique identifier starting with the letters "MN" |
| Event_Type | Required | The MA event type is "Military Activity". |
| Event_Date | Required | The date the event occurred or, in the case of a multi-day event, the date that the event started. |
| Earliest_Reported_Date | Required | The Earliest Reported Date (ERD) is the publication date for the first article that reported the MA event. |
| Country | Required | Country where the MA event occurred. Location information must conform to a gazetteer value. |
| State | Required | State where the event occurred. Location information must conform to a gazetteer value. |
| City | Required | City where the event occurred. Location information must conform to a gazetteer value. |
| Approximate_Location | Optional | Possible values are "True" and "False." Location is Approximate if: 1) the location is not in a gazetteer and the closest GDB City location is used, or 2) if the location is described in relation to a city in a gazetteer value. An event is flagged Approximate Location if the exact location of the event is unknown, or if the exact location if not available in a gazetteer value. The following location information would indicate the use of the Approximate Location flag when no other location information for an event is available: <br><br> • near Riyadh <br> • on the border between Latakia and Turkey <br> • the road linking Sijariya and Tel Mushaihida <br> • west of Saqalawiyah <br><br> More detailed information on Approximate Location is provided in the Appendix. |
| Latitude | Optional | Latitude for the City as provided in the gazetteer. |
| Longitude | Optional | Longitude for the City as provided in the gazetteer. |
| Event_Subtype | Required | The Event Subtype is used to describe the event that occurred. The event subtype is determined through the Actor. The following table provides a detailed description of each event type. Excerpts from articles detailing each event type can be found in the Appendix. <br><br> Use "Conflict" if: <br> Incident where a State Actor (police, military, or other sanctioned state/government security forces) initiates an encounter or takes action. <br><br> Examples include: airstrikes, engaging in combat, hostage rescues involving force, drone strikes, raids (includes arrests). <br> Cannot be conducted by a coalition including the United States, Great Britain, Canada, Australia, New Zealand or Israel. |

| GSR Field | Optional or Required | Description |
|---|---|---|
|  |  | Use "Force Posture" if: Incident where a State Actor (police, military, or other state/government security forces) **DOES NOT** use force and/or are not met with resistance. Examples include: 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. |
| **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, and must be a State Actor (SA). See separate Actor Tables for the list of MA actors, available in the Appendix. |
| **News_Source** | Required | Name of primary GSS. See the Appendix for a list of approved GSS. |
| **First_Reported_Link** | Required | URL link from the GSS news source used to collect the MA event. |
| **GSS_Link** | Required | URL link from the GSS news source that provides the most complete event information. |
| **Other_Links** | Optional | URL for additional news reports. Multiple URLs may be encoded in this field; separate URLs with a semi-colon. |
| **Revision_Date** | Required | Date of most recent update to the Event ID. |

### SAMPLE MA EVENT IN JSON FORMAT:

```
{
    "Event_ID": "MN35501",
    "Event_Type": "Military Activity",
    "Event_Date": "2016-03-01",
    "Earliest_Reported_Date": "2016-03-01",
    "Country": "Iraq",
    "State": "Anbar",
    "City": "Al Fallūjah",
    "Approximate_Location": "True",
    "Latitude": 33.3491,
    "Longitude": 43.786,
    "Event_Subtype": "Conflict",
    "Actor": "Iraqi Military",
    "News_Source": "NINA News",
    "First_Reported_Link":
"https://www.ninanews.com/News_Details.aspx?0p88GUyXgTrqUXokWgwKuQ%253d%253d"
,
    "GSS_Link":
"https://www.ninanews.com/News_Details.aspx?0p88GUyXgTrqUXokWgwKuQ%253d%253d"
,
    "Other_Links": "",
    "Revision_Date": "2016-08-17",
},
```

## MA WARNING FORMAT

The following table provides descriptions of the warning fields required for MA warning submissions.

TABLE 5. MA WARNING FIELDS

| Warning Fields | Descriptions |
|---|---|
| Warning_ID | The Warning ID is a unique alphanumeric string provided by the challenge participant. To update a warning, submit a revised warning with the same Warning ID. The Mercury Challenge scoring system will automatically create a sequence number for the updated warning. |
| Actor | See the Actor Tables for the lists of MA actors, available in the Appendix. Each warning must contain exactly one Actor. |
| Event_Type | Military Activity |
| Event_Subtype | Conflict or Force Posture |
| Event_Date | The predicted event date. Encode in yyyy-mm-dd format. |
| Country | The country where the event will occur. Each field is populated with a location from a gazetteer value. |
| State | The state where the event will occur. Each field is populated with a location from a gazetteer value. |
| City | The city where the event will occur. Each field is populated with a location from a gazetteer value. |
| Latitude | The latitude of the city where the event will occur. Each field is populated with a location from a gazetteer value. |
| Longitude | The longitude of the city where the event will occur. Each field is populated with a location from a gazetteer value. |

## MA WARNING SAMPLE CODE

Mercury Challenge Participants submit warnings for MA in the following format:

```
{    "Warning_ID": "BaserateConflict1",
      "Event_Type": "Military Activity",
      "Event_Subtype": "Conflict"
      "Event_Date": "2016-02-18",
      "Country": "Syria",
      "State": "Aleppo",
      "City": "Al Maḥmūdīyah",
      "Latitude": 35.9143,
      "Longitude": 38.0711,
      "Actor": "Syrian Arab Military"}
```

Note that while a GSR event may contain multiple Actors, a warning contains only one Actor.

Participants should not include any fields other than the listed required and optional fields. Doing so may cause a validation error for your warning submission.

## MA SCORING

Two measures are calculated for each MA warning-event (W-E) pair: Quality Score (QS) and Lead Time (LT). The QS is the measure of accuracy of the W-E pair, on a scale from 0 to 1.

MA are scored for Location (LS), Date (DS), and two Facets: Actor (AS), and Event Subtype (ES). These measures are used to calculate the Quality Score (QS).

The following Matching Process is used for MA scoring:

- Linear Assignment is performed using the Hungarian Algorithm to maximize the aggregate Quality Score among matched event-warning pairs.
- Warnings and events may only be matched if they are in the same country and have both Location Score > 0 and Date Score > 0.

## LOCATION SCORING (LS):

The Value Range is 0.0 to 1.0.

LS is linear in distance between event and warning, with a cutoff at 100 km distance.

If the location of the event in the GSR is flagged as Approximate Location, then LS is 1 for distance up to 16.67 km and drops off linearly to 0 at 100 km.

In the following example, the GSR event location is not flagged as Approximate:

Distance = 30km => LS = 0.70 (1 – 30/100)

In this example, the GSR event location is flagged as Approximate:

Distance = 50km => LS = 0.60 (1 – 33.33/83.33)

## DATE SCORING (DS)

The Value Range is 0.0 to 1.0.

The DS is linear in time difference between event and warning, with a cutoff of 4 days.

Example:  2016-04-06 and 2016-04-08 differ by 2 days, so DS = 0.50 (2/4)

## FACET SCORING

Facet Scores take on values of 0 if the warning and event differ, 1 if they agree on the facet.

If the event encodes the facet value as "Unknown" then the warning is assumed to agree no matter what value is assigned.

If the event encodes multiple actors and the warning matches any of them, it's considered to agree.

TABLE 6. FACETS FOR MA SCORING

| Score Name | Facet |
|---|---|
| Actor Score (AS) | Actor |
| Event Subtype Score (ES) | Event Subtype |

The following equation shows the calculation of Quality Score (QS) for MA:

```
QS = (LS + DS + AS + ES)/4
```

## LEAD TIME (LT)

Lead Time (LT) is the difference in days between the warning receipt and the earliest reported date for the event.

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

If this Lead Time value is less than one day, the warning is considered late, it will not be scored, and it does not contribute to the calculation of any performance measures.  The day ends at 21:00:00 UTC, which equates to midnight local in most of the Mercury countries.

Note:  For some events the submission date may be after the event date but before the earliest reported date.  This is still a valid warning and will have a positive lead time.

# NON-VIOLENT CIVIL UNREST EVENTS

## NON-VIOLENT CIVIL UNREST DEFINITION

Non-violent Civil Unrest (CU) events are physical acts that occur in public venues such as demonstrations, marches, and protests. Events that do not involve a physical presence are not considered to be CU. For example, protests that occur on the internet or letter writing campaigns are not CU events.

IARPA generates ground truth for individual non-violent civil unrest events, but challenge participants will forecast *counts* of non-violent civil unrest. The ground truth for counts is computed from the actual individual events. IARPA will make those individual events available to participants, updated on a monthly basis. This section describes how these events are encoded.

The following examples are considered <u>non-violent</u> Civil Unrest:

- Threats of violence, e.g., yelling, cursing
- Police arresting protestors
- Hunger strikes
- Throwing things that would not cause harm, e.g., eggs
- Brandishing clubs, machetes, etc.
- Fireworks
- Blockades of streets
- Theft without damage

The following are examples of situations that would be NOT be considered Non-violent Civil Unrest and should not be forecasted:

- Reports of general instability within a country, natural disasters, scandals, accidents, and criminal events
- Protests that do not require a physical presence such as protests that occur on the internet or letter-writing campaigns
- Civic engagement with officials.  For example, a civic group opposed to some government policy meets with government officials to discuss their concerns, rather than conducting a public protest
- Future protests that are in the planning stages
- Events that are prevented from taking place
- One-person protests
- Criminal and terrorist activity

The following are examples of events of a violent nature that would also be NOT be considered Non-violent Civil Unrest and should NOT be forecasted:

- Clashes with police, e.g., police using tear gas or high pressure water hoses
- Clashes between opposing groups resulting in injuries
- Pushing or being forcibly removed from an area
- Hostage taking
- Self-inflicted wounds, e.g., protesters sewing their mouths shut
- Throwing hard objects which could cause injury or damage
- Hitting with clubs, machetes, etc.
- Burning tires, burning barricades, burning cars or buses, burning buildings

- Looting shops (where the shop is damaged)
- Defacement of property, e.g., painting graffiti on buildings

## NON-VIOLENT CIVIL UNREST LOCATIONS

Non-violent Civil Unrest event counts are encoded for Egypt and Jordan. Predictions include:

- Daily count of non-violent civil unrest events in Egypt;
- Weekly count of non-violent civil unrest events in Jordan;
- Weekly count of non-violent civil unrest events in a 75km radius of Tahrir Square, Egypt;
- Monthly count of non-violent civil unrest events in the Jordanian provinces of Amman, Madaba, and Irbid.

## CU GOLD STANDARD SOURCES (GSS)

CU GSS are representative local, native-language news sources. GSS collection techniques include use of Google Advanced Search (online newspaper versions) and Nexis (print newspaper versions). The following table lists the current set of CU GSS by country and website, including their availability through Nexis (indicated by an X).

TABLE 7. MERCURY CHALLENGE NON-VIOLENT CIVIL UNREST GSS

| 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.

Additional information on event collection from the GSS is found in the Appendix.

## NON-VIOLENT CIVIL UNREST GSR FIELDS

The non-violent Civil Unrest GSR will be provided in two formats. The first format will be a list of individual events and will be available to use as a reference to gain additional insight at the single event level. The second format will be event counts and will be used for scoring.

GSR fields applicable to non-violent Civil Unrest, along with brief encoding instructions, are defined in Table 8. Fields which are only included in the reference listing of individual events are denoted with an asterisk.

TABLE 8. NON-VIOLENT CIVIL UNREST GSR FIELDS

| GSR Fields | Description |
|---|---|
| Event_ID | All GSR single event (reference) records and count format records are assigned an Event ID unique to that record. |
| Event_Type | The Event Type is Civil Unrest. |
| Event_Date | The date the event occurred or, in the case of a multi-day event, the date that the event started. |
| Earliest_Reported_Date | The Earliest Reported Date (ERD) is the publication date for the first article that reported the CU event. |
| Country | Country where the CU event occurred. May be Egypt or Jordan. |
| State | For Monthly events in Jordan. May be Amman, Madaba or Irbid. |
| City | For Weekly events occurring within 75 km of Tahrir Square. |
| Case_Count | Number of events in day/week/month |
| Approximate_Location * | Possible values are "True" and "False." Location is Approximate if: 1) the location is not in a gazetteer and the closest GDB City location is used, or 2) if the location is described in relation to a city in a gazetteer value. An event is flagged Approximate Location if the exact location of the event is unknown, or if the exact location if not available in a gazetteer value. The following location information would indicate the use of the Approximate Location flag when no other location information for an event is available:<br><br>• near Riyadh<br>• on the border between Latakia and Turkey<br>• the road linking Sijariya and Tel Mushaihida<br>• west of Saqalawiyah<br><br>More detailed information on Approximate Location is provided in the Appendix. |
| Latitude* | The latitude of the city where the event will occur. Each field is populated with a location from a gazetteer value. |
| Longitude* | The longitude of the city where the event will occur. Each field is populated with a location from a gazetteer value. |
| News Source* | Name of the primary GSS. See Table 5 for a list of Mercury Challenge approved news sources. |
| First_Reported_Link* | URL documenting the first report of the CU event. |
| GSS_Link* | URL of the report from a GSS source that best describes the CU event. |
| Other_Links* | URL for additional news reports. If there are multiple URLs, separate with a semi-colon. |
| Revision_Date* | Date of most recent update to the Event ID. |

## SAMPLE NON-VIOLENT CIVIL UNREST CODE IN JSON FORMAT

Sample non-violent Civil Unrest entry in single event (reference) format:

```
{
    "Event_ID": 125,
    "Event_Type": "Civil Unrest",
    "Event_Date": "2016-03-04",
    "Earliest_Reported_Date": "2016-03-05",
    "Country": "Jordan",
    "State": "Irbid",
    "City": "Ar Ramthā",
    "Approximate_Location": "False",
    "Latitude": 32.5592,
    "Longitude": 36.0069,
    "News_Source": "Al-Rai",
    "First_Reported_Link": "http://www.alrai.com/article/772622.html",
    "GSS_Link": "http://www.alrai.com/article/772622.html",
    "Other_Links": "",
    "Revision_Date": "2016-08-17",
}
```

Sample non-violent Civil Unrest entry in Daily count format:

```
{
    "Case_Count": 13,
    "Country": "Egypt",
    "Earliest_Reported_Date": "2015-04-30",
    "Event_Date": "2015-04-30",
    "Event_ID": "CU_Count_Egypt_2015-04-30",
    "Event_Type": "Civil Unrest",
}
```

Sample non-violent Civil Unrest entries in Weekly count format:

```
{
    "Case_Count": 16,
    "Country": "Jordan",
    "Earliest_Reported_Date": "2015-05-04",
    "Event_Date": "2015-04-29",
    "Event_ID": "CU_Count_Jordan_2015-04-29",
    "Event_Type": "Civil Unrest",
}
```

Sample non-violent Civil Unrest entry in Monthly count format:

```
{
    "Case_Count": 11,
    "Country": "Jordan",
    "State": "Amman",
    "Earliest_Reported_Date": "2015-05-01",
    "Event_Date": "2015-04-15",
    "Event_ID": "CU_Count_Amman_April_2015",
    "Event_Type": "Civil Unrest",
}
```

## NON-VIOLENT CIVIL UNREST WARNING FORMAT

The following fields are required for CU Event Warnings:

TABLE 9. NON-VIOLENT CIVIL UNREST WARNING FIELDS

| Warning Fields | Descriptions |
|---|---|
| **Warning ID** | The Warning ID is a unique alphanumeric string provided by the challenge participant. To update a warning, submit a revised warning with the same Warning ID. The Mercury Challenge scoring system will automatically create a sequence number for the updated warning. |
| **Event_Type** | Civil Unrest |
| **Event_Date** | The predicted event date. For daily events, the day being forecast. For weekly events, the Wednesday of the week being forecast. For monthly events, the 15$^{th}$ of the month being forecast. |
| **Country** | Egypt or Jordan. Required for all warnings. |
| **State** | For Monthly events, Amman, Madaba or Irbid. Not required for Daily or Weekly Events |
| **City** | For Weekly count of events in within a 75 km radius of Tahrir Square use Tahrir Square. Not required for other non-violent CU events. |
| **Case_Count** | Number of events predicted |

## SAMPLE NON-VIOLENT CIVIL UNREST WARNING FORMAT

Warnings may only include the fields listed in Table 9. Examples are listed below:

```
{
    "Warning_ID":"CU1",
    "Event_Type":"Civil Unrest",
    "Event_Date":"2018-05-09",
    "Country":"Egypt",
    "Case_Count": 10
}


{
    "Warning_ID":"CU2",
    "Event_Type":"Civil Unrest",
    "Event_Date":"2018-05-09",
    "Country":"Egypt",
    "City": "Tahrir Square",
    "Case_Count": 3
}
```

## NON-VIOLENT CIVIL UNREST SCORING

Two measures are calculated for Weekly Count, a Quality Score (QS), and Lead-Time (LT). Warnings and Events are matched by location and date. Weeks follow the ISO pattern. Months follow a calendar month (CM) pattern. Note that warnings for each ISO week or CM are only scored against an event for the same ISO week or CM.

| Mercury Challenge ISO Week | First Day of the ISO Week |
|---|---|
| 1 | 2017-01-02 |
| 2 | 2017-01-09 |
| 3 | 2017-01-16 |
| 4 | 2017-01-23 |
| 5 | 2017-01-30 |

TABLE 11. CALENDAR MONTH DATE SAMPLE

| Mercury Challenge CM | First Day of the CM |
|---|---|
| 1 | 2017-01-01 |
| 2 | 2017-02-01 |
| 3 | 2017-03-01 |
| 4 | 2017-04-01 |
| 5 | 2017-05-01 |

The QS for CU Warnings depends on the absolute difference between the forecasted count and the actual count and is calculated as follows:

```
Quality Score = 1- (abs(Predicted – Actual)/max(Predicted, Actual, 4))
```

TABLE 12. QS EXAMPLES FOR MONTHLY OR WEEKLY CASE COUNTS

| Actual | Predicted | QS |
|---|---|---|
| 0 | 0 | 1.00 |
| 0 | 2 | 0.50 |
| 1 | 2 | 0.75 |
| 5 | 6 | 0.83 |
| 11 | 12 | 0.92 |
| 998 | 999 | 1.00 |

Lead Time (LT) is the difference in days between the warning receipt and the Event Date for the event.

Weekly count Event Dates are Wednesday of that week

Monthly count Event Dates are the 15th of that month regardless of the days in the month

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

If this value is less than one day, the warning is considered late, it will not be scored, and it does not contribute to the calculation of any performance measures. The day ends at 23:59:59 UTC.

# DISEASE EVENT

## DISEASE DEFINITION

The disease event for the Mercury Challenge is a weekly case count of confirmed MERS cases in Saudi Arabia.

## WEEKLY CASE COUNT DISEASE EVENTS

For MERS in Saudi Arabia weekly disease case counts are collected and are encoded by Epidemiological Week (EW). The Mercury Challenge conforms to the Centers for Disease Control and Prevention standard definition of EW, which defines the EW as beginning on Sunday and ending on Saturday. (See Table 17.)

### WEEKLY CASE COUNT COUNTRIES

The Weekly MERS Case Count Country is for Saudi Arabia.

### WEEKLY CASE COUNT GSS

Primary and Back-up sources for this event class are provided in the following table.

TABLE 13. WEEKLY CASE COUNT PRIMARY AND BACK-UP SOURCES

| Disease | Primary Sources | Back-Up Sources |
|---|---|---|
| MERS specific to Saudi Arabia | EMPRES-i | HealthMap Google Advanced Search |

HealthMap is the preferred back-up source for Weekly Case Count events, and is used to confirm data collected from EMPRES-i.

### WEEKLY CASE COUNT GSR FIELDS

There are 12 GSR fields for Weekly Case Count Diseases. Eleven are required and one is optional. These fields are defined as follows:

TABLE 14. WEEKLY CASE COUNT GSR FIELDS

| GSR Fields | Optional or Required | Description |
|---|---|---|
| Event_ID | Required | All GSR records have an Event ID unique to that record. |
| Event_Type | Required | Event type is "Disease" |
| Country | Required | For this count, only the country record "Saudi Arabia" is required. |
| Disease | Required | Disease is "MERS" |
| Case_Count | Required | This field contains the number of confirmed Weekly Case Count cases. |
| Event_Date | Required | The first day of the reported EW, encoded in yyyy-mm-dd format. See Table 17 for a list of sample EWs. |
| Earliest_Reported_Date | Required | The publication date of the first official confirmation of the Weekly Case Count Disease case. If no disease occurred in that EW, enter the start date for the EW. |
| News_Source | Required | The News Source is the name of the primary GSS for the event. |
| First_Reported_Link | Required | URL of the source that first reported the Weekly Case Count Disease case confirmation. |
| GSS_Link | Required | URL of the GSS source that that provides the most complete event information. |
| Other_Links | Optional | URLs for other supporting information, if available.  Separate multiple URLs with a semi-colon. |
| Revision_Date | Required | Date of most recent update to the Event ID. |

SAMPLE WEEKLY CASE COUNT EVENT IN JSON FORMAT:

```
{
    "Case_Count": 0,
    "Country": "Saudi Arabia",
    "Disease": "MERS",
    "Earliest_Reported_Date": "2015-05-03",
    "Event_Date": "2015-05-03",
    "Event_ID": "Disease_Saudi_Arabia_MERS_2015-05-03",
    "Event_Type": "Disease",
    "First_Reported_Link": "http://empres-i.fao.org/eipws3g/#h=1",
    "GSS_Link": "http://empres-i.fao.org/eipws3g/#h=1",
"News_Source": "UN Empres-i.fao",
    "Other_Links": "-",
    "Revision_Date": "2016-08-17"
},
```

WEEKLY CASE COUNT WARNING FORMAT

The following fields are required for Weekly Case Count warnings:

TABLE 15. WEEKLY CASE COUNT DISEASE WARNING FIELDS

| Warning Fields | Definitions |
|---|---|
| Warning ID | The Warning ID is a unique alphanumeric string provided by the challenge participant.  To update a warning, submit a revised warning with the same Warning ID.  The Mercury Challenge scoring system will automatically create a sequence number for the updated warning. |
| Event_Type | Disease |
| Case_Count | This field contains the weekly case count of confirmed cases. |
| Disease | MERS |
| Date | The first day of the associated EW (for the Mercury Challenge the first day of the EW is always Sunday), encoded in the yyyy-mm-dd format. See Table 17. |
| Country | Saudi Arabia |

## SAMPLE CASE COUNT WARNING IN JSON FORMAT:

Mercury Challenge Participants submit warnings for each Weekly Case Count event in the following JSON format:

```
{
    "Warning_ID": MERSSA1",
    "Event_Type": "Disease"
    "Disease": "MERS",
    "Date": "2016-02-10",
    "Country": "Saudi Arabia",
    "Case_Count": 6
}
```

## WEEKLY CASE COUNT DISEASE SCORING

Two measures are calculated for Weekly Case Count, a Quality Score (QS), and Lead-Time (LT). Warnings and Events are matched by country, date and disease.  Weeks follow the Epidemiological Week (EW)  pattern, beginning on Sunday and ending on Saturday.  Note that warnings for each EW are only scored against an event for the same EW.

The QS for Case Count Warnings depends on the absolute difference between the forecasted count and the actual count and is calculated as follows:

Quality Score = 1- (abs(Predicted – Actual)/max(Predicted, Actual, 4))

TABLE 16. QS EXAMPLES FOR MONTHLY OR WEEKLY CASE COUNTS

| Actual | Predicted | QS |
|--------|-----------|------|
| 0 | 0 | 1.00 |
| 0 | 2 | 0.50 |
| 1 | 2 | 0.75 |
| 5 | 6 | 0.83 |
| 11 | 12 | 0.92 |
| 998 | 999 | 1.00 |

Table 17.  EPIDEMIOLOGICAL WEEK DATES SAMPLE

| Mercury EW | First Day of the EW |
|:----------:|:-------------------:|
| 1 | 2017-01-01 |
| 2 | 2017-01-08 |
| 3 | 2017-01-15 |
| 4 | 2016-01-22 |
| 5 | 2016-01-29 |

Lead Time (LT) is the difference in days between the warning receipt and the 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.

If this value is less than one day, the warning is considered late, it will not be scored, and it does not contribute to the calculation of any performance measures.

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.

# ACRONYMS AND KEY DEFINITIONS

| | |
|---|---|
| CU | Non-violent Civil Unrest |
| DARPA | Defense Advanced Research Projects Agency |
| Epiweek | Reference for Epidemiological Week (EW) |
| ERD | 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 Military Activity, either Conflict or Force Posture |
| 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 |
| ICEWS | Integrated Conflict Early Warning System: a data repository developed by DARPA. |
| 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 |
| QA | Quality Assurance |
| QS | Quality Score |
| Participant | Research team responsible for developing automated predictive models for societal events. |
| W-E | Warning-event |
| Warning | Participant event forecast |

# RESOURCES

- **Integrated Crisis Early Warning Systems (ICEWS) Event Data**
  Data produced for the Integrated Crisis Early Warning System (ICEWS) for the Defense Advanced Research Projects Agency (DARPA) and Office of Naval Research (ONR).
  https://dataverse.harvard.edu/dataverse/icews

- **UIUC Cline Center Event Data**
  The University of Illinois Urbana-Champaign has produced datasets focused on the extraction of civil unrest from unstructured data sources through the Social, Political, Economic Event Database (SPEED) Project.
  http://www.clinecenter.illinois.edu/data/event/

- **ACLED Event Data**
  The Armed Conflict Location & Event Data Project (ACLED) is a disaggregated conflict analysis and crisis mapping project. ACLED is the highest quality, most widely used, realtime data and analysis source on political violence and protest in the developing world. Practitioners, researchers and governments depend on ACLED for the latest reliable information on current conflict and disorder patterns.
  https://www.acleddata.com/

- **World Bank Data Sources (GDP, population, etc)**
  This is a portal designed to make World Bank data easy to find, download, and use. All of the data found here can be used free of charge with minimal restrictions
  https://data.worldbank.org/

- **Common Crawl – Global News Scraper**
  This news crawler is based on StormCrawler and covers data containing articles from news sites all around the world.
  http://commoncrawl.org/2016/10/news-dataset-available/

- **Predicting the international spread of Middle East respiratory syndrome (MERS)**
  A study on the spread of MERS using a statistical model to analyze data from airline transportation data
  https://bmcinfectdis.biomedcentral.com/track/pdf/10.1186/s12879-016-1675-z

- **‘Beating the News’ with EMBERS: Forecasting Civil Unrest using Open Source Indicator**
  http://people.cs.vt.edu/naren/papers/kddindg1572-ramakrishnan.pdf
  EMBERS for Early Model Based Event Recognition using Surrogates is a system developed by industry and academia to forecast events using open source indicators

- **Geopolitical Forecasting Skill in Strategic Intelligence**
  A meta-analysis on the efficacy of 3622 geopolitical forecasts from strategic intelligence reports
  https://onlinelibrary.wiley.com/doi/pdf/10.1002/bdm.2055

- **Official Github Basics Guide**
  A guide to the basics of GitHub from its creators
  https://guides.github.com

- **Creating, cloning, and archiving GitHub repos**
  A GitHub help article on how to copy a GitHub code repository
  https://help.github.com/categories/creating-cloning-and-archiving-repositories/

- **Simple guide to git**
  A popular and simple guide to GitHub that's been translated into international languages
  http://rogerdudler.github.io/git-guide/