MITRE TECHNICAL REPORT 160246

MITRE

Open Source Indicators Program Handbook

This work was accomplished in support of the Intelligence Advanced Research Projects Activity (IARPA) Open Source Indicators Program, BAA number IARPA-BAA-11-11. This technical data was first produced for the U.S. Government under contract 2009-0917826-016, and is subject to the Rights in Data-General Clause 52.227-14, Alt. IV (DEC 2007). The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, expressed or implied, of IARPA or the U.S. Government or any of the authors' host affiliations.

© 2017 The MITRE Corporation. All Rights Reserved.

Terry Reed
Kim Belvin
Molly Hickman
Maeve Kokolus
Francisco Mendoza
Ana Phelps
Kurt Reisler

Approved for Public Release;
Distribution Unlimited. Case Number 16-4297

This page intentionally left blank.

Abstract

The Open Source Indicators (OSI) research program investigates techniques to make automated predictions of significant societal events through the continuous and automated analysis of publically available data such as news media, social media, informational websites, and satellite imagery. Societal events of interest include civil unrest, disease outbreaks, and election results. Geographic areas of interest include countries in Latin America (LA) and the Middle East and North Africa (MENA).

Foreword

This forward serves as an introduction to the Open Source Indicators (OSI) project information posted on the Dataverse. OSI project information includes:

- OSI Program Handbook
- OSI Gold Standard Report (GSR)
- OSI Scoring Code

The OSI Program Handbook contains the OSI program description and approach used to evaluate the progress of three competing research teams or Performers (BBN, HRL, and Virginia). The MITRE Corporation provided independent Test and Evaluation (T&E) for the OSI program. The Handbook discusses the evaluation of Performer forecasts by comparing those forecasts to a validated set of ground truth. Further, the Handbook describes the sources used for ground truth, the methodology to collect and encode ground truth information, how ground truth information was validated, and how ground truth was compared to Performer forecasts based on a set of program metrics.

The OSI GSR contains a record for each validated OSI event including civil unrest, disease, and elections. The reader should refer to the OSI Program Handbook Section 2 for detailed descriptions of each field in a GSR record. The OSI GSR contains encoded events for selected Latin America countries. Events were encoded for the period of May 2013 through March 2015.

The OSI Government team controls the OSI scoring code used to compare and assess performer forecasts to validated ground truth; requests for copies may be sent to the Points of Contact listed below.

Points of Contact:
OSI Program Handbook and GSR—Terry Reed (MITRE T&E Task Lead), treed@mitre.org, 703-983-2557 or Murray McDonald (MITRE IARPA Project Lead), jmmcdona@mitre.org, 703-983-7275
OSI Scoring Code—Jason Matheny (OSI Program Manager), jason.matheny@iarpa.gov, 301-851-7451 or Naren Ramakrishnan (Virginia Tech Principal Investigator), naren@cs.vt.edu, 571-858-3331

This page intentionally left blank.

Table of Contents

1 Introduction 1-1
1.1 Purpose 1-1
1.2 Background 1-1
1.3 Approach 1-1
1.4 Handbook Roadmap 1-3
2 OSI Event Classes 2-1
2.1 Civil Unrest (CU), City-Level (01) and Widespread (07) Events 2-2
2.2 Vote, Event Class (02) 2-8
2.3 Disease, Event Class (03) 2-11
2.3.1 Rare Disease Event Classes 2-12
2.3.2 Chikungunya Event Class 2-16
2.3.3 Saudi Arabia Middle East Respiratory Syndrome (MERS) Event Class 2-18
2.3.4 Polio Event Class 2-21
2.3.5 Disease Season Event Classes 2-23
2.3.6 Influenza-like Illness Event Class (Retired) 2-28
2.4 Integrated Crisis Early Warning System (ICEWS) 2-31
3 Bibliography 3-1
Appendix A Civil Unrest Encoding Guidance and Keywords A-1
A.1 Civil Unrest Event Identification A-1
A.1.1 What Constitutes the Boundaries of an Event A-1
A.1.2 How to Distinguish Continuous Events from Multiple Events A-2
A.1.3 Events with Multiple Encodings A-3
A.1.4 Identifying Crowd Size A-4
A.1.5 Identifying Event Date A-4
A.1.6 Identifying Earliest Reported Date A-5
A.1.7 Identifying Locations A-5
A.1.8 Identifying Violence A-5
A.2 Civil Unrest Event Reason Encoding Guidance A-6
A.3 Civil Unrest Population Encoding Guidance A-10
Appendix B Civil Unrest (CU) Event Discovery B-1
B.1 Advanced Google Search Terms B-2
B.2 Lexis-Nexis Search Terms B-3
Appendix C Selected OSI Elections C-1

Appendix D Monthly Scoring and Performance Measures D-1
D.1 Update Sequences D-1
D.2 Program Performance Measures D-1
D.2.1 Common Performance Measure Calculations D-2
D.2.2 Quality Score Calculations D-3
Appendix E Base Rate Model E-1
Appendix F GSR Quality Control (QC) F-1
F.1 Civil Unrest and Disease Quality Review F-1
F.2 Semi-Automated Verification and Validation (V&V) F-1
F.3 Warning Ingest Evaluation Criteria F-2
F.4 Scoring Quality Control (QC) F-2
Appendix G OSI Gazetteer Database (OGDB) G-1
Appendix H Instructions for Performers H-1
H.1 Audit Trail H-1
H.2 Replication and Ablation Tests H-2
H.3 Trade-offs: Precision v. Recall and Lead-time v. Quality Score H-2
H.4 Automated Warning Generation H-2
Appendix I Program Milestones I-1
Appendix J Acronyms J-1

List of Figures

Figure 1. High Level OSI Approach 1-1
Figure 2. Civil Unrest LS QS Calculation, Maximum Score is 4 2-7
Figure 3. LS Scorer Civil Unrest QS Example 2-7
Figure 4. Traditional Civil Unrest QS Calculation, Maximum Score is 4 2-8
Figure 5. Traditional Scorer Civil Unrest QS Example 2-8
Figure 6. Rare Disease QS Calculation, Maximum Score is 4. 2-15
Figure 7. Rare Disease QS Scoring Example 2-15
Figure 8. Mexico 2011-2012 Flu Season (Example) 2-25
Figure 9. LS Civil Unrest QS Calculation, Maximum Score is 4 ..... D-3
Figure 10. LS Scorer Civil Unrest QS Example D-3
Figure 11. Traditional Civil Unrest QS Calculation, Maximum Score is 4 ..... D-4
Figure 12. Traditional Scorer Civil Unrest QS Example D-4
Figure 13. Rare Disease QS Calculation Example D-4
Figure 14. Rare Disease QS Scoring Example D-5
Figure 15. Base Rate Warning Development Process E-1

List of Tables

Table 1. Program Performance Measures 1-3
Table 2. Civil Unrest Event Types 2-2
Table 3. Civil Unrest Violence Indicators 2-3
Table 4. Civil Unrest Populations 2-3
Table 5. Latin America Civil Unrest Countries 2-4
Table 6. Civil Unrest Gold Standard Sources (GSS) 2-4
Table 7. Civil Unrest GSR Fields 2-5
Table 8. Civil Unrest Warning Fields 2-6
Table 9. Vote Event Types 2-8
Table 10. Vote Event Countries 2-9
Table 11. Vote GSR Fields 2-9
Table 12. Vote Warning Fields 2-10
Table 13. Disease Summary 2-12
Table 14. Rare Disease Countries 2-13
Table 15. Rare Disease Ground Truth Sources 2-13
Table 16. Rare Disease GSR Fields 2-13
Table 17. Rare Disease Warning Fields 2-14
Table 18. OSI Epidemiological Week Dates Sample 2-16
Table 19. Chikungunya Case Count Disease Countries 2-16
Table 20. Chikungunya Ground Truth Sources 2-16
Table 21. Chikungunya GSR Fields 2-17
Table 22. Chikungunya Warning Fields 2-17
Table 23. OSI Epidemiological Week Dates Sample 2-18
Table 24. Saudi Arabia MERS Ground Truth Sources 2-19
Table 25. Saudi Arabia MERS Disease GSR Fields 2-19
Table 26. Saudi Arabia MERS Disease Warning Fields 2-20
Table 27. OSI Epidemiological Week Dates Sample 2-21
Table 28. Polio Event Class Countries 2-21
Table 29. Polio Ground Truth Sources 2-21
Table 30. Polio GSR Fields 2-22
Table 31. Polio Warning Fields 2-22
Table 32. OSI Epidemiological Week Dates Sample 2-23
Table 33. Disease Season Event Code Definitions 2-23
Table 34. Disease Season Countries 2-25
Table 35. Disease Season Ground Truth Sources 2-25
Table 36. Disease Season GSR Fields 2-25
Table 37. Disease Season Warning Fields 2-27
Table 38. OSI Epidemiological Week Dates Sample 2-28
Table 39. ILI Countries 2-28
Table 40. ILI GSR Fields 2-30
Table 41. ILI Warning Fields 2-30
Table 42. ICEWS Countries 2-32
Table 43. ICEWS Warning Fields 2-32
Table 44. Civil Unrest Violence Indicators (from Section 2.1) A-5
Table 45. Civil Unrest Event Reason Encoding Guidance A-7

Table 46. Civil Unrest Population Encoding Guidance A-10
Table 47. Civil Unrest Ground Truth Sources (from Section 2.1) B-2
Table 48. Selected OSI Elections C-1
Table 49. Overview of Program Performance Measures D-1
Table 50. Summary of Computer Performance Measures by Event Type D-2
Table 51. Common Performance Measure Calculations D-2

This page intentionally left blank.

1 Introduction

1.1 Purpose

The goal of the Open Source Indicators (OSI) Program is to make automated predictions of significant societal events through the continuous and automated analysis of publicly available data such as news media, social media, informational websites, and satellite imagery. Societal events of interest include civil unrest, disease outbreaks, and election results. Geographic areas of interest include countries in Latin America (LA) and the Middle East and North Africa (MENA).

This Handbook is intended to serve as a reference document for the Program and a companion to publicly released event data from the Program. It provides guidance regarding the types of events considered; the submission of automated predictions or "warnings;" the development of ground truth; the test and evaluation of submitted warnings; performance measures; and other programmatic information.

1.2 Background

IARPA initiated a solicitation for OSI Research Teams, also called Performers, in late summer 2011 for one base year and two option years of research. MITRE was selected as the Test and Evaluation (T&E) Team in November 2011. Following a review of proposals, three Performers (BBN, HRL, and Virginia Tech (VT)) were selected. The OSI Program officially began in April 2012; manual event encoding and formal T&E ended in March 2015. Note that while each Performer is represented by a lead organization, in practice each Team is composed of a diverse set of academic and industry organizations working together collaboratively.

In March 2014, IARPA down-selected to one Performer Team led by VT with the requirement that VT integrate the best practices and systems from all OSI participants.

1.3 Approach

OSI consists of three parts: 1) the development of automated warnings, 2) the development of ground truth in the form of a Gold Standard Report (GSR), and 3) test and evaluation, which compares received warnings to ground truth to generate a set of performance measures or score.

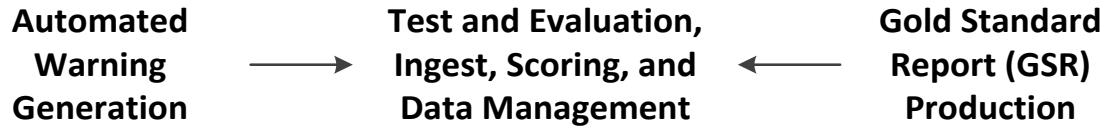


Figure 1. High Level OSI Approach

Based on the collection and assessment of a diverse set of open source information, Performers develop and continually improve their systems that generate warnings automatically with no human in the loop. These warnings predict societal events of interest for a set of countries in LA and MENA. Each warning contains information relevant to the specific type of societal event.

This information includes:

- The type of warning (e.g., Civil Unrest—Employment and Wages)
- Whether the event was violent or not violent
- Location (the Country, State, and City where the event is forecast to occur)
- Population (for a civil unrest event, who participated in the event, such as Teachers)
- When the event will occur (month/day/year)
- The Performer's estimate of the probability of event occurrence.

Performers forward warnings to the MITRE OSI Warning Ingest process continuously, 24 hours a day and 7 days a week (24x7), and the OSI Ingest Process likewise operates 24x7. Upon receipt of a warning, the Ingest Process parses each warning, evaluates each warning according a set of ingest criteria, and adds a date/time-stamp as well as an identifier unique to each Performer. If the warning does not pass the Ingest Process evaluation, the warning is rejected and an email is sent to the Performer. Performers may resubmit rejected warnings; however, each resubmitted warning will receive a new date/time-stamp, which affects the calculation of the Lead Time program performance measure, described below. Accepted warnings are maintained in the OSI Warning Database. See Appendix F for more information on warning ingest evaluation criteria.

On a monthly basis, a GSR is assembled and delivered to the OSI Program Manager (PM) and Performers. The GSR is developed through the data discovery and analysis of ground truth data sources. From that process, an "event" is encoded with a set of information specific to the event type. For example, a Civil Unrest event contains an event identifier (ID), date of occurrence, date that the event was first reported (earliest reported date or ERD), location of the event (Country, State, City), type of event (e.g., Housing), whether the event was violent or not violent, and the population participating in the event (e.g., Refugees/Displaced). In addition, each GSR event contains background information documenting the source of the event, brief description of the event, and the encoder's comments, as needed. Note the OSI encoding ontology is derived from both the Integrated Data for Events Analysis (IDEA) [1] and Text Analysis By Augmented Replacement Instructions (TABARI) [2] encoding typologies.

Upon completion of the GSR, events are compared with warnings during the monthly scoring process. Using a linear optimization program (based on the Kuhn–Munkres or Hungarian algorithm [3]), warnings and events are matched based on a Quality Score (QS), which is a measure of the accuracy of the match. The optimization program looks only for the "best match"; the QS must be non-zero, otherwise no minimum QS threshold is applied in determining the best match. During this process only one warning is matched to one event, i.e. a warning may not be matched to multiple events and an event may not be matched to multiple warnings.

The scoring process supports "look-back" to the previous month for possible matches. For example, an event occurring on the 2nd of the month may be matched with a warning for the 30th of the previous month. The look-back period is determined by the PM. Following the development of warning-event (W-E) pairs, performance measures (assembled into a scoring report) are developed and sent to the PM and Performer. Performers use the monthly scores and the GSR as a basis to improve their systems' forecasts; the PM uses monthly scores to assess performance and determine areas for improvement.

The set of overall program performance measures follows. Note the calculation of the QS is unique to the W-E pair type, e.g., civil unrest vice disease.

Table 1. Program Performance Measures

Performance Measure	Definition
Quality Score (QS)	Assessment of the accuracy of the W-E pair match.
Lead-Time (LT)	Difference between the warning-receipt time-stamp and ground truth ERD in days.
Precision	Number of W-E pairs for the scoring period divided by the number of warnings received for the same period.
Recall	Number of W-E pairs for the scoring period divided by the number of GSR events for the same period.
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	Mean precision for all W-E pairs, reported for each country for a scoring period.
Mean Recall	Mean recall for all W-E pairs, reported for each country for a scoring period.
Mean Probability	Mean probability score (Brier score) for all W-E pairs, reported for each country for a scoring period.

1.4 Handbook Roadmap

The Handbook consists of Section 2: OSI Event classes and a set of 10 appendices.

Section 2 of this Handbook was developed as a "one-stop-shop" for each OSI event class. Each subsection in Section 2 is written to stand-alone. Because of this, there is some redundancy across the subsections, e.g., acronyms are redefined in each subsection for the reader's convenience. Each subsection contains a complete set of information for an event class including: countries, ground truth sources, event and warning formats, warning ingest requirements, and performance measures/scoring. In this way, the reader may quickly go to the item of interest without reading the entire Handbook. The appendices provide more detailed information as follows:

- Appendix A: Civil Unrest Encoding Guidance—detailed encoding guidance used by the analysts who encode civil unrest ground truth events.
- Appendix B: Civil Unrest Event Discovery—discussion of the techniques used to discover relevant civil unrest ground truth events.
- Appendix C: PM-Selected OSI Elections—list of OSI elections.
- Appendix D: Scoring and Performance Measures—discussion of all program performance measures and how they are calculated.
- Appendix E: Base Rate Model—PM-developed reference model derived from historical GSR civil unrest and disease events.
- Appendix F: Quality Control—a discussion of: 1) Civil Unrest and Disease Quality Review, 2) GSR Semi-automated Verification and Validation (V&V), 3) Warning Ingest Evaluation Criteria, and 4) Scoring Quality Control (QC).

- Appendix G: OSI Gazetteer Database (OGDB)—discussion of the database used as the location-naming standard for OSI.
- Appendix H: Instructions for Performers
- Appendix I: Program Milestones—performance goals for the performer by Program year.
- Appendix J: Acronyms

General Definitions:

Event Class: the broad category of the societal event of interest, such as Civil Unrest, Vote, or Disease

Event Type: a specific event category within an event class, such as Civil Unrest: Employment and Wages

Event Code: the 3-4 digit identifier associated with an event type, such as 0111 represents a Civil Unrest: Employment and Wages non-violent event.

Warning: Performer event forecast

2 OSI Event Classes

This section contains a detailed discussion of all OSI event classes, how they are encoded and evaluated:

- Civil Unrest (CU): City-Level (01), Widespread (07), and Integrated Crisis Early Warning Systems (ICEWS) (05 and 06) Events
- Vote, Event Class (02)
- Disease, Event Class (03)
    - Rare Disease (RD) Event Classes
    - Saudi Arabia MERS Event Class
    - Chikungunya Event Class
    - Polio Event Class
    - Disease Season Event Classes
    - Influenza-like Illness (ILI) Event Class (Retired)

A detailed set of information is provided for each event class; this information is sufficient to correctly encode and score each event class. Note that the information is presented in the same order and same format; this information is specific to the event class.

1. Description—general description of the event class.
2. Countries—the associated countries. Note that not all event classes are used for every OSI country.
3. Ground Truth Sources—a discussion of the ground truth sources used to develop GSR events.
4. Event Format—a definition for each of the 19 fields in a GSR event record.
5. Warning Format—a definition of each of the 7 fields in a warning record.
6. Performance Measures—definitions of performance measures. Note some performance measures are common for all event classes while other are specific to an event class.

In addition, there are two processes common to each event class: Warning Ingest and the management of Update Sequences.

1. Warning Ingest—Performers email warnings to MITRE. When the MITRE mail server receives warnings, they are date/time-stamped. Periodically, the ingest process parses and archives warnings for scoring at a later time.

   Warnings that do not conform to the format described above are rejected by the warning ingest process. Performers 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.

   Update sequences may contain only one type of event code; different event codes may not be mixed in an update sequence.

If the last member of an update sequence has a probability of 0.0, the update sequence warning is withdrawn from scoring. A withdrawn update sequence may be reinstated with the submission of a new warning with a non-zero probability. Although withdrawn warnings are not scored, they are considered in the calculation of Precision.

2. Update Sequences—As more information is available, the Performer 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 update sequence warnings submitted at least one day in advance of the reported date are used to calculate the update sequence score.

2.1 Civil Unrest (CU), City-Level (01) and Widespread (07) Events

OSI CU events are physical acts that occur in public venues such as, demonstrations, marches, and protests. For example, protests that occur on the internet or letter writing campaigns that do not require a physical presence are not considered to be OSI CU.

There are two types of CU events—city-level events and widespread events.

- City-level events—event occurs in a (Country, State, City) location triplet. Note a city-level event may also be part of a larger widespread event.
- Widespread events—a widespread event is characterized by protestors assembling for a common, specific reason (as indicated by event type) at several locations across a state or in several cities. Widespread events are encoded at the country level. As discussed below, the GSR provides a link between a widespread event and any associated city-level events. Note a widespread event may not necessarily be centrally orchestrated. For example, during June 2013 Brazilian students protested against bus fare hikes in all major metropolitan areas across the country; here there was a common reason but no central organization across the country and thus no associated widespread event.

A CU event is encoded with a 4-digit code. The first 2 digits indicate whether the event is a city-level (01) or widespread (07) event. The third digit represents one of 6 event types. The fourth digit represents violence where 1 is non-violent and 2 is violent.

The six CU event types are summarized in Table 2. A more detailed description of each event code including examples is provided in Appendix A: Civil Unrest Encoding Guidance.

Table 2. Civil Unrest Event Types

Event Code	Event Description
011x (city-level) 071x (widespread)	**Employment and Wages**—protests regarding employment, wages, job benefits, and working conditions.
012x (city-level) 072x (widespread)	**Housing and Shelter**—protests regarding housing, temporary shelter, evictions, relocations, and eminent domain issues.
013x (city-level) 073x (widespread)	**Land, Energy, and Resources**—protests regarding access to or the management of energy, public resources (such as clean water), social services, health resources, property rights (not housing), environmental issues, and eviction from land (not homes).
014x (city-level) 074x (widespread)	**Other Economic Issues**—protests regarding government policies or actions related to the management of the nation's economy, e.g., taxes, fees, subsidies, trade agreements.

Event Code	Event Description
015x (city-level) 075x (widespread)	**Other Government and Political Issues**—protests related to governmental policies, mandates, regulations, elections, political rallies, and general prison issues.
016x (city-level) 076x (widespread)	**Other Civil Unrest**—protests without a stated reason, events that do not fit into the previous five categories, also specific events such as religious or cultural marches, peace demonstrations, protests against foreign governments, protests against private institutions, and awareness marches (e.g., breast cancer awareness).

Again, the fourth digit of the event code indicates violence, where 1 is non-violent and 2 is violent. Note that in the absence of additional information, all civil unrest events are assumed to be non-violent. The following table presents violence indicators, which may be present in the text of the news report or in any accompanying photograph.

Table 3. Civil Unrest 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 sticks, bars, machetes, etc.	• Throwing things that would not cause harm (e.g., eggs) • Brandishing sticks, bars, 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

Further, a CU event is encoded for every population participating in the event. The following is the list of OSI populations (See Appendix A for examples of each population type.) Note that some population groups are more specific than others. For instance, Teachers and Medical workers are a particular kind of Labor force.

Table 4. Civil Unrest Populations

Population
General Population
Business
Ethnic
Legal
Education
Religious
Media
Medical
Labor
Refugees/Displaced

Agricultural

Civil Unrest Countries

Civil Unrest is encoded for ten LA countries:

Table 5. Latin America Civil Unrest Countries

Countries
Argentina
Brazil
Chile
Colombia
Ecuador
El Salvador
Mexico
Paraguay
Uruguay
Venezuela

Civil Unrest Gold Standard Sources (GSS)

All CU events are discovered from local news sources. Event discovery techniques include Advanced Google Search (limited to the newspaper website), Lexis-Nexis queries, and Newspaper Direct review. Advanced Google Search is used to discover events in online media. Lexis-Nexis and Newspaper Direct are used to discover events in print media. While most CU events are encoded in both online and print media, there is not a 100% overlap. For the period of May 2013 through October 2014, 3.5% of all CU events were only reported in print media. Appendix B: CU Event Discovery and Encoding Process provides more details including a discussion of GSS selection criteria and the keywords used for Advanced Google Search and Lexis-Nexis queries.

The following table lists CU GSS or ground truth sources by country and website including their availability through either Lexis-Nexis or Newspaper Direct.

Table 6. Civil Unrest Gold Standard Sources (GSS)

Country	News Source	Website (for Advanced Google Search)	Lexis-Nexis	Newspaper Direct
Argentina	Clarín	http://www.clarin.com/		X
	La Nación	http://www.lanacion.com.ar/		X
Brazil	O Globo	http://oglobo.globo.com/	X	
	O Estado de São Paulo	http://www.estadao.com.br/		
	Jornal do Brasil	http://www.jb.com.br/		
Chile	La Tercera	http://www.latercera.com/		X
	Las Últimas Notícias	http://www.lun.com/		
	El Mercurio	http://www.emol.com/		
Colombia	El Espectador	http://www.elespectador.com/		
	El Tiempo	http://www.eltiempo.com/		
	El Colombiano	http://www.elcolombiano.com/		
Ecuador	El Universo	http://www.eluniverso.com/		
	El Comercio	http://www.elcomercio.com/	X	

Country	News Source	Website (for Advanced Google Search)	Lexis-Nexis	Newspaper Direct
El Salvador	El Diáro de Hoy	http://www.elsalvador.com/		
	La Prensa Gráfica	http://www.laprensagrafica.com/		
	El Mundo	http://elmundo.com.sv/		
Mexico	La Jornada	http://www.jornada.unam.mx		
	Reforma*	http://www.reforma.com/	X	
	Milenio	http://www.milenio.com/		X
Paraguay	ABC Color	http://www.abc.com.py/		
	Ultima Hora	http://www.ultimahora.com/		X
	La Nacíon	http://www.lanacion.com.py/		
Uruguay	El País	http://www.elpais.com.uy/		
	El Observador	http://www.elobservador.com.uy/		
Venezuela	El Universal	http://www.eluniversal.com/		
	El Nacional	http://www.el-nacional.com/	X	
	Ultimas Notícias	http://www.ultimasnoticias.com.ve/		

* Reforma is available only in a print version and thus cannot be searched via Advanced Google Search.

Civil Unrest Event Format

Each GSR record consists of 19 fields, but not all fields are applicable to every event class. GSR fields for CU are defined as follows:

Table 7. Civil Unrest GSR Fields

Civil Unrest GSR Fields
Event ID—Unique identifier for each GSR event.
Event Sub-ID—Encode as 0.
Entry/Revision Date—Date of last GSR update for this event ID.
Record Status: 0 = Active; 1 = Inactive—Encode as 0 unless the event is a city-level event that is part of a larger widespread event. In that case, encode the event ID of the associated widespread event.
Country—Country where the CU event occurred; spelling and diacritic marks conform to the OGDB.
State—For city-level events (01) encode the state where the CU event occurred ensuring that spelling and diacritic marks conform to the OGDB. For widespread events (07), encode a dash (-).
City— For city-level events (01) encode the city where the CU event occurred ensuring that spelling and diacritic marks conform to the OGDB. For widespread events (07), encode a dash (-).
Event Code—4-digit number unique to the type of event; see Table 2 for the list of CU event codes. For city-level events (01) encode violence as discussed in Table 3. Encode widespread events as violent if more than half of the associated city-level events are violent.
Population—one of 11 population types. See Table 4 and Appendix A for more information.
Date—the date the event occurred or, in the case of a multi-day event, the date that the event started, encode in mm/dd/yyyy format.
Earliest Reported Date (ERD)—publication date of the local news source where the event is first reported, encode in mm/dd/yyyy format.
News Source—name of primary GSS. See Table 6 for a list of OSI approved news sources.
Headline—the article headline in the native language.
Event Description— a 2-3 sentence extract from the article that best describes the CU event.
First Reported Link—URL documenting the first report of the CU event.
Other Links (GSS Link)—URL of the report that best describes the CU event.
Other Links—URL for additional news reports. Multiple URLs may be encoded in this field. A semi-colon separates multiple URLs (;).
Crowd Size—analyst's best estimate of crowd size as determined from the article text or accompanying photograph, if available. A crowd size greater than or equal to 50 people is Large, less than 50 is Small, undetermined crowd size is Unknown.

Civil Unrest GSR Fields
Encoding Comment—Optional.

Civil Unrest Warning Format

OSI Performers submit warnings for CU in the following format:
{Warning ID, [Population Type, Event Code, Predicted date, (Country, State, City)], probability} #Optional comment

For example:

{15033, [Education, 0111, 06/24/2014, (Brazil, Rio de Janeiro, Rio de Janeiro)], 1.00}

Table 8. Civil Unrest Warning Fields

Civil Unrest Warning Fields
Warning ID—is Performer defined with a -# (dash number) suffix. Warnings may be updated by the Performer if desired. The first warning submitted has an optional dash zero (-0) suffix, e.g., 12345-0 or 12345. For updates to the first warning, advance the suffix by one (1), e.g., 12345-1. The set of updates for a single warning is called an update sequence.
Population— one of 11 population types. See Table 4 and Appendix A for more information.
Event Code—4-digit number indicating the event type and whether the predicted event will be violent.
Predicted Date—the predicted date of disease onset, encode in mm/dd/yyyy format.
Location Triplet—the (country, state, city) location triplet where the event will occur. For city-level events, each field is populated with a location from the OGDB. For widespread events, only the country is listed; the state and city fields should contain a dash (-). Ensure that spelling and diacritic marks conform to the OGDB.
Probability—a number between 0 and 1, rounded to 2 decimal places, indicating the likelihood of occurrence.
Comment—the Performer may include an optional comment along with the warning submission. The comment text should be on a separate line, right below the warning, starting with the hash or pound symbol, #. Comments may include reasons for sending the warning, additional information about the warning, or information about how this warning relates to other warnings. Each comment is limited to 600 characters including the hash mark (#) at the beginning of the comment, with no line breaks.

Civil Unrest Performance Measures

Two measures are calculated for each CU warning-event (W-E) pair: Quality Score (QS) and Lead Time (LT). The QS is the measure of accuracy of the CU W-E pair, on a scale from 0 to 4.

Four measures (Precision, Recall, Average QS, and Average LT) are calculated for the set of monthly W-E pairs by event code.

LT is the difference between the ground truth ERD and the date/time-stamp indicating when the warning was received. 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.

Performance measures include:

- **Precision:** number of W-E Pairs / Total warnings ingested
- **Recall:** number of W-E pairs / Total number of GSR events
- **Average QS:** the average of the individual QS for all W-E pairs; this is reported for each country and as a summary for all countries.

- **Average LT:** the average of individual LT for all W-E pairs; this is reported for each country and as a summary for all countries.
- **Average Probability Score:** the average of the Brier scores for the probabilities.

Civil Unrest Scoring

Regarding the calculation of CU QS, two distinct scoring approaches were applied during OSI: location scoring (LS) and traditional CU scoring. LS evolved following a series of sensitivity analyses to determine whether any component input to the QS unfairly affected the end result. Following this analysis, in early 2014 the PM decided to use the LS approach, which is described in the following pages.

The following figures 2 through 5 show the QS calculation for each scoring approach including an example for each.

Civil Unrest LS QS Scoring

The LS CU scorer bases the QS on the equally-weighted sum of a "location score" and the event date scores. The location score is based on a user selected maximum distance number. Currently, 300 km is used as the maximum distance between any two locations.

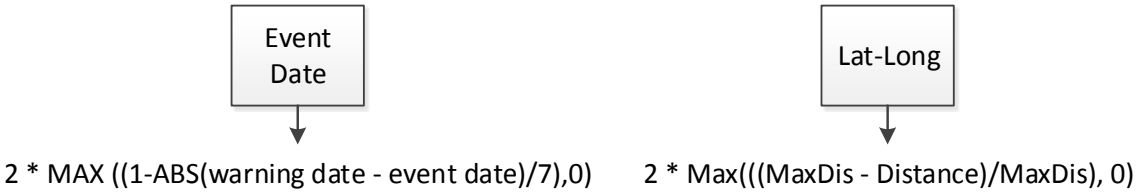


Figure 2. Civil Unrest LS QS Calculation, Maximum Score is 4

An example of the LS CU QS scoring follows:

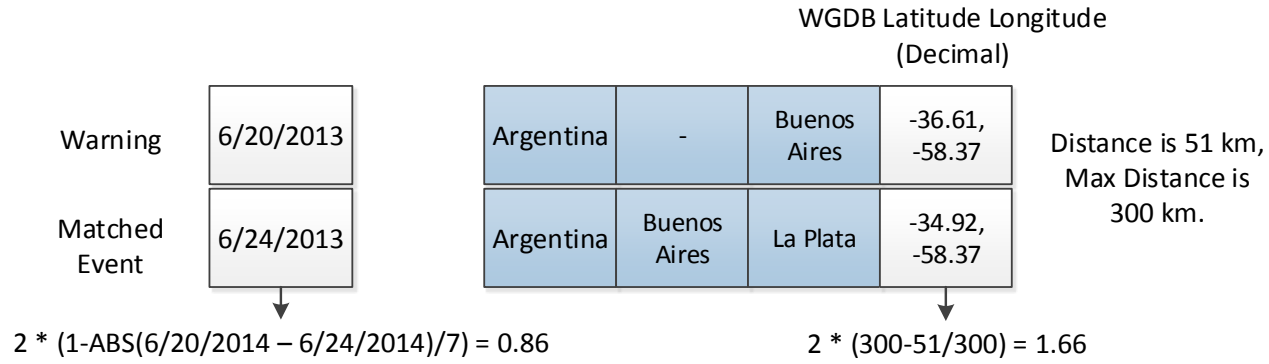


Figure 3. LS Scorer Civil Unrest QS Example

From this example, the date score is 0.86 and the weighted location score is 1.66; the QS for this W-E pair is 2.52.

Civil Unrest Traditional QS Scoring

The Traditional CU scorer bases the QS on a match for all components of the W-E pair: event date, location (scores for country, state, and city), event type (scores for event code and violence), and population.

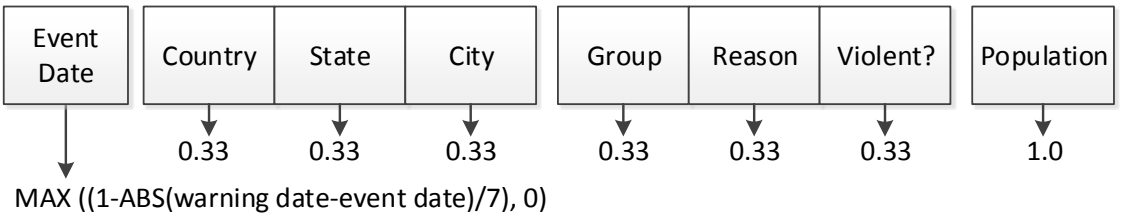


Figure 4. Traditional Civil Unrest QS Calculation, Maximum Score is 4

An example of traditional CU QS scoring follows:

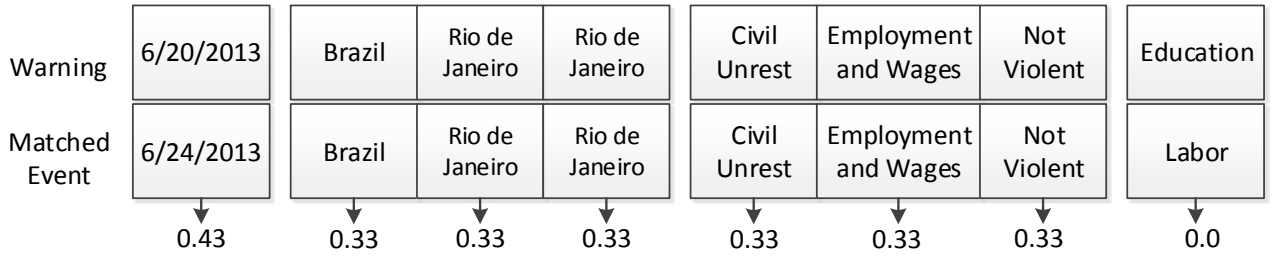


Figure 5. Traditional Scorer Civil Unrest QS Example

For this example, the location score is 1.0, the event score is 1.0, the date score is 0.43, and the population score is 0; the QS for this W-E pair is 2.43.

LT is the difference between the ground truth ERD and the date/time-stamp indicating when the warning was received. If this value is less than one day, the warning is considered late and will not be scored.

2.2 Vote, Event Class (02)

OSI addresses both national and local elections (021) as well as referenda (022). Elections of interest are pre-selected by the OSI PM; see Appendix C for the complete list of OSI elections. (Local elections were included only for major cities and states.)

The following table lists all OSI Vote event types.

Table 9. Vote Event Types

Event Code	Vote Type
0211	President or Prime Minister
0212	Governor
0213	Mayor
0221	Referendum “Yes” vote
0222	Referendum “No” vote

Vote Countries

The following tables list the OSI countries that hold elections, which may be of interest.

Table 10. Vote Event Countries

LA			MENA	
Argentina	El Salvador		Bahrain	Syria
Brazil	Mexico		Egypt	Saudi Arabia
Chile	Paraguay		Iraq	
Colombia	Uruguay		Jordan	
Ecuador	Venezuela		Libya	

Vote Event Format

Each GSR record consists of 19 fields, not all fields are applicable to every event class. GSR fields for Vote are defined as follows:

Table 11. Vote GSR Fields

Vote GSR Fields
Event ID—Unique identifier for each GSR event.
Event Sub-ID—Encode as 0.
Entry/Revision Date—Date of last GSR update for this event ID.
Record Status: 0 = Active; 1 = Inactive—Encode as 0.
Country—Country where the election was held; ensure that spelling and diacritic marks conform to the OGDB.
State—For event types 0212, 0213 (governor, mayor) encode the state associated with the election. For 0221 and 0222 (referendum—yes, no), encode the state only if this is a state- or city-level referendum otherwise place a dash (-) in the field. Ensure that spelling and diacritic marks conform to the OGDB.
City—For event type 0213, encode the city associated with the election. For 0221 and 0222, encode the city only if this is a city-level referendum otherwise place a dash (-) in the field. Ensure that spelling and diacritic marks conform to the OGDB.
Event Code—4-digit number unique to the type of event; see Table 9 for the list of Vote event codes.
Population—For event codes 0211, 0212, and 0213 encode the name of the winning candidate. For event codes 0221 and 0222, encode the ballot question.
Date—date of the election encoded in mm/dd/yyyy format.
Earliest Reported Date (ERD)—publication date of first report of confirmed election results, encoded in mm/dd/yyyy format.
News Source—name of the news source with the first report of confirmed election results.
Headline—headline of the first report with confirmed election results.
Event Description—a 2-3 sentence extract from the first report of confirmed election results that describes the election results.
First Reported Link—URL of the first report of confirmed election results.
Other Links (GSS Link)—Primary ground truth source.
Other Links—URL for supporting information if available.
Crowd Size—Blank for this event code.
Encoding Comment—Optional

Vote Warning Format

OSI Performers submit warnings for each Vote code in the following format:
{Warning ID, [(Name of predicted winner or name of ballot initiative, Event Code, Predicted date, (Location triplet)], probability} #Optional comment

For example:

{15033, [Arthur Neto, 0213, 10/06/2014, (Brazil, Amazonas, Manaus)], 1.00}

{15034, [Amendment to Article 120 of the CP, 0221, 10/09/2011, (Paraguay, -, -)], 1.00}

Table 12. Vote Warning Fields

Vote Warning Fields
Warning ID—is Performer defined with a -# (dash number) suffix. Warnings may be updated by the Performer if desired. The first warning submitted has an optional dash zero (-0) suffix, e.g., 12345-0 or 12345. For updates to the first warning, advance the suffix by one (1), e.g., 12345-1. The set of updates for a single warning is called an update sequence.
Population—For event codes 0211, 0212, and 0213 encode the name of the winning candidate. For event codes 0221 and 0222, encode the ballot question.
Event Code—4-digit number unique to the type of event; see Table 9 for the list of Vote event codes.
Predicted Date—the date of the election, encoded in mm/dd/yyyy format.
Location Triplet—the (country, state, city) location triplet for the election. Encode a dash (-) if the state and/or the city are not relevant. Ensure that spelling and diacritic marks conform to the OGDB.
Probability—a number between 0 and 1, rounded to 2 decimal places, indicating the likelihood of occurrence.
Comment—the Performer may include an optional comment along with the warning submission. The comment text should be on a separate line, right below the warning, starting with the hash or pound symbol, #. Comments may include reasons for sending the warning, additional information about the warning, or information about how this warning relates to other warnings. Each comment is limited to 600 characters including the hash mark (#) at the beginning of the comment, with no line breaks.

Vote Performance Measures

Two measures are calculated for each Vote warning/event pairing: Quality Score and Lead Time.

If the population fields of the warning-event pair (W-E) match, the QS is 4. Otherwise, the QS is 0.

LT is the difference between the ground truth ERD and the date/time-stamp indicating when the warning was received. If this value is less than one day, the warning is considered late and will not be scored.

Other performance measures include:

- **Precision:** number of W-E Pairs / Total warnings ingested
- **Recall:** number of W-E pairs / Total number of GSR events
- **Average QS:** the average of the individual QS for all W-E pairs; this is reported for each country and as a summary for all countries.
- **Average LT:** the average of individual LT for all W-E pairs; this is reported for each country and as a summary for all countries.
- **Average Probability:** the average of the Performer-provided probability statistic.

2.3 Disease, Event Class (03)

Diseases are grouped into five categories based on common warning formats, event formats, and performance measures:

- Rare Disease event classes—see Section 2.3.1
  - 0311—Bolivian Hemorrhagic Fever (LA)
  - 0312—Cholera (LA)
  - 0313—Hantavirus (LA)
  - 0314—Yellow Fever (LA)
  - 0315—Dengue (MENA)
  - 0317—MERS (non-Saudi MENA)
  - 0319—Chikungunya (MENA)
- Chikungunya event class—see Section 2.3.2
  - 0316—Chikungunya (LA)
- Saudi Arabia MERS event class—see Section 2.3.3
  - 036—MERS (Saudi Arabia)
- Polio event class—see Section 2.3.4
  - 0318—Polio
- Disease season event classes (includes Flu, RSV, and Dengue for LA)—see Section 2.3.5
  - 033x - LA Influenza (Flu)
  - 034x - LA Respiratory Syncytial Virus (RSV)
  - 035x - LA Dengue.
  - Additional data for the season event classes includes the following:
    - Start Date—03x1
    - Peak Week—03x2
    - End Date—03x3
    - Peak Week Case Counts—03x4
    - Total Disease Season Case Counts—03x5
- Influenza-like Illness (ILI) event class (Retired in May 2014—033)—see Section 2.3.6

The following table lists all OSI diseases. For each disease, the table presents the OSI event code, the disease name, the associated disease event class and countries, as well as when the disease was included in the OSI disease set. Some diseases were included from program initiation while others were added or retired during the course of the program.

Table 13. Disease Summary

Event Code	Disease	Disease Event Class	Countries	Date of Inclusion
0311	Bolivian Hemorrhagic Fever (LA)	Rare Disease Event	Argentina, Brazil, Chile, Colombia, Ecuador, El Salvador, Mexico, Paraguay, Uruguay, Venezuela	At Program initiation
0312	Cholera (LA)	Rare Disease Event	Same as above	At Program initiation
0313	Hantavirus (LA)	Rare Disease Event	Same as above	At Program initiation
0314	Yellow Fever (LA)	Rare Disease Event	Same as above	At Program initiation
0315	Dengue (MENA)	Rare Disease Event	Bahrain, Egypt, Iraq, Jordan, Libya, Saudi Arabia, Syria	Added 6/2014
0317	MERS (non-Saudi MENA)	Rare Disease Event	Bahrain, Egypt, Iraq, Jordan, Libya, Syria	Added 6/2014
0319	Chikungunya (MENA)	Rare Disease Event	Bahrain, Egypt, Iraq, Jordan, Libya, Saudi Arabia, Syria	Added 6/2014
0316	Chikungunya (LA)	Weekly Case Count – Cumulative Cases	Anguilla, Brazil, Colombia, Dominica, Dominican Republic, El Salvador, French Guiana, Guadaloupe, Saint Barthelemy, Saint Martin, Sint Maarten, Suriname, Venezuela	Added 6/2014
036	MERS (Saudi Arabia)	Weekly Case Count	Saudi Arabia	Added 6/2014
0318	Polio	Polio Event Class	Bahrain, Egypt, Iraq, Jordan, Libya, Saudi Arabia, Syria	Added 6/2014
033x	LA Influenza (Flu)	Season Event Class	Bolivia, Chile, Costa Rica, Ecuador, El Salvador, Guatemala, Mexico, Nicaragua, Panama, Peru	Added 6/2014
034x	LA Respiratory Syncytial Virus (RSV)	Season Event Class	Same as above	Added 6/2014
035x	LA Dengue	Season Event Class	Mexico, Peru, Venezuela	Added 6/2014
033	Influenza-Like Illness (ILI),	Retired – Season Event Class	Argentina, Bolivia, Chile, Colombia, Costa Rica, Ecuador, El Salvador, French Guiana, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru	Retired 5/2014 – replaced with 033x and 034x

2.3.1 Rare Disease Event Classes

OSI Rare Diseases include: Bolivian Hemorrhagic Fever, Cholera, Hantavirus, Yellow Fever, Dengue, Middle East Respiratory Syndrome (MERS), and Chikungunya. An OSI Rare Disease event is the first occurrence of a confirmed disease case within the previous 6 months for a specific location, where the location is defined by the (Country, State, City) location triplet. Cases must be confirmed by an authoritative source such as an official statement from the country's health department.

Rare Disease Countries

The following table shows those countries where Rare Diseases are encoded and the event code associated with each Rare Disease.

Table 14. Rare Disease Countries

Event Code	Disease	Countries
0311	Bolivian Hemorrhagic Fever (LA)	Argentina, Brazil, Chile, Colombia, Ecuador, El Salvador, Mexico, Paraguay, Uruguay, Venezuela
0312	Cholera (LA)	Same as above
0313	Hantavirus (LA)	Same as above
Event Code	**Disease**	**Countries**
0314	Yellow Fever (LA)	Same as above
0315	Dengue (MENA)	Bahrain, Egypt, Iraq, Jordan, Libya, Saudi Arabia, Syria
0317	MERS (non-Saudi MENA)	Bahrain, Egypt, Iraq, Jordan, Libya, Syria
0319	Chikungunya (MENA)	Bahrain, Egypt, Iraq, Jordan, Libya, Saudi Arabia, Syria

Rare Disease Ground Truth Sources

Ground truth sources for Rare Disease are listed in the following table.

Table 15. Rare Disease Ground Truth Sources

Event Code	Disease	Ground Truth Sources
0311	Bolivian Hemorrhagic Fever (LA)	Advanced Google Search for health authority announcement and disease confirmation in local news sources; Program for Monitoring Emerging Diseases (ProMED) (http://www.promedmail.org/)
0312	Cholera (LA)	Advanced Google Search for local sources, ProMED
0313	Hantavirus (LA)	Advanced Google Search for local sources, Chile MOH (http://web.minsal.cl/), ProMED
0314	Yellow Fever (LA)	Advanced Google Search for local sources, ProMED
0315	Dengue (MENA)	Advanced Google Search for local sources, ProMED
0317	MERS (non-Saudi MENA)	Advanced Google Search for local sources, UN Site for Zoonotic Diseases (http://empres-i.fao.org/eipws3g/#h=1), ProMED
0319	Chikungunya (MENA)	Advanced Google Search for local sources, ProMED

Note that ProMED is used only for event validation and only if there is a question regarding the primary ground truth source.

Rare Disease Event Format

Each GSR record consists of 19 fields, not all fields are applicable to every event class. GSR fields for Rare Disease are defined as follows:

Table 16. Rare Disease GSR Fields

Rare Disease GSR Fields
Event ID—Unique identifier for each GSR event.

Rare Disease GSR Fields
Event Sub-ID—Encode as 0.
Entry/Revision Date—Date of last GSR update for this event ID.
Record Status: 0 = Active; 1 = Inactive—Encode as 0.
Country—Country where the disease was contracted; spelling and diacritic marks conform to the OGDB.
State—State where the disease was contracted; spelling and diacritic marks conform to the OGDB.
City—City where the disease was contracted; spelling and diacritic marks conform to the OGDB.
Event Code—4-digit number unique to the type of event; see Table 15 for the list of Rare Disease event codes.
Population—is "General Population".
Date—the date of disease onset. If this date cannot be determined, encode as the earliest reported date; encode in mm/dd/yyyy format.
Earliest Reported Date (ERD)—the publication date of the first official confirmation of the Rare Disease case; encode in mm/dd/yyyy format.
News Source—name of ground truth source. For Rare Disease, this may be a newspaper (e.g., www.eltribuno.info), country level Ministry of Health (e.g., http://web.minsal.cl/), or a PM-approved database (e.g., UN Site for Zoonotic Diseases).
Headline—If the news source is a newspaper, this is the headline. Otherwise, this field is optional.
Event Description—If the news source is a newspaper, this is a 2-3 sentence extract from the article that describes confirmation of the Rare Disease. Otherwise, the field is optional.
First Reported Link—URL documenting the Rare Disease case confirmation.
Other Links (GSS Link)—URL for supporting information if available.
Other Links—URL for supporting information if available.
Crowd Size—Blank for this event code.
Encoding Comment—Optional.

Rare Disease Warning Format

OSI Performers submit warnings for each Rare Disease event code in the following format: {Warning ID, [General Population, Event Code, Predicted date, (Location triplet)], probability} #Optional comment

For example:

{15033, [General Population, 0313, 05/02/2014, (Chile, Maule, Talca)], 1.00}

Table 17. Rare Disease Warning Fields

Rare Disease Warning Fields
Warning ID—is Performer defined with a -# (dash number) suffix. Warnings may be updated by the Performer if desired. The first warning submitted has an optional dash zero (-0) suffix, e.g., 12345-0 or 12345. For updates to the first warning, advance the suffix by one (1), e.g., 12345-1. The set of updates for a single warning is called an update sequence.
Population—is "General Population".
Event Code—as described in Table 15, the Event Code indicates the type of event.
Predicted Date— predicted data of disease onset, encoded in mm/dd/yyyy format.
Location Triplet—the (country, state, city) location triplet where the patient contracted the disease. For Rare Disease, the location triplet must contain the country and state. If no city is indicated, encode the city with a dash (-). Use the location spelling (diacritic marks and capitalization) as listed in the OGDB.
Probability—a number between 0 and 1, rounded to 2 decimal places, indicating the likelihood of occurrence.
Comment—the Performer may include an optional comment along with the warning submission. The comment text should be on a separate line, right below the warning, starting with the hash or pound symbol, #. Comments may include reasons for sending the warning, additional information about the warning, or information about how this warning relates to other warnings. Each comment is limited to 600 characters including the hash mark (#) at the beginning of the comment, with no line breaks.

Rare Disease Performance Measures

Two measures are calculated for each Rare Disease warning/event pairing: Quality Score (QS) and Lead Time (LT). Four additional measures (Precision, Recall, Average QS, and Average LT) are calculated for the set of monthly warning-event (W-E) pairs by event code.

Quality Score (QS)

The QS is the measure of accuracy of the Rare Disease W-E pairing based on a comparison of date, location, and event type. Note that while the maximum result is 3; all QS are normalized to a maximum of 4. The following figure shows how the Rare Disease QS is calculated.

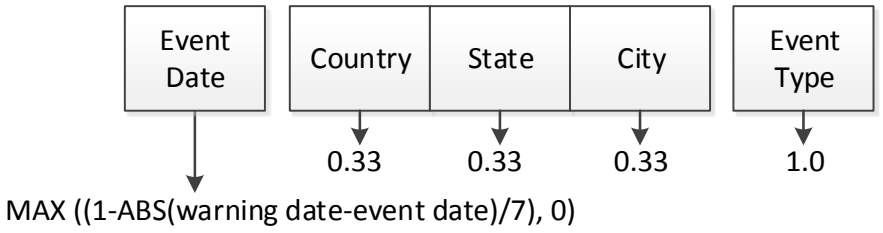


Figure 6. Rare Disease QS Calculation, Maximum Score is 4.

An example of a Rare Disease scoring follows.

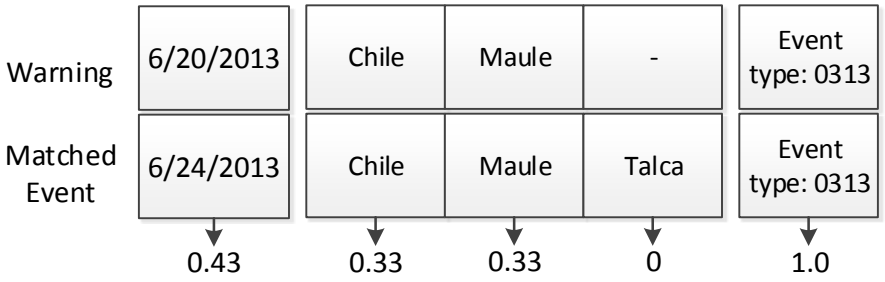


Figure 7. Rare Disease QS Scoring Example

The resulting component score is 2.09; when normalized to a maximum of 4, the QS score is 2.79.

Lead Time (LT)

LT is the difference between the ground truth ERD and the date/time-stamp indicating when the warning was received. 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.

Additional Measures

Other performance measures include:

- **Precision:** number of W-E Pairs / Total warnings ingested
- **Recall:** number of W-E pairs / Total number of GSR events
- **Average QS:** the average of the individual QS for all W-E pairs; this is reported for each country and as a summary for all countries.

- **Average LT:** the average of individual LT for all W-E pairs; this is reported for each country and as a summary for all countries.
- **Average Probability:** the average of the Performer-provided probability statistic.

2.3.2 Chikungunya Event Class

For Chikungunya, cumulative weekly case counts are collected from an authoritative reporting source and are encoded by Epidemiological Week (EW). For OSI, the EW begins on Wednesday and ends on Tuesday; each year consists of 52 EWs. The following table shows an OSI EW extract as an example.

Table 18. OSI Epidemiological Week Dates Sample

OSI EW	First Day of the EW
51	12/17/2014
52	12/24/2014
1	12/31/2014
2	1/7/2015
3	1/14/2015

Chikungunya Countries

Chikungunya countries are listed in the following table:

Table 19. Chikungunya Case Count Disease Countries

Event ID	Disease	Countries	Encoding
0316	Chikungunya (LA)	Anguilla, Brazil, Colombia, Dominica, Dominican Republic, El Salvador, French Guiana, Guadaloupe, Saint Barthelemy, Saint Martin, Sint Maarten, Suriname, Venezuela	Cumulative case count

Chikungunya Ground Truth Sources

Chikungunya ground truth sources are listed in the following table.

Table 20. Chikungunya Ground Truth Sources

Event Code	Disease	Ground Truth Sources
0316	Chikungunya (LA)	PAHO/World Health Organization (WHO) Chikungunya Statistics: (http://www.paho.org/hq/index.php?option=com_topics&view=article&id=343&Itemid=40931)

Chikungunya Event Format

Each GSR record consists of 19 fields, not all fields are applicable to every event class. GSR fields for Chikungunya are defined as follows:

Table 21. Chikungunya GSR Fields

Chikungunya Disease GSR Fields
Event ID—Unique identifier for each GSR event.
Event Sub-ID—Encode as 0.
Entry/Revision Date—Date of last GSR update for this event ID.
Record Status: 0 = Active; 1 = Inactive—Encode as 0.
Country—Country where the disease was contracted; spelling and diacritic marks conform to the OGDB. Chikungunya Diseases are encoded at the country level only.
State—contains a dash (-).
City—contains a dash (-).
Event Code—the Chikungunya event code is 0316.
Population—this field contains the cumulative count of confirmed Chikungunya cases.
Date—the first day of the reported EW, encoded in mm/dd/yyyy format.
Earliest Reported Date (ERD)—the publication date of the first official confirmation of the Chikungunya Disease case.
News Source—name of the ground truth source. For Chikungunya Disease, this may be the PAHO/World Health Organization, UN Site for Zoonotic Diseases, or one of the alternative ground truth sources.
Headline—If the news source is a newspaper, this is the headline. Otherwise, this field is optional.
Event Description—Optional.
First Reported Link—URL documenting the Chikungunya Disease case confirmation.
Other Links (GSS Link)—URL for supporting information if available.
Other Links—URL for supporting information if available.
Crowd Size—Blank for this event code.
Encoding Comment—Optional

Chikungunya Warning Format:

OSI Performers submit warnings for Chikungunya in the following format:

{Warning ID, [Cumulative case Count, Event Code, Predicted date, (Location triplet)], probability} #Optional comment

For example:

{23681, [2258, 0316, 11/05/2014, (French Guiana, -, -)], 1.00}

Table 22. Chikungunya Warning Fields

Warning Fields Definitions
Warning ID—is Performer defined with a -# (dash number) suffix. Warnings may be updated by the Performer if desired. The first warning submitted has an optional dash zero (-0) suffix, e.g., 12345-0 or 12345. For updates to the first warning, advance the suffix by one (1), e.g., 12345-1. The set of updates for a single warning is called an update sequence.
Population—the cumulative weekly case count of confirmed Chikungunya cases.
Event Code—the event code for Chikungunya is 0316.
Predicted Date—the first day of the associated EW (for OSI the first day of the EW is always Wednesday), encoded in the mm/dd/yyyy format.

Warning Fields Definitions
Location Triplet—the country where the disease occurs. Format the location triplet as follows: (country, dash, dash), e.g., (French Guiana, -, -). Use the country spelling (diacritic marks and capitalization) as listed in the OGDB.
Probability—a number between 0 and 1, rounded to 2 decimal places, indicating the likelihood of occurrence.
Comment—the Performer may include an optional comment along with the warning submission. The comment text should be on a separate line, right below the warning, starting with the hash or pound symbol, #. Comments may include reasons for sending the warning, additional information about the warning, or information about how this warning relates to other warnings. Each comment is limited to 600 characters including the hash mark (#) at the beginning of the comment, with no line breaks.

Chikungunya Performance Measures

Two measures are calculated for Weekly Case Count Disease warning-event (W-E) pairs by country per EW, a Quality Score (QS) and Lead-Time (LT). Note that warnings for each EW are only scored against an event for the same EW.

The QS for Case Count Warnings is calculated as follows:

$$\left\{1 - \frac{ABS(Predicted\ Case\ Count - Actual\ Case\ Count)}{MAX(Predicted\ Case\ Count, Actual\ Case\ Count)}\right\} * 4$$

LT is the difference between the ground truth ERD and the date/time-stamp indicating when the warning was received. 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.

Other performance measures include:

- **Average QS:** the average of the individual QS for all W-E pairs; this is reported for each country and as a summary for all countries.
- **Average LT:** the average of individual LT for all W-E pairs; this is reported for each country and as a summary for all countries.

2.3.3 Saudi Arabia Middle East Respiratory Syndrome (MERS) Event Class

For Saudi Arabia MERS, weekly case counts are collected from an authoritative reporting source and are encoded by Epidemiological Week (EW). For OSI, the EW begins on Wednesday and ends on Tuesday; each year consists of 52 EWs. The following table shows an OSI EW extract as an example.

Table 23. OSI Epidemiological Week Dates Sample

OSI EW	First Day of the EW
51	12/17/2014
52	12/24/2014
1	12/31/2014
2	1/7/2015
3	1/14/2015

The Saudi Arabia MERS Event Code is 036.

Saudi Arabia MERS Ground Truth Source

Saudi Arabia MERS ground truth sources are listed in Table 24.

Table 24. Saudi Arabia MERS Ground Truth Sources

Event Code	Disease	Ground Truth Sources
036	MERS (Saudi Arabia)	Primary Source: UN Site for Zoonotic Diseases (http://empres-i.fao.org/eipws3g/#h=1), Backup Sources: Advanced Google Search for local news sources, the Saudi Arabia Ministry of Health http://www.moh.gov.sa/en/CCC/pressreleases/pages/default.aspx?PageIndex=2), WHO Disease Outbreak Notices (http://www.who.int/csr/don/en/), and, ProMED (http://www.promedmail.org/).

Note that ProMED is used only for event validation and only if there is a question regarding the primary ground truth source.

Saudi Arabia MERS Event Format

Each GSR record consists of 19 fields, not all fields are applicable to every event class. GSR fields for Weekly Case Count Disease are defined as follows:

Table 25. Saudi Arabia MERS Disease GSR Fields

Saudi Arabia MERS Disease GSR Fields
Event ID—Unique identifier for each GSR event.
Event Sub-ID—Encode as 0.
Entry/Revision Date—Date of last GSR update for this event ID.
Record Status: 0 = Active; 1 = Inactive—Encode as 0.
Country—Country where the disease was contracted; spelling and diacritic marks conform to the OGDB. Saudi Arabia MERS Diseases are encoded at the country level only.
State—contains a dash (-).
City—contains a dash (-).
Event Code—the Saudi Arabia MERS event code is 036.
Population—this field contains the count of confirmed Saudi Arabia MERS cases.
Date—the first day of the reported EW, encoded in mm/dd/yyyy format. See Table 23 for a list of sample EWs.
Earliest Reported Date (ERD)—the publication date of the first official confirmation of the Saudi Arabia MERS Disease case.
News Source—name of the ground truth source. For Saudi Arabia MERS Disease, this may be the PAHO/World Health Organization, UN Site for Zoonotic Diseases, or one of the alternative ground truth sources.
Headline—If the news source is a newspaper, this is the headline. Otherwise, this field is optional.
Event Description—Optional.
First Reported Link—URL documenting the Saudi Arabia MERS Disease case confirmation.
Other Links (GSS Link)—URL for supporting information if available.
Other Links—URL for supporting information if available.
Crowd Size—Blank for this event code.
Encoding Comment—Optional

Saudi Arabia MERS Warning Format:

OSI Performers submit warnings for Saudi Arabia MERS in the following format:

{Warning ID, [Case Count, Event Code, Predicted date, (Location triplet)], probability}
#Optional comment

For example:

{12748, [7, 036, 11/05/2014, (Saudi MERS, -, -)], 1.00}

Table 26. Saudi Arabia MERS Disease Warning Fields

Warning Fields Definitions
Warning ID—is Performer defined with a -# (dash number) suffix. Warnings may be updated by the Performer if desired. The first warning submitted has an optional dash zero (-0) suffix, e.g., 12345-0 or 12345. For updates to the first warning, advance the suffix by one (1), e.g., 12345-1. The set of updates for a single warning is called an update sequence.
Population—this field contains the weekly case count of confirmed cases.
Event Code—the Saudi Arabia MERS event code is 036.
Predicted Date—the first day of the associated EW (for OSI the first day of the EW is always Wednesday), encoded in the mm/dd/yyyy format. See Table 23 for sample EWs.
Location Triplet—the country where the disease occurs. Format the location triplet as follows: (country, dash, dash), e.g., (Saudi Arabia, -, -). Use the country spelling (diacritic marks and capitalization) as listed in the OGDB.
Probability—a number between 0 and 1, rounded to 2 decimal places, indicating the likelihood of occurrence.
Comment—the Performer may include an optional comment along with the warning submission. The comment text should be on a separate line, right below the warning, starting with the hash or pound symbol, #. Comments may include reasons for sending the warning, additional information about the warning, or information about how this warning relates to other warnings. Each comment is limited to 600 characters including the hash mark (#) at the beginning of the comment, with no line breaks.

Saudi Arabia MERS Disease Performance Measures

Two measures are calculated for Saudi Arabia MERS, a Quality Score (QS) and Lead-Time (LT). Note that warnings for each EW are only scored against an event for the same EW.

The QS for Case Count Warnings is calculated as follows:

$$\left\{1 - \frac{ABS(Predicted\ Case\ Count - Actual\ Case\ Count)}{MAX(Predicted\ Case\ Count, Actual\ Case\ Count)}\right\} * 4$$

LT is the difference between the ground truth ERD and the date/time-stamp indicating when the warning was received. 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.

Other performance measures include:

- **Average QS:** the average of the individual QS for all W-E pairs; this is reported for each country and as a summary for all countries.
- **Average LT:** the average of individual LT for all W-E pairs; this is reported for each country and as a summary for all countries.

2.3.4 Polio Event Class

This event class includes only one disease, Polio. While the GSR includes the number of weekly cases of Polio per country, warnings note the likelihood of Polio occurring in a given Epidemiological Week (EW). For OSI, the EW begins on Wednesday and ends on Tuesday; each year consists of 52 EWs. The following table shows an OSI EW extract as an example.

Table 27. OSI Epidemiological Week Dates Sample

OSI EW	First Day of the EW
51	12/17/2014
52	12/24/2014
1	12/31/2014
2	1/7/2015
3	1/14/2015

Polio Disease Countries

The following table shows those countries for which Polio cases are encoded and the event code associated with the disease.

Table 28. Polio Event Class Countries

Event ID	Disease	Countries
0318	Polio	Bahrain, Egypt, Iraq, Jordan, Libya, Saudi Arabia, Syria

Polio Disease Ground Truth Sources

Ground truth sources for Polio are listed in the following table.

Table 29. Polio Ground Truth Sources

Event Code	Disease	Ground Truth Sources
0318	Polio	Primary Source: WHO/CDC Global Polio Eradication Initiative (http://www.polioeradication.org/Dataandmonitoring/Poliothisweek.aspx) Backup Sources: Advanced Google Search for local sources, WHO Disease Outbreak Notices (http://www.who.int/csr/don/archive/disease/poliomyelitis/en/), and ProMED (http://www.promedmail.org/)

Note that ProMED is used only for event validation and only if there is a question regarding the primary ground truth source.

Polio Event Format

Each GSR record consists of 19 fields, not all fields are applicable to every event class. GSR fields for Weekly Case Count Disease are defined as follows:

Table 30. Polio GSR Fields

Polio GSR Fields
Event ID—Unique identifier for each GSR event.
Event Sub-ID—Encode as 0.
Entry/Revision Date—Date of last GSR update for this event ID.
Record Status: 0 = Active; 1 = Inactive—Encode as 0.
Country—country where the disease was confirmed; spelling and diacritic marks conform to the OGDB. Polio is encoded at the country level only.
State— contains a dash (-).
City— contains a dash (-).
Event Code—the event code for Polio is 0318.
Population—an integer representing weekly confirmed disease case counts.
Date—the first day of the EW where confirmed case count data is available, encoded in mm/dd/yyyy format. If this information is not available, the Date is the Earliest Reported Date (ERD). See Table 27 for a list of sample EWs.
Earliest Reported Date (ERD)—the publication date of the first official confirmation of Polio case(s).
News Source—name of ground truth source. For Polio, this is the WHO/CDC Global Polio Eradication Initiative or one of the alternative ground truth sources.
Headline—If the news source is a newspaper, this is the headline. Otherwise, this field is optional.
Event Description—Optional.
First Reported Link—URL documenting the Polio case confirmation.
Other Links (GSS Link)—URL for supporting information if available.
Other Links—URL for supporting information if available.
Crowd Size— Blank for this event code.
Encoding Comment—Optional

Polio Warning Format:

OSI Performers submit warnings for Polio in the following format:
{Warning ID, [Probability of Occurrence, Event Code, Predicted date, (Location triplet)], probability} #Optional comment

For example:

{12978, [0.917, 0318, 10/01/2014, (Syria, -, -)], 1.00}

Table 31. Polio Warning Fields

Warning Fields Definitions
Warning ID—is Performer defined with a -# (dash number) suffix. Warnings may be updated by the Performer if desired. The first warning submitted has an optional dash zero (-0) suffix, e.g., 12345-0 or 12345. For updates to the first warning, advance the suffix by one (1), e.g., 12345-1. The set of updates for a single warning is called an update sequence.
Population—this field contains a number between 0 and 1 indicating the probability that one or more Polio cases were confirmed during the EW.
Event Code—is 0318.
Predicted Date—the first day in the associated EW (for OSI the first day of the EW is always Wednesday), encoded in the mm/dd/yyyy format, e.g., 02/05/2014. See Table 27 for a list of sample EWs.
Location Triplet—the country where the disease occurs. Format the location triplet as follows: (country, dash, dash), e.g., (Syria, -, -). Use the country spelling (diacritic marks and capitalization) as listed in the OGDB.
Probability—a number between 0 and 1, rounded to 2 decimal places, indicating the likelihood of occurrence.
Comment—the Performer may include an optional comment along with the warning submission. The comment text should be on a separate line, right below the warning, starting with the hash or pound symbol, #. Comments

Warning Fields Definitions
may include reasons for sending the warning, additional information about the warning, or information about how this warning relates to other warnings. Each comment is limited to 600 characters including the hash mark (#) at the beginning of the comment, with no line breaks.

Polio Performance Measures

Two measures are calculated for Polio warning-event (W-E) pairs per country per EW, a Quality Score (QS) and Lead-Time (LT). Note that warnings for each EW are only scored against an event for the same EW.

Although the number of confirmed case counts is encoded in the GSR, for the calculation of QS a non-zero number is treated as a one and zero is treated as zero. The QS for Polio is the Brier score normalized to a maximum of 4.

$$\text{QS is: } 4 * (1.0 - ((Actual\ probability - Predicted\ probability)^{2.0}))$$

LT is the difference between the ground truth ERD and the date/time-stamp indicating when the warning was received. 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.

Other performance measures include:

- **Average QS:** the average of the individual QS for all W-E pairs; this is reported for each country and as a summary for all countries.
- **Average LT:** the average of individual LT for all W-E pairs; this is reported for each country and as a summary for all countries.

2.3.5 Disease Season Event Classes

There are three seasonal OSI diseases: LA Influenza (Flu), LA Respiratory Syncytial Virus (RSV), and LA Dengue. The OSI Disease Season is described by five event codes; these codes are defined in Table 33. As discussed in Table 33, some Disease Season event codes are encoded by Epidemiological Week (EW). For OSI, the EW begins on Wednesday and ends on Tuesday; each year consists of 52 EWs. The following table shows an OSI EW extract as an example.

Table 32. OSI Epidemiological Week Dates Sample

OSI EW	First Day of the EW
51	12/17/2014
52	12/24/2014
1	12/31/2014
2	1/7/2015
3	1/14/2015

Table 33. Disease Season Event Code Definitions

Event Code	Definitions
0331 (Flu) 0341 (RSV) 0351 (Dengue)	Start Date—the Epidemiological Week (EW) indicating the start of the disease season. The steps to determine the Start Date for the current disease season follow: The Rolling Threshold is derived from a minimum of one year of historical data and a maximum of three years of historical data. Determine the Rolling Threshold by calculating the 40th percentile of all observed values for the available historical period (1 to 3 years). Calculate the 40th percentile on a rolling basis by EW. The disease season Start Date is the third consecutive EW with case counts that exceed the Rolling Threshold. As described above, the Start Date is the first day of the EW.
0332 (Flu) 0342 (RSV) 0352 (Dengue)	Peak Week—for the period bounded by the Start and End Dates, this is the EW with the greatest number of disease cases. The Peak Week date is the first day of the associated EW.
Event Code	**Definitions**
0333 (Flu) 0343 (RSV) 0353 (Dengue)	End Date—the EW indicating the end of the disease season. The steps to determine the End Date for the current disease season follow: The End Date uses the same Rolling Threshold as used to determine the Start Date. The disease season End Date is the third consecutive EW with case counts below the Rolling Threshold. Again, the End Date is the first day of the EW.
0334 (Flu) 0344 (RSV) 0354 (Dengue)	Peak Week Case Counts—this is the number of disease cases for the Peak Week.
0335 (Flu) 0345 (RSV) 0355 (Dengue)	Total Case Counts—for the disease season bounded by the Start and End dates, this is the sum of all disease case counts for each EW in the season.

The following figure shows an example of all 5 Disease Season measures for the 2011-2012 Mexico Flu season.

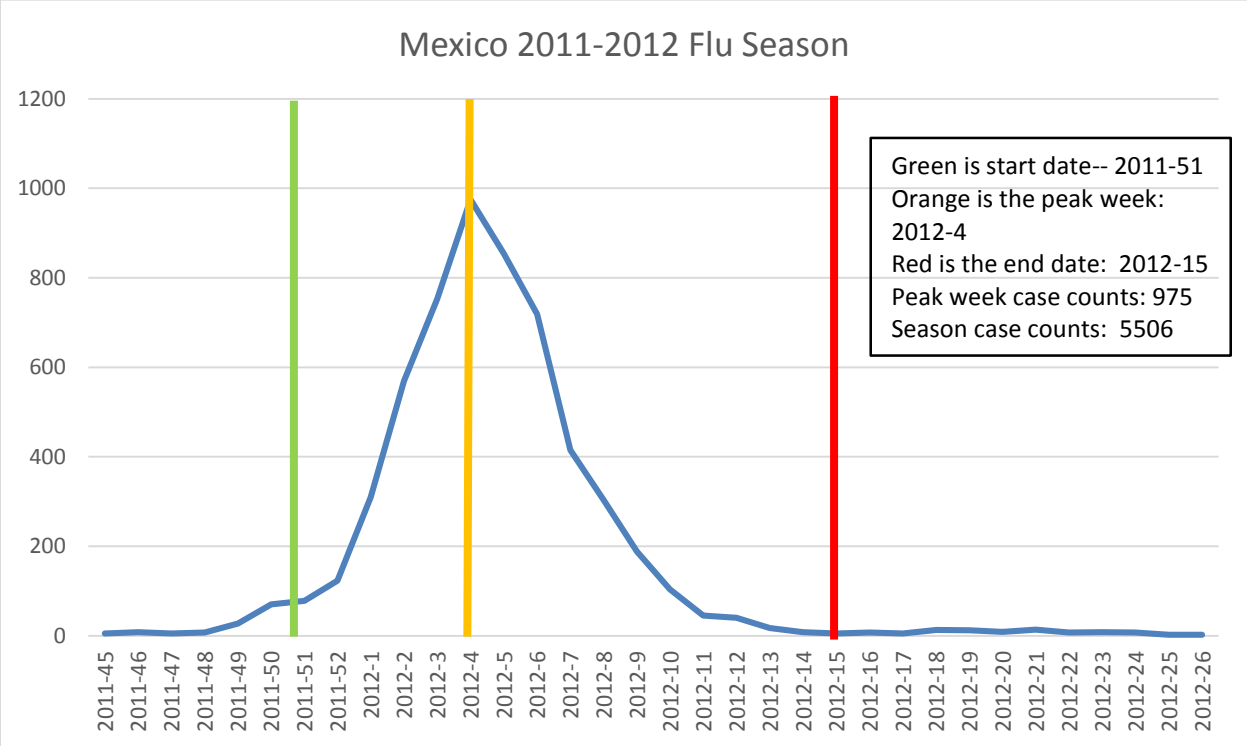


Figure 8. Mexico 2011-2012 Flu Season (Example)

Disease Season Countries

The LA Flu and RSV countries as well as the LA Dengue countries are shown in the following table.

Table 34. Disease Season Countries

Flu and RSV Countries			Dengue Countries
Bolivia	Guatemala		Mexico
Chile	Mexico		Peru
Costa Rica	Nicaragua		Venezuela
Ecuador	Panama		
El Salvador	Peru		

Flu and RSV Ground Truth Data Source

The ground truth data source for all LA Flu and RSV countries is the Pan American Health Organization (PAHO) (http://ais.paho.org/phip/viz/ed_flu.asp). PAHO publishes a database containing weekly influenza-like illness measures for Latin America.

In the PAHO database, the "VSR" column contains RSV case counts by EW per country. The "N_pos_flu" column contains Flu case counts by EW per country. As of 10 November 2014, "N_pos_flu" contains the following flu strains: B Victoria, B Yamagata, B linaje no determinado, FLUAH1, FLUAH3, FLUANoSu, FLUH1N1. Note that the flu strains included in the N_pos_flu case counts may change over time as new flu strains emerge in Latin America.

Dengue Ground Truth Sources

The dengue ground truth source is the Ministry of Health for the given country as shown in Table 2-8.

Table 35. Disease Season Ground Truth Sources

LA Dengue Ministry of Health Ground Truth Sources
Mexico Weekly Epidemiological Bulletin— http://www.epidemiologia.salud.gob.mx/dgae/boletin/intd_boletin.html
Peru Periodic Epidemiological Updates— http://www.app.minsa.gob.pe/bsc/Detalle_IndBSC.asp?lcind=59&lcobj=4&lcper=1&lcfreg=30/10/2014
Venezuela Weekly Epidemiological Bulletin— http://www.bvs.org.ve/php/level.php?lang=es&component=35&item=65

Disease Season Event Format

Each GSR record consists of 19 fields, not all fields are applicable to every event class. GSR fields for disease seasons are defined as follows:

Table 36. Disease Season GSR Fields

Disease Season GSR Fields
Event ID—Unique identifier for each GSR event.
Event Sub-ID—Encode as 0.
Entry/Revision Date—Date of last GSR update for this event ID.
Record Status: 0 = Active; 1 = Inactive—Encode as 0.
Country—Country where the disease was contracted; spelling and diacritic marks conform to the OGDB. Disease Season events are encoded at the country level only.
State—contains a dash (-).
City— contains a dash (-).
Disease Season GSR Fields
Event Code—4-digit number unique to the type of event; see Table 33 for more information.
Population— For Event Codes 03x4 and 03x5, this field is an integer representing weekly confirmed case counts. For Event Codes 03x1, 03x2, and 03x3, this field should contain "NA."
Date—the first day of the EW where confirmed case count data is available, encoded in mm/dd/yyyy format. See Table 32 for a list of sample EWs. For event codes 03x4 and 03x5, the date is the same date as the Disease Season end date, 03x3.
Earliest Reported Date—the date of the first report of case counts per EW-country in the PAHO database. For event codes 03x4 and 03x5, the date is the same date as for the Disease Season end date, 03x3.
News Source—name of ground truth source, e.g., Mexico Ministry of Health BOLETÍN EPIDEMIOLÓGICO
Headline—Blank for this event code.
Event Description—Blank for this event code.
First Reported Link—Blank for this event code.
Other Links (GSS Link)—URL for primary ground truth source.
Other Links—URL for supporting information if available.
Crowd Size— Blank for this event code.
Encoding Comment—Optional.

Disease Season Warning Format:

OSI Performers submit warnings for each disease season event code in the following format:
{Warning ID, [Case Count, Event Code, Predicted date, (Location triplet)], probability}
#Optional comment

There are two types of disease season warnings: Date and Case Count.

Sample Disease Season Date Warnings (03x1, 03x2, 03x3)

{12345-0, [NA, 0331, 12/25/13, (Chile, -, -)], .65}
Comment for warning 12345-0; Start Date for the disease season; Population field is NA.

{12346-0, [NA, 0332, 02/05/2014, (Chile, -, -)], 1.00}
Comment for warning 12346-0; Peak Week for the disease season; Population field is NA.

{12347-0, [NA, 0333, 03/26/2014, (Chile, -, -)], .5}
Comment for warning 12347-0; End Date for the disease season; Population field is NA.

Sample Case Count Warnings (03x4, 03x5)

{12348-0, [2489, 0334, 08/07/2014, (Chile, -, -)], .8}
Comment for warning 12348-0; default to today's date to conform to warning format requirements. This field is not used for scoring.

{12348-0, [10847, 0335, 08/07/2014, (Chile, -, -)], .35}
Comment for warning 12349-0; default to today's date to conform to warning format requirements. This field is not used for scoring.

Table 37. Disease Season Warning Fields

Warning Fields Definitions
Warning ID—is Performer defined with a -# (dash number) suffix. Warnings may be updated by the Performer if desired. The first warning submitted has an optional dash zero (-0) suffix, e.g., 12345-0 or 12345. For updates to the first warning, advance the suffix by one (1), e.g., 12345-1. The set of updates for a single warning is called an update sequence.
Population—for Event Codes 03x4 and 03x5, this field is integer representing weekly confirmed case counts. For Event Codes 03x1, 03x2, and 03x3, this field should contain "NA."
Event Code—4-digit number unique to the type of event; see Table 33 for more information.
Predicted Date—the first day in the associated EW (for OSI the first day of the EW is always Wednesday), in the mm/dd/yyyy format, e.g., 02/05/2014. See Table 32 for a list of sample EWs. (This field is not used for Event Codes 03x4 and 03x5; populate with any date such as today's date.)
Location Triplet—the country where the disease occurs. Format the location triplet as follows: (country, dash, dash), e.g., (Chile, -, -). Use the country spelling (diacritic marks and capitalization) as listed in the OGDB.
Probability—a number between 0 and 1, rounded to 2 decimal places, indicating the likelihood of occurrence.
Comment—the Performer may include an optional comment along with the warning submission. The comment text should be on a separate line, right below the warning, starting with the hash or pound symbol, #. Comments may include reasons for sending the warning, additional information about the warning, or information about how this warning relates to other warnings. Each comment is limited to 600 characters including the hash mark (#) at the beginning of the comment, with no line breaks.

Disease Season Case Count Warnings Performance Measures

Two measures are calculated for a Case Count Warning, a Quality Score (QS) and Lead-Time (LT).

The QS for Case Count Warnings (03x4 and 03x5) is calculated as follows:

$$\left\{1 - \frac{ABS(Predicted\ Case\ Count - Actual\ Case\ Count)}{MAX(Predicted\ Case\ Count, Actual\ Case\ Count)}\right\} * 4$$

For 03x4 (peak week case counts) events, LT is calculated as the difference between the date/time the warning was received and the first day of the ground truth Peak Week. If this value is less than one day, the warning is considered late will not be scored.

For 03x5 events (total disease season case counts), LT is calculated as the difference between the date/time the warning was received and the ERD of the End Date of the season, i.e. the maximum number of disease cases cannot be determined until after the season ends. 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.

Disease Season Date Warnings Performance Measures

Two measures are calculated for Date Warnings, QS and LT.

The QS for Date Warnings (03x1, 03x2, 03x3) is calculated as follows:

$MAX\left(0{,}1 - ABS\left(PredictedEW - \frac{ActualEW}{NMax}\right)\right) * 4$, where Nmax is user-selectable; Nmax was set to 4 for OSI.

NMax is the allowable match window measured in weeks. For example, if Nmax = 4 and PredictedEW-ActualEW = 5, then there is no match. If NMax > End Date - Start Date, then NMax = End Date - Start Date.
LT is the difference between the ground truth ERD and the date/time-stamp indicating when the warning was received. 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.

2.3.6 Influenza-like Illness Event Class (Retired)

The disease event class, 033-Influenza-Like Illness (ILI), was retired from use in May 2014 and replaced with disease season codes 033x (LA influenza) and 034x (LA Respiratory Syncytial Virus (RSV)). This section will address ILI ground truth, encoding and warnings formats, as well as ILI performance measures.

For ILI diseases are encoded by Epidemiological Week (EW). For OSI, the EW begins on Wednesday and ends on Tuesday; each year consists of 52 EWs. The following table shows an OSI EW extract as an example.

Table 38. OSI Epidemiological Week Dates Sample

OSI EW	First Day of the EW
51	12/17/2014
52	12/24/2014
1	12/31/2014
2	1/7/2015
3	1/14/2015

ILI Countries

ILI countries are presented in the following table.

Table 39. ILI Countries

Event ID	Disease	Countries
033	Influenza-like Illness—this event class was retired in May 2014 and replaced with 033x and 034x.	Argentina, Bolivia, Chile, Colombia, Costa Rica, Ecuador, El Salvador, French Guiana, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru

ILI Ground Truth Source

All ILI ground truth was collected from the Pan-American Health Organization (PAHO) Influenza and other Respiratory Viruses under Surveillance, 2010-2014 website (http://ais.paho.org/phip/viz/ed_flu.asp).

ILI Event Format

Each GSR record consists of 19 fields, not all fields are applicable to every event class. GSR fields for ILI are defined as follows:

Table 40. ILI GSR Fields

ILI GSR Fields
Event ID—Unique identifier for each GSR event.
Event Sub-ID—Encode initial event a 0; increment for subsequent updates.
Entry/Revision Date—Date of last GSR update for this event ID.
Record Status: 0 = Active; 1 = Inactive—Encode as 0.
Country—Country where the disease was contracted; spelling and diacritic marks conform to the OGDB. ILI events are encoded at the country level only.
State—contains a dash (-).
City—contains a dash (-).
Event Code—033
Population—an integer representing weekly confirmed case counts.
Date—the first day of the EW where confirmed case count data is available, encoded in mm/dd/yyyy format. See Table 38 for a list of sample EWs.
Earliest Reported Date— the date of the first report of case counts per EW-country in the PAHO database.
News Source—PAHO
Headline—Blank for this event code.
Event Description—ILI EW Case Count
First Reported Link—http://ais.paho.org/phip/viz/ed_flu.asp
Other Links (GSS Link)—Optional
Other Links—Optional
Crowd Size— Blank for this event code.
Encoding Comment—Optional.

ILI Warning Format:

OSI Performers submit warnings for each disease season event code in the following format:
{Warning ID, [Case Count, Event Code, Predicted date, (Location triplet)], probability}
#Optional comment

For example:

{12748, [34, 033, 02/12/2014, (Venezuela, -, -)], 1.00}

Table 41. ILI Warning Fields

Warning Fields Definitions
Warning ID—is Performer defined with a -# (dash number) suffix. Warnings may be updated by the Performer if desired. The first warning submitted has an optional dash zero (-0) suffix, e.g., 12345-0 or 12345. For updates to the first warning, advance the suffix by one (1), e.g., 12345-1. The set of updates for a single warning is called an update sequence.
Population—an integer representing weekly case counts.
Event Code—033
Predicted Date—the first day in the associated EW (for OSI the first day of the EW is always Wednesday), encoded in the mm/dd/yyyy format, e.g., 02/05/2014. See Table 38 for a list of sample EWs.
Location Triplet—the country where the disease occurs. Format the location triplet as follows: (country, dash, dash), e.g., (Argentina, -, -). Use the country spelling (diacritic marks and capitalization) as listed in the OGDB.
Probability—a number between 0 and 1, rounded to 2 decimal places, indicating the likelihood of occurrence.
Comment—the Performer may include an optional comment along with the warning submission. The comment text should be on a separate line, right below the warning, starting with the hash or pound symbol, #. Comments may include reasons for sending the warning, additional information about the warning, or information about how this warning relates to other warnings. Each comment is limited to 600 characters including the hash mark (#) at the beginning of the comment, with no line breaks.

ILI Performance Measures

Two measures are calculated for Weekly Case Count Disease warning-event (W-E) pairs per country per EW, a Quality Score (QS) and Lead-Time (LT).

The QS for Case Count Warnings is calculated as follows:

$$\left\{1 - \frac{ABS(Predicted\ Case\ Count - Actual\ Case\ Count)}{MAX(Predicted\ Case\ Count, Actual\ Case\ Count)}\right\} * 4$$

LT is the difference between the ground truth ERD and the date/time-stamp indicating when the warning was received. 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.

Other performance measures include:

- **Precision:** number of W-E Pairs / Total warnings ingested
- **Recall:** number of W-E pairs / Total number of GSR events
- **Average QS:** the average of the individual QS for all W-E pairs; this is reported for each country and as a summary for all countries.
- **Average LT:** the average of individual LT for all W-E pairs; this is reported for each country and as a summary for all countries.
- **Average Probability:** the average of the Performer-provided probability statistic.

2.4 Integrated Crisis Early Warning System (ICEWS)

ICEWS is a data repository developed by the Defense Advanced Research Projects Agency (DARPA) and operated by Lockheed Martin. The ICEWS repository contains geopolitical events automatically encoded from news stories published by over 6000 international, national, regional, and local news sources; the ICEWS system automatically parses and then encodes news stories following the Conflict and Mediation Event Observations (CAMEO) taxonomy. In addition, ICEWS employs a statistical method to aggregate news stories into "Events of Interest" such as Domestic Political Crisis (DPC).The ICEWS repository includes encoded news stories beginning with January 2001.

There are two OSI ICEWS-based event classes:

- Domestic Political Crises, 05—ICEWS uses a statistical method to determine, each month, whether a country has experienced a DPC.
- Civil Unrest, 06—based on 014 (Protest), 017 (Coerce), and 018 (Assault) CAMEO categories and subcategories.

ICEWS Countries

Table 42. ICEWS Countries

LA			MENA	
Argentina	El Salvador		Bahrain	Syria
Brazil	Mexico		Egypt	Saudi Arabia
Chile	Paraguay		Iraq	
Colombia	Uruguay		Jordan	
Ecuador	Venezuela		Libya	

ICEWS Ground Truth and Event Format

Lockheed Martin provides OSI with access to the ICEWS 05 and ICEWS 06 data set. This information is extracted into OSI ICEWS database tables, which MITRE updates as new ICEWS data are provided by Lockheed Martin.

ICEWS events are not encoded in the OSI GSR; scoring is done directly from the OSI ICEWS database tables. Note the ERD for an ICEWS 06 event is the day that data are posted by Lockheed Martin. For an ICEWS 05 event, the ERD is the last day of the month.

ICEWS Warning Format

OSI Performers submit warnings for each ICEWS event code in the following formats:

ICEWS 05—Domestic Political Crisis

{Warning ID, [Probability of DPC, 05, Predicted date, (Location triplet)], probability} #Optional comment

ICEWS 05 Sample Warning:

{V9731, [.7, 05, 10/15/2013, (Mexico, -, -)], 1.00}

ICEWS 06—Civil Unrest

{Warning ID, [Forecasted number of protest events, 06, Predicted date, (Location triplet)], probability} #Optional comment

ICEWS 06 Sample Warning:

{12748, [34, 033, 02/12/2014, (Venezuela, -, -)], 1.00}

Table 43. ICEWS Warning Fields

Warning Fields Definitions
Warning ID—is Performer defined with a -# (dash number) suffix. Warnings may be updated by the Performer if desired. The first warning submitted has an optional dash zero (-0) suffix, e.g., 12345-0 or 12345. For updates to the first warning, advance the suffix by one (1), e.g., 12345-1. The set of updates for a single warning is called an update sequence.
Population—for ICEWS 05 this is the Probability of DPC; for ICEWS 06 this is the Forecasted number of events.

Warning Fields Definitions
Event Code—05 or 06
Predicted Date—for 05 this is the 15th of every month, i.e. mm/15/yyyy. For 06 this is the first day of the associated ISO week, e.g., 11/27/2014.
Location Triplet—ICEWS 05 and 06 are encoded at the country-level only. Format the location triplet as follows: (country, dash, dash), e.g., (Argentina, -, -). Use the country spelling (diacritic marks and capitalization) as listed in the OGDB.
Probability—set to 1 for both ICEWS 05 and 06. (The Performer does not forecast a probability for this type of warning.)
Comment—the Performer may include an optional comment along with the warning submission. The comment text should be on a separate line, right below the warning, starting with the hash or pound symbol, #. Comments may include reasons for sending the warning, additional information about the warning, or information about how this warning relates to other warnings. Each comment is limited to 600 characters including the hash mark (#) at the beginning of the comment, with no line breaks.

ICEWS Ingest

MITRE updates the OSI ICEWS database when Lockheed Martin sends a notification that new data is available.

ICEWS Performance Measures

Note that:

1. ICEWS is scored for each country and for each capital.
2. Special capital city areas are used for Brazil (Distrito Federal and Federal District), Egypt (Tahrir Square and Cairo), and Mexico (Distrito Federal and Mexico City).
3. ICEWS 06 warnings are scored weekly; ICEWS 06 warnings are scored monthly.

ICEWS 05

The Quality Score (QS) is the Brier score normalized to a maximum of 4.

$$\text{QS is: } 4 * (1.0 - ((Actual\ DPC\ probability - Predicted\ DPC\ probability)^{2.0}))$$

Average QS: the roll-up average of all country QS scores.

ICEWS 06

QS is determined as follows:

- Actual = MAX(Actual event counts, 0)
- Predicted = MAX(Predicted event counts, 0)
- Occurrence_Score = 1 if Actual and Predicted are both zero or both non-zero, 0 otherwise.
- Accuracy_Score = 1 – ABS(Actual – Predicted)/MAX(Actual, Predicted, 4)
- QS = 0.5*Occurrence_Score + 3.5*Accuracy_Score

LT is the difference between the ground truth ERD and the date/time-stamp indicating when the warning was received. 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.

Other performance measures include:

- **Average QS:** the average of the individual QS for all W-E pairs; this is reported for each country and as a summary for all countries.
- **Average LT:** the average of individual LT for all W-E pairs; this is reported for each country and as a summary for all countries.

3 Bibliography

[1] J. B. C. O. J. J. C. T. D. Bond, "Integrated Data for Events Analysis (IDEA): An Event Typology for Automated Events Data Development," *Journal of Peace Research,* vol. 40, no. 6, pp. 733-745, 2003.

[2] P. Schrodt, "Automated Production of High-Volume, Real-Time Political Event Data," in *American Political Science Association,* 2010.

[3] H. W. Kuhn, "The Hungarian Method for the Assignment Problem," *Naval Research Logistics Quarterly,* pp. 83-97, 1955.

Appendix A Civil Unrest Encoding Guidance and Keywords

A.1 Civil Unrest Event Identification

A Civil Unrest (CU) event is any mention of a demonstration, protest, or strike in support of or in opposition to a common cause (e.g., economic, political, societal) as reported in the CU ground truth news media (see Table 6, Section 2). CU events may also include positive events such as donation drives or religious marches.

Example of a Positive CU Event: A natural disaster took place that left people without homes/shelter. A few days later, a donation drive was organized to assist people affected by the disaster. Although not completely spontaneous, the event was scheduled and organized in only a few days, more than a 1,000 people attended the event (the event itself was basically a festival, with food and music and where people would bring in the donations). This event was encoded as General Population, Other Civil Unrest.

To be considered a CU event, more than one person must be involved in the event. For example, a hunger strike by one person would not count as an event but if two people are undertaking a hunger strike together then it would be considered a CU event.

The news media source must indicate the date when the event took place or, if it is an ongoing event, it must indicate the date when the event began. If the article's publication date cannot be verified, the event cannot be encoded. An event date may be verified by a non-GSR media source, which would then be listed in the "other links" GSR field.

Events that are not considered CU include:

- Reports of general instability within a country, natural disasters, scandals, accidents, and criminal events.
- A civic group that is against some government policy and, instead of taking to the streets to protest, actually meets with government officials to discuss their concerns.
- A future protest that is in the planning stages or a protest that is about to happen - a good indication of a future event is the use of the future tense in the news article.
- An event that is prevented from taking place.
- A one-person protest.

Example of a non-CU Event: An event did not take place because the protesters were stopped by the National Guard before they could get to the place where they planned to protest. Workers from Puerto Cabello either drove together or separately (the article does not specify) towards Caracas (the place where the protest was going to take place), however the National Guard set up a checkpoint in the road (somewhere in between Puerto Cabello and Caracas). When the protesters got to the checkpoint, they were told to go home (no mention if they were forced to do so). Therefore, the event did not occur.

A.1.1 What Constitutes the Boundaries of an Event

The temporal granularity of an event is one day with most events taking place during a single 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 is

"occupied" for several days, it would be unlikely that the occupation would end one day and resume the next. In this case, the occupation would be treated as one event.

Situations that would be encoded as a single event include the following:

- An event that starts stops and starts again within the same day is still the same event.
- An event that moves from one location to another within the same day is still the same event.
- An event that 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 that changes from non-violent to violent the same day, it is one event and is encoded as violent.
- Events that are continuous over the course of multiple days, such as a 48-hour strike, should be encoded as one event based on its onset date.

A.1.2 How to Distinguish Continuous Events from Multiple Events

In some situations, an ongoing event is encoded as multiple events.

- The article mentions that the event stopped and then started again on another day.
- The event turned violent after the first day. The day on which the event turns violent becomes a separate event.
- A continuous event is punctuated by specific acts of protest.
- A counter protest begins after the first day of an event.

Example of Continuous versus Repeated Event: An article may say that a building was occupied on the 3rd of July and a subsequent article may say that the same building was occupied on the 4th of July by the same group of people. However, neither article specifies whether the building was occupied continuously for 2 days. In the absence of any other information, the event would be treated as being continuous and encoded only once using the 3rd of July as the event date. However, if one of the articles mention that the building was held for 3 hours, then it is clear that the protesters stopped at night and returned the next day. In this case, the occupation on the 4th of July is encoded as a separate event.

Example of Change in Violence: An event that goes on for several days and remains non-violent for the first few of days, but later becomes violent. This would be encoded as two events; one encoding for the days that it was non-violent and a second encoding for when it becomes violent. For example, students peacefully occupy a university building. On the 3rd day of the occupation, the police try to dislodge the students; the students repel the police by throwing Molotov cocktails at the police. The students continue to occupy the building for several more days. The initial building occupation is encoded as non-violent. However, the day of the police action and extreme response by the students is encoded as a new violent event.

Example of Continuous Protest with Intermittent Acts of Protest: On the 3rd day of a strike, employees march through the center of town holding signs stating their grievance. The initial "strike" event would be encoded with the date on which the strike began and the location of the town where the strikers are employed. A second encoding for the "march" event would have the date and location where the march took place and the size (large, small, or unknown) of the number of people in the march. Most likely, the location of the two events will be the same but

not necessarily. Conversely, if an article refers to the 3rd day of a strike with no mention of an additional form of protest, it is not a new event. The link to that article should be added to the "Additional Links" of the strike event encoded on the first day of the strike.

Example of Counter Protest: There is a student protest that is ongoing and has lasted for a few days. The protesters are blocking the entrance to their university - resulting in the suspension of classes. On the 4th day of the protest, another group of students that wishes to attend class begins to protest against the protesters (a counter protest). Since the counter protest started on a different day than the original protest, a separate event should be encoded.

A.1.3 Events with Multiple Encodings

Some events require multiple event encodings given the following circumstances:

- More than one population is participating in the event.
- The protest is occurring concurrently in more than one city.
- The event is Widespread – occurring simultaneously in cities throughout a region, state, or the entire country.

Multiple Populations: One event would have separate encodings for each population group that participates in the same event. For example, a protest in which residents of a neighborhood, construction workers, and students all protest the lack of reliable public transportation would be encoded as three events, one for each population (general population, labor, and education).

That said, for a population to deserve its own event encoding, there must be two or more individuals involved in the event, e.g., if one priest joins the aforementioned protest, there should not be a separate encoding for "Religious," but if a priest and a bishop join the protest, there should be a separate encoding.

Multiple Locations: A protest may take place in more than one location at the same time. For example, people protesting the failure of the government to repair damage to their community after a major flood demonstrate in the state capital while another group blocks the highway near their village. This would be encoded as two events.

If, on the other hand, a protest moves from one city to another on the same day then this is not two events – only the location of the first city would be encoded.

Widespread Events: A large number of protests across a geographic region characterized by protestors assembling for a common, specific reason (as indicated by event type) on the same day. A widespread event may not necessarily be centrally orchestrated. For example, during June 2013 in Brazil—student protests against bus fare hikes occurred in all major metropolitan areas; here there was a common reason but no central organization of the protests across the country.

The news article may clearly state that the event is widespread. However, in some cases the article may not explicitly describe the protest as widespread. Widespread events are often associated with a day of special significance, such as Labor Day, or the anniversary of an historic or tragic event. The event must take place in more than one city on the same day. For example, teachers protesting in several locations around the capital would not be a widespread event. However, teachers protesting across a state or in several cities would be a widespread event. Regardless of whether the widespread events are region-wide (e.g., coastal region), state, or countrywide—all are encoded the same way. If the article explicitly states that the event is

widespread, a note is added to the GSR comments field indicating the region/state specified in the text.Encode a Widespread Event as follows:

- Encode separate events for all the cities listed in the article (i.e. part of the widespread event). Include the widespread event ID in the "Record Status" GSR field for each associated city-level event. If an article says there were multiple protests across the region, but no specific cities are listed, encode only the widespread event with the appropriate event code and "-" (dashes) in the city and state fields.
- Encode the associated "widespread" event without specifying the state or City by putting a dash (-) in state and city fields.
    - Encode the primary population and event type. If there are multiple populations, add a clarifying comment.
    - Encode Size as "large".
    - Encode Violent only if there is violence in over half of the associated cities or the article mentions it was a violent widespread event.
    - Encode the event with the 07 event code.
    - Include all URLs associated with the city-level events in the "Other Links" GSR field.

A.1.4 Identifying Crowd Size

Crowd size is encoded as Very Large—if 1,000 or more people reported; Large—if 50-999 people reported; Small—if less fewer than 50 people reported; or Unknown—if crowd size cannot be determined. Crowd size is identified in two ways:

- The news article states the number of people involved. (Note "dozens" is encoded as Large.)
- A photo of the event shows a crowd size greater than 50 people. There are two caveats with respect to using photos to determine crowd size. First, the photo must be associated with the event (i.e. it must not be a stock or archived photo). Second, a photo of a small group of people is not sufficient to determine that the crowd size is Small – it is possible that not everyone participating in the event was captured in the photo.

A.1.5 Identifying Event Date

Most of the media sources used for the GSR are online which enables a reporter to post an article on the same day as the event occurred. However, when encoding from a printed news source (e.g., NewspaperDirect or Lexis-Nexis), the event date may be prior to the date of publication.

In the best case scenario, the text of the article will clearly state when the event happened—as indicated by words like "today", "yesterday", a particular day of the week, or a date. Also, evaluate the date of publication. Often articles will have the exact time the article was posted which must also be taken into account. 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 mentions that the event took place today, then checking the hour of publication is critical. If the article does not otherwise mention when the event took place and it is print media, assume that that it took place the day before the article was published. If it is an online source and there is no time stamp, assume it took place on the date of publication.

A.1.6 Identifying Earliest Reported Date

The Earliest Reported Date (ERD) is the publication date of the news source that first reported the event. If there is no publication date associated with the article, the article may be used for information only and may not be used as the primary source for event encoding.

A.1.7 Identifying Locations

All locations are defined by a location triplet (Country, State (Province, department, etc.), City). An article may refer to a place name by an abbreviated or colloquial form (e.g., "Los Angeles" is commonly referred to as "LA"; Tegucigalpa" is commonly referred to as "Tegus"). In addition, the spelling of place names may vary. To ensure uniformity of location names, the OSI Gazetteer Database (OGDB) is the location standard for OSI.

- Be sure to verify place names in the OGDB to ensure proper spelling and correct use of diacritic marks for all city and state names. Use the full name of the place as it appears in the OGDB (e.g., Ciudad de México).
- If the location is not in the OGDB, encode the event at the nearest city (within 50 kilometers) that is in the OGDB. If a nearby location cannot be determined or if the article does not identify the city, encode the state and place "NA" in the city field. Make a note in the comment field.
- Some cities have the same name as their administrative district, e.g., Caracas, Caracas; a few cities form their own administrative unit and are not part of a state, e.g., Buenos Aires. In either case, follow the OGDB standard.
- An article may have a city name in the lead of the article. Although this may correspond to the event location, this is not always the case. 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 on a particular street, then perform an Internet search to verify the location of the building or street.
- Embassy grounds are treated as being internal to the country in which they reside, e.g., encode the geographic location of the embassy.

A.1.8 Identifying Violence

Violence is defined as a physical confrontation between people or damage done to property. If no judgment can be made whether the event was violent or not, then encode the event as non-violent.

Table 44. Civil Unrest Violence Indicators (from Section 2.1)

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

Violent	Non-Violent
• Throwing hard objects which could cause injury or damage • Hitting with sticks, bars, machetes, etc.	• Throwing things that would not cause harm (e.g., eggs) • Brandishing sticks, bars, 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

A.2 Civil Unrest Event Reason Encoding Guidance

Event Reasons give insight into the common cause against which (or in support of which) a group is protesting, e.g., government, political, religious. What follows are tips for correctly encoding Event Reasons, and a table with examples of each of the six categories.

1. If more than one reason is given for a protest, try to determine which is the primary reason. For example, in Mexico, teachers protested the disappearance of a large group of students that had been detained by local police at the mayor's behest. In addition, at the same event, there was a small subset of teachers who were protesting for better salaries. In this case, the primary reason for the protest is clearly "Other Government and Political Issues"—not "Employment and Wages."

2. If it is not possible to determine a primary reason, use a top-down approach. The goal is to encode as specific a category as possible. The category "Other Civil Unrest" should only be used when none of the other reasons applies. Only a few events should fall into the "Other Civil Unrest" category.

   Begin at the top of the list of event reasons—the first category is "Employment and Wages." Question whether the reason falls into this category. If yes, encode the event as such; if not, evaluate the next reason; and continue with this approach until the appropriate reason is found.

   Example of Top-Down Event Reason Encoding:

   > *Residents protest the lack of government support while their neighborhoods undergo major renovation. They demand the city provide housing relocation, pay commercial rates for property owners, and provide soft loans for local businesses to stay afloat during construction. There are several aspects to this protest: Housing and Shelter; Land, Energy, and Resources; Other Economic Issues; and, Other Government and Political Issues. Using the "top down" approach, the appropriate event reason is Housing and Shelter.*

   In this example, people were protesting the scarcity of food due to the economic crisis in Venezuela, which could be construed as a land/resource, economic, or government issue—the Top-Down results in the selection of "Land, Energy and Resource."

3. If the event does not fit into event categories 1 through 5, then it may be encoded as the sixth category, "Other Civil Unrest." This should be used only when none of the other categories applies.

4. An encoding guidance table follows.

Table 45. Civil Unrest Event Reason Encoding Guidance

Event Reason Encoding Guidance	
Civil Unrest Event Reasons	***Examples***
Employment and Wages • Lack of employment • Unemployment rates • Working conditions • Loss of jobs • Poor job growth • Loss of wages • Other forms of compensation such as employee rights, employer rights, benefits • Retirement benefits • Bonuses	• The February 2012 strike by the Bolivian Medical Council to reject the government's increase of their work day without compensation • Homeless people protesting for better jobs • Politicians who go on strike for better retirement benefits • Prisoners rioting for better pay for prison labor • Retirees protesting for increase in their pension • Workers in Mexico protesting against government labor reforms • Ambulance drivers in Paraguay protesting against the privatization of ambulance service in the country
Housing and Shelter • Home ownership, rental, lease, or other formal contractual arrangements • Public housing • Temporary shelter • Housing conditions that lead to civil unrest • Prison living conditions • Evictions from homes (not raw land) • Relocations • Eminent Domain Issues involving homes (not raw land)	• The October 2012 protest by residents of El Sito in Panamá City against their eviction • Peasants protesting against eviction from their homes ("viviendas") • Refugees marching to city hall demanding that the government provide them with houses

Event Reason Encoding Guidance	
Civil Unrest Event Reasons	***Examples***
Land, Energy and Resources • Restriction on use or lack of availability of energy sources such as, gasoline, heating oil, natural gas, water, and electricity • Issues with public/private utilities • Cost of energy resources • Cost of resources and utilities • Lack of community resource such as social services • Lack of community services such as trash pickup • Lack of natural resources such as coal, oil, water, forests, minerals • Lack of health resources such as services and materials for health and mental welfare • Loss of ownership of property (not homes) or mineral rights • Environmental issues related to land or some other resource • Land protests • Eviction from land (not homes) • Development of infrastructure such as roads, dams, and bridges	• The March 2012 protests in Honduras over a hike in oil prices • Residents protest the construction of an oil pipeline • Protests for city bus service • Protests against an increase in the cost of city bus service • Protests against the construction of a bridge • Residents protesting for access to better healthcare • Vendors protesting their eviction from public land (e.g., city streets) • Peasants protesting against eviction from their land ("tierras") • Residents protesting against the transformation of a local park into a parking lot • Townspeople protesting against a factory dumping waste into a local lake • 300 indigenous people block roads in protest of deforestation of lands they claim belonged to their ancestors • Clean-up following storms
Other Economic Issues Government policies focused on the management of the nation's economy—this category includes both anti- and pro-government policies: • Control of monetary system • Control of trade • Fees, taxes, business rights, parking meters • Public budget issues • Subsidies • Economic development funds • Damage compensation paid by the state • Fishing quotas	• The September 2011 farmers protest against onion imports in Quibor, Venezuela • A group of students protesting tax reforms that will impact education costs • Farmers protesting against cuts in agricultural subsidies • Local residents protesting the installation of parking meters in their neighborhood

Event Reason Encoding Guidance	
Civil Unrest Event Reasons	***Examples***
Other Government and Political Issues • Government policies • Government mandates • Government regulations • Demands for increased safety or security • State funerals • Elections including elected officials and candidates • Political rallies or marches • Events involving past, current, or future political figures (those running for office) • Events involving government officials • Events involving protests related to public institutions or agencies (university, police, fire department) • Demands for release of detainees • General prison riots (unless a more specific reason is given) • Environmental issues (not resource or land related) directed towards the government or some government policy (ex. environmental policies). • Animal rights issues directed towards the government or some government policy • Mismanagement of government funds or programs	• The October 2011 nationwide strike by students in Chile calling for education reform • Protests demanding new safety laws for taxi drivers • Protests against bullfighting (bullfighting concerns government laws and regulations) • Residents protesting against increase in crime or protesting for better security • Riots at polling stations • Protests demanding that the government bans the use of pesticides that are harmful to the environment • Protests demanding that the government creates stricter laws against animal abuse • Students protesting against the dean of a public university • Protest of a government figure or appointee who mismanages programs or funds
Other Civil Unrest • Religious marches or demonstrations • Cultural protests or marches • Social protests or marches • Marches or demonstrations for peace • Awareness marches and demonstrations • Events involving protests related to private institutions (private schools, private universities, private companies) • Animal rights and environmental issues (not resource or land related) directed towards a private company or institution (not involving the government) • Protests against foreign governments/agencies • Any event where the reason cannot be determined, or that does not fit into any other event reason	• The May 2012 protest in Honduras's Mosquito Coast region against the DEA's role in fatal shootings (DEA is a US government agency) • Protests against the price of soccer tickets • Activists protesting against a makeup company using animals for testing • Women marching to raise awareness of violence against women • Protest against the British government in front of the British embassy in Buenos Aires

A.3 Civil Unrest Population Encoding Guidance

1. One event entry should be created for each population that takes part in an event. For example, if there is a protest in which the residents of a neighborhood, construction workers, and students all participate in the event, the protest is encoded as three separate events, one for each population (general population, labor, and education).
2. If the event population does not fit into any of the specific population categories or the population is not specified within the article, then it should be encoded as "General Population."
3. If a population falls under the "Labor" category but at the same time falls under one of the other more specific categories, then it should be encoded using the more specific category that best reflects the group's profession. For example, teachers, indigenous workers, attorneys, doctors, and agricultural workers asking for better pay are all "Labor" related populations. However, instead of encoding them all as "Labor", they should be encoded separately as Education, Ethnic, Legal, Medical, and Agricultural. Administrative and support staff for educational, medical, and legal institutions are generally included in these specialized fields unless the article states that only support staff are protesting in which case the population is encoded as "Labor."
4. If the population falls under more than one population category (not including "Labor" or "General Population"), pick the category that is most relevant to the event. For example, indigenous students protesting for indigenous student scholarships would warrant a judgment call and would most likely be encoded as Ethnic, since the protest is specifically for scholarships for indigenous students. *Note that this case is different from an event with 2 population groups protesting together for the same cause, which would be encoded as 2 separate events.

Table 46. Civil Unrest Population Encoding Guidance

Population Encoding Guidance	
Populations	***Examples***
Business • Business Owners • Vendors • Merchants • Marketplace traders • Import/Exporters • Service providers (not employees) • Owner/operators	• Street vendors protesting their eviction from public land • Shopkeepers protesting against lack of air-conditioning in Botafogo mall • Taxi drivers protesting for regulation of "pirate" taxis • Fishermen protesting for better monitoring of fishing areas
Ethnic • Indigenous groups • Specifically named racial groups (afrodescendientes, mestizos, etc.)	• Wayúu (natives) starting a hunger strike in front of the Supreme Court building • Aborigines protesting against oil companies in Formosa, Argentina
Legal • Attorneys • Judges • Judicial workers	• Lawyers protesting against the supreme courts excessive spending habits

Population Encoding Guidance	
Populations	***Examples***
Education • Teachers • Students • Parents of students • Relatives of students	• Teachers protesting for better wages • A group of students protesting tax reforms that will impact education costs • Parents and family members of students protesting tax reforms that will impact education costs
Religious • Clergy • Missionaries • Nuns • Religious Groups • Churchgoers (residents, townsfolk, etc. that practice said religion and partake in the event due to religious motives)	• Clergy and about five thousand churchgoers marching for peace
Media • Reporters • News Crews	• Reporters protesting because they were not allowed access to the National Assembly in Caracas
Medical • Doctors • Nurses • Patients • Healthcare providers • Medical personnel/staff (e.g., radiologists, ambulance drivers)	• A strike by the Bolivian Medical Council • Patients at local hospital protesting for better services • Ambulance drivers in Paraguay protesting against the privatization of ambulance service in the country
Labor • Employees • Workers • Strikers • Retirees	• Army soldiers protesting for increase in their salary and benefits • Politicians going on strike for better retirement benefits • Prisoners rioting for better pay for prison labor • Retirees protesting for increase in their pension • Workers in Mexico protesting against government labor reforms • Dismissed workers protesting grounds for dismissal
Refugees/Displaced • People evicted from their homes • People displaced due to natural or manmade disasters • People displaced due to conflict • Homeless populations • Vagrants	• Refugees marching to city hall demanding that the government provide them with houses

Population Encoding Guidance	
Populations	***Examples***
Agricultural • Farmers • Livestock farmers • Fish farmers (not fishermen) • Peasants "Campesinos" (if the article provides the necessary information to determine that the population at hand is composed of farmers/agricultural workers)	• Peasants ("campesinos") belonging to the national agricultural association marching against the state governor (article provides the necessary information to determine that the population at hand is composed of farmers/agricultural workers) • Farmers protesting against cuts in agricultural subsidies
General Population • Groups listed as "residents" • Prisoners • Peasants or "campesinos" (for cases when the article does not provide the necessary information to determine that the population at hand is composed of farmers/agricultural workers) • Events where the specific population cannot be determined in the article (Example - "A group of angry people," etc.) • Young people (if they are not specified as being students) • Populations that do not fall under any of the other more specific population encodings	• The people of Caracas marching, demanding that they be told Hugo Chavez's health condition • Prisoners beginning hunger strike • Activists protesting against a makeup company using animals for testing • Women marching to raise awareness of violence against women • Peasants ("Campesinos") marching against the state governor (Note: Here, the article does not specifically talk about farmers/agricultural workers) • Residents protesting for access to better healthcare

Appendix B **Civil Unrest (CU) Event Discovery**

The following steps outline the process for determining OSI Civil Unrest (CU) and Widespread (WS) event ground truth sources and performing CU and WS event discovery with those sources.

- Identify top 2-3 native news sources from the country of interest.
    - In Fall 2012, MITRE formed a set of working groups composed of native Spanish and Portuguese speakers. Working group members evaluated a preliminary set of news sources derived from www.4imn.com; this website provides newspaper rankings based on Google Page Rank, Alexa Traffic Rank, and a proprietary method to determine page rank. The group encoded events from these news sources for over 2 months. Depending on the number and quality of events discovered from each news source, the news source was evaluated and possibly eliminated from the evaluation process. The result of this assessment was the selection of 2-3 of the most widely read and most popular newspapers in each country of interest.
- Identify search keywords to be used for event discovery.
    - Keywords or search terms were identified and refined by the working group. The goal was to select target keywords that returned articles of interest without too much extraneous information (i.e., too many articles to read and encode in a timely manner). The working group found a "happy medium" regarding the selection of keywords. As with the selection of news sources, the list of keywords was refined through use.
- Select event discovery methods.
    - The working group determined that Advanced Google Search is the best way to find online articles in each news source. Results from using a news source's own search function were inconsistent and incomplete.
    - The working group then determined that not all OSI-related articles were available online. The group selected Lexis Nexis and NewspaperDirect for print news source event discovery. Lexis Nexis may be queried using the same keywords as used for Advanced Google Search. NewspaperDirect does not offer a search function, instead the analysts read a PDF of the print version of the daily newspaper.
- Assign two native speakers (Spanish and/or Portuguese) to encode events in each country.
- Perform event discovery on a daily or nearly daily basis.
    - With Google Advanced Search, the analyst searches a specific newspaper website, using all search keywords at once and limiting the search result to a single day. The analyst reviews search results and encodes relevant events.
    - If Lexis Nexis is available for a specific news source, MITRE provides each analyst with Lexis Nexis search results twice a week. Again, the analyst reviews the results and encodes relevant events.
    - If the news source is available on NewspaperDirect, the analyst skims headlines to discover articles, reviews those articles, and encodes relevant events.

The following table lists the final set of ground truth sources by country and website including their availability through either Lexis-Nexis or Newspaper Direct, indicated by an X. (Note this table from Section 2.1 is reproduced here for the convenience of the reader.)

Table 47. Civil Unrest Ground Truth Sources (from Section 2.1)

Country	News Source	Website (for Advanced Google Search)	Lexis-Nexis	Newspaper Direct
Argentina	Clarín	http://www.clarin.com/		X
	La Nación	http://www.lanacion.com.ar/		X
Brazil	O Globo	http://oglobo.globo.com/	X	
	O Estado de São Paulo	http://www.estadao.com.br/		
	Jornal do Brasil	http://www.jb.com.br/		
Chile	La Tercera	http://www.latercera.com/		X
	Las Últimas Notícias	http://www.lun.com/		
	El Mercurio	http://www.emol.com/		
Colombia	El Espectador	http://www.elespectador.com/		
	El Tiempo	http://www.eltiempo.com/		
	El Colombiano	http://www.elcolombiano.com/		
Ecuador	El Universo	http://www.eluniverso.com/		
	El Comercio	http://www.elcomercio.com/	X	
El Salvador	El Diáro de Hoy	http://www.elsalvador.com/		
	La Prensa Gráfica	http://www.laprensagrafica.com/		
	El Mundo	http://elmundo.com.sv/		
Mexico	La Jornada	http://www.jornada.unam.mx		
	Reforma *	http://www.reforma.com/	X	
	Milenio	http://www.milenio.com/		X
Paraguay	ABC Color	http://www.abc.com.py/		
	Ultima Hora	http://www.ultimahora.com/		X
	La Nacíon	http://www.lanacion.com.py/		
Uruguay	El País	http://www.elpais.com.uy/		
	El Observador	http://www.elobservador.com.uy/		
Venezuela	El Universal	http://www.eluniversal.com/		
	El Nacional	http://www.el-nacional.com/	X	
	Ultimas Notícias	http://www.ultimasnoticias.com.ve/		

* Reforma is available only in a print version and thus cannot be searched via Advanced Google Search.

B.1 Advanced Google Search Terms

Spanish Keywords

bloqueo mostración manifestación protesta demostración mitin concentración motín disturbio huelga levantamiento rebelión sublevación piquete manifestantes inconformes activistas boicotear marcha marchar marcharon marchas cacerolazo manifestaron marchan

Portuguese Keywords

ativistas obstrução embarreiramento boicote manifestação manifestantes demonstração marcha marcharam passeata piquete protesto demonstração concentração comício motim distúrbio tumulto bloqueio greve rebelião revolta insubordinacão subversão

B.2 Lexis-Nexis Search Terms

Spanish Keywords

((cacerolazo OR huelga OR huelgas OR protestan OR protesta OR protestar OR protestó OR protestaron OR piquet OR manifestantes OR inconformes OR protestando OR manifestante OR protestas OR manifestaron OR activista OR militante OR "disturbios civiles" OR “se reunieron” OR boicotear OR concentrarse OR “se concentraron” OR manifestarse OR "se manifestaron" OR reuniones OR concentraciones OR manifestaciones OR concentrarse OR manifestarse OR "disturbios civiles" OR "disturbio civil" OR "perturbación civil" OR "disturbios civiles" OR "perturbaciónes civiles" OR "disturbio civil" OR "revuelta civil" OR "interrupción civil" OR "trastorno civil" OR "disturbios civiles" OR "revueltas civiles" OR "interrupciones civiles" OR "trastornos civiles" OR "desorden civil" OR "conflictos civiles" OR "lucha civil" OR "disturbios civiles" OR levantamiento OR disturbio OR motín OR amotinamiento OR sublevación OR antidisturbios OR "control de disturbios" OR amotinaron OR disturbios OR motines OR amotinamientos OR sublevaciones OR disturbios OR amotinamientos OR sublevaciones OR rebelión OR rebelaron OR levantamiento OR revuelta OR sublevación OR rebeliones OR levantamientos OR revueltas OR sublevaciones OR insurrección OR insurgente OR insurgentes OR turba OR turbas OR (paro AND (labor OR transporte OR servicios OR trabajadores OR nacional OR petrolero OR estudiantes OR estudiantil OR universidad OR obreros)) OR (("civil" OR "civiles" OR muchedumbre OR multitud OR muchedumbres OR multitudes OR sindicato OR sindicatos OR trabajadores OR empleados OR residentes OR miembros OR popular OR extendido OR mayor OR "gran escala" OR masa OR masivo OR votantes OR facción OR partido OR manifestantes OR manifestante OR inconformes OR proponente OR partidario OR partidarios OR opositores OR detractores OR críticos OR adherentes OR policía OR grupo OR grupos OR estudiantes OR activistas OR militantes OR activista OR militante OR multitud OR multitudes OR autoridades OR masas OR organizadores OR huelguistas OR rebeldes) AND (parada OR paradas OR desfile OR desfiles OR marcharon OR marchó OR marchas OR marcha OR marchando OR marchar OR protesta OR protestar OR protestó OR protestaron OR protestando OR disturbios OR desórdenes OR malestar OR conflictividad OR perturbación OR ruptura OR interrumpido OR trastorno OR huelga OR huelgas OR lucha OR conflictos OR bloqueo OR impedimento OR obstrucción OR bloquear OR impedir OR obstruir OR bloqueado OR impedido OR obstruido OR bloquearon OR impidieron OR obstruyeron OR bloqueos OR impedimentos OR obstrucciones OR reunirse OR reunieron OR concentrarse OR concentraron OR manifestarse OR manifestaron OR reuniones OR concentraciones OR manifestaciones OR sentada OR encierro OR ocupación OR "huelga de brazos caídos" OR "huelga pasiva" OR sentadas OR encierros OR ocupaciones OR "huelgas de brazos caídos" OR "huelgas pasivas" OR enfrentamiento OR enfrentar OR choque OR "gas lacrimógeno" OR barricada OR rodear OR obstrucción OR impedimento OR encercar OR encercaron OR encerco OR obstruir OR obstruyeron OR impedir OR impidieron OR impedido OR barricadas OR obstrucción OR cercas OR retén OR “puesto de control” OR obstáculos OR retenes OR "puestos de control" OR encuentros)))

Note: “Grupo” is removed for Venezuela’s “El Nacional” since “grupo” is a part of the source name – the “National Group.”

Portuguese Keywords

(ativista OR bloqueio OR manifestação OR hantavirose OR protesto OR protestei OR protestou OR protestamos OR protestaram OR protestava OR protestavamos OR protestavam OR comício OR motim OR distúrbio OR tumulto OR bloqueio OR "corte de estrada" OR greve OR revolta OR sublevação OR Parista OR Paristas OR passeata OR marcha OR caminhada OR demonstração OR barricada OR barreira OR movimento OR protestam)

Appendix C **Selected OSI Elections**

OSI includes selected national, state, and local elections. Relevant OSI elections are listed in the following table.

Table 48. Selected OSI Elections

National Election	State Election	Mayoral Election
Argentina President	None – Autonomous	Buenos Aires* Mayor
	Cordoba Governor	Cordoba Mayor
Bahrain Council of Representatives		Manama, Municipal Council
Belize Prime Minister	None – Civil servant	Belmopan* City Council
Bolivia President	La Paz Governor	None – Elected by City Council
	Santa Cruz Governor	
	Tarija Governor	
Brazil President	None – Federal District	Brasilia* Governor
	Sao Paulo Governor	Sao Paulo Mayor
	Rio de Janeiro Governor	Rio de Janeiro Mayor
	Bahia Governor	Salvador Mayor
	Ceara Governor	Fortaleza Mayor
	Minas Gerais Governor	Belo Horizonte Mayor
	Amazonas Governor	Manaus Mayor
	Parana Governor	Curitiba Mayor
	Pernambuco Governor	Recife Mayor
	Rio Grande do Sul Governor	Porto Alegre Mayor
Chile President	None – Intendant appointed by President	Santiago* Mayor
Colombia President	None – Capital District	Bogota* Mayor
	Valle del Cauca Governor	Cali Mayor
	Antioquia Governor	Medlin Mayor
	Atlantico Governor	Barranquilla Mayor
Costa Rica President	None – No elected provincial officials	San Jose* Mayor
Ecuador President	None – Appointed by President	Quito* Mayor
		Guayaquil Mayor
Egypt President	None	None
El Salvador President	None – Governor appointed	San Salvador* Mayor
None in French Guiana –Prefect appointed	None	Cayenne* Mayor
Guatemala President	None – Appointed by President	Guatemala City* Mayor
Guyana President	None	Georgetown* Mayor
Honduras President	None – Governor appointed	Tegucigalpa* Mayor
Mexico President	None – Federal District	Mexico City* Mayor
	Mexico Governor	Ecatepec Mayor
	Jalisco Governor	Guadalajara Mayor
	Puebla Governor	Puebla Mayor
	Oaxaca Governor	Juarez Mayor
	Baja California Governor	Tijuana Mayor
Nicaragua President	None	Managua* Mayor
Panama President	None – Appointed by President	
Paraguay President	None – Autonomous Capital District	Asuncion* Mayor

National Election	State Election	Mayoral Election
Peru President	None – Appointed by President	Lima* Mayor
Suriname President	None – Appointed by President	Paramaribo* Mayor
Syria President	None	None
Uruguay President	Montevideo Governor	Montevideo* Mayor
Venezuela President	None – Capital district	Caracas* Mayor
	Zulia Governor	Maracaibo Mayor
	Carabobo Governor	Valencia Mayor
	Aragua Governor	Maracay Mayor

*Capital city

Appendix D Monthly Scoring and Performance Measures

Upon completion of the GSR, events are compared with both Performer warnings and Base Rate warnings (See Appendix E for a discussion of Base Rate) in the scoring process. Using a linear optimization program (based on the Kuhn–Munkres or Hungarian algorithm), warnings and events are matched based on a non-zero Quality Score (QS), which is a measure of the accuracy of the match. During this process, only one warning is matched to one event and the reverse. The scoring process supports user defined "look-back" to the previous month for possible matches. Following the development of the warning-event (W-E) pairs, additional performance measures are developed and forwarded to the PM and Performer. The Performer uses both the GSR and monthly scores as a basis to improve their systems' forecasts; the PM uses monthly scores to assess performance and determine areas for improvement.

D.1 Update Sequences

As more information is available, the Performer 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.

D.2 Program Performance Measures

The set of overall program performance measures follows. The calculation of each measure is specific to each event type. (The following table was present in Section 1; it is reproduced here for easy reference.)

Table 49. Overview of Program 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.
Precision	Number of warnings received during a scoring period divided by the number of W-E pairs for the same period. (#W-E pairs/#Warnings)
Recall	Number of events in the GSR for a scoring period divided by the number of W-E pairs for the same period. (#Actual Events/#W-E pairs)
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	Mean precision for all W-E pairs, reported for each country for a scoring period.
Mean Recall	Mean recall for all W-E pairs, reported for each country for a scoring period.
Mean Probability	Mean probability for all W-E pairs, reported for each country for a scoring period.

As presented in Section 2, not all measures are relevant to every event type. The following table summarizes the computed performance measures for all OSI event classes. An "X" in a cell indicates that the measure is computed for the associated event types. Note that Average Probability is computed from the Performer/Base Rate reported Probability measure.

Table 50. Summary of Computer Performance Measures by Event Type

Event Type	Quality Score	Lead Time	Precision	Recall	Probability	Average QS	Average LT	Average Probability
Civil Unrest—01xx and 07xx	X	X	X	X		X	X	X
Vote—02xx	X	X	X	X		X	X	X
Rare Disease—0311, 0312, 0313, 0314, 0315, 0317, 0318	X	X	X	X		X	X	X
Chikungunya—0316								
Saudi Arabia MERS—036	X	X				X	X	
Polio—0318	X	X				X	X	
Disease Season Date Warnings—03x1, 03x2, 03x3	X	X						
Disease Season Case Count Warnings—03x4, 03x5	X	X						
Influenza-like Illness (ILI) - 033	X	X	X	X		X	X	
ICEWS 05	X						X	
ICEWS 06	X	X				X	X	
ICEWS 06	X	X				X	X	

D.2.1 Common Performance Measure Calculations

Some performance measures are calculated in the same way for all event types. They are listed in the following table:

Table 51. Common Performance Measure Calculations

Performance Measure	Calculation
LT*	Difference between the event's earliest reported date (ERD) and the date/time when the warning was received.
Precision	Total number of W-E Pairs / Total warnings ingested; this is reported for each country and as a summary for all countries.
Recall	Total number of W-E pairs / Total number of GSR events; this is reported for each country and as a summary for all countries.
Average QS	The average of the individual QS for all W-E pairs; this is reported for each country and as a summary for all countries.
Average LT*	The average of individual LT for all W-E pairs; this is reported for each country and as a summary for all countries.
Average Probability	The average of individual reported Probability measures; this is reported for each country and as a summary for all countries.

* Note: LT is calculated differently for event codes 03x4 and 03x5 (Disease Season Case Counts).

For 03x4 (peak week case counts) events, LT is calculated as the difference between the date/time the warning was received and the first day of the ground truth Peak Week. If this value is less than one day, the warning is considered late will not be scored.

For 03x5 events (total disease season case counts), LT is calculated as the difference between the date/time the warning was received and the earliest reported data (ERD) of the End Date of the season, i.e. the maximum number of disease cases cannot be determined until after the season ends. 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.

D.2.2 Quality Score Calculations

Civil Unrest QS Calculations

OSI uses two different methods to calculate Civil Unrest QS depending on whether the "traditional" vs Location Scoring (LS) scorer is used.

LS QS Calculation

$$\text{MAX}\left(\left(2 * \left(1 - \frac{\text{ABS(warning date} - \text{event date)}}{7}\right)\right), 0\right) + 2 * \left(\text{MAX}\left(\frac{\text{MaxDis} - \text{Distance}}{\text{MaxDis}}\right), 0\right),$$

where MaxDis is user defined (currently set to 300 KM) and Distance is the distance in km between the location specified in the warning and the location specified in the GSR. Section 2.1 presents an illustration of LS QS calculation—these figures are replicated here for convenience.

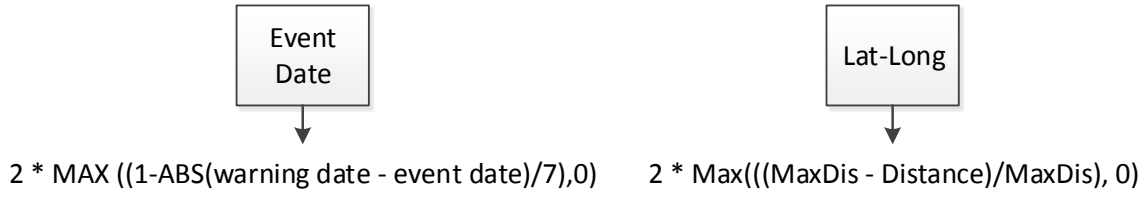


Figure 9. LS Civil Unrest QS Calculation, Maximum Score is 4

An example of the LS CU QS scoring follows:

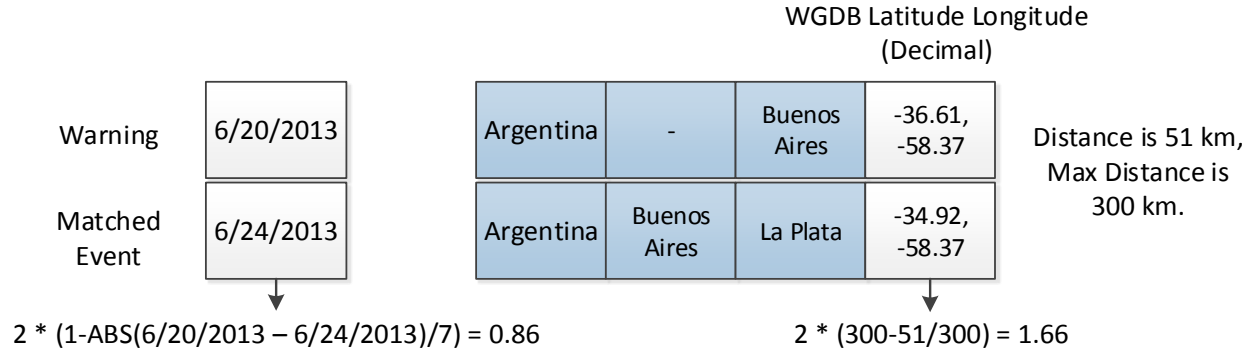


Figure 10. LS Scorer Civil Unrest QS Example

Traditional QS Calculation

$$2 * \left(1 - ABS\frac{warning\ date - event\ date}{7}\right) + Location\ Score + Event\ Score + Population\ Score,$$

where each component of the score has a maximum value of 1. Section 2.1 presents an illustration of Tradition QS calculation—these figures are replicated here for convenience.

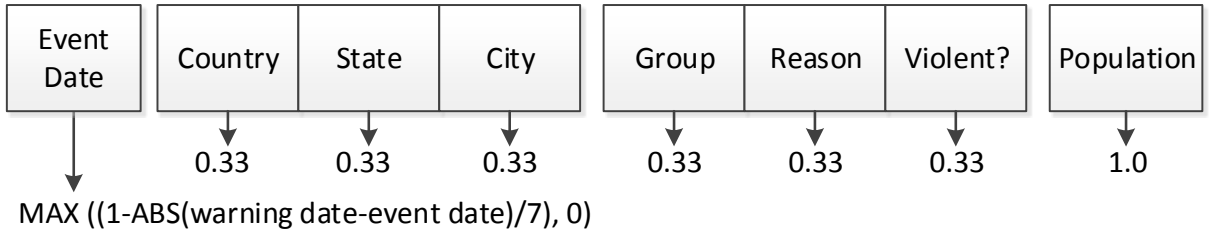


Figure 11. Traditional Civil Unrest QS Calculation, Maximum Score is 4

An example of traditional CU QS scoring follows:

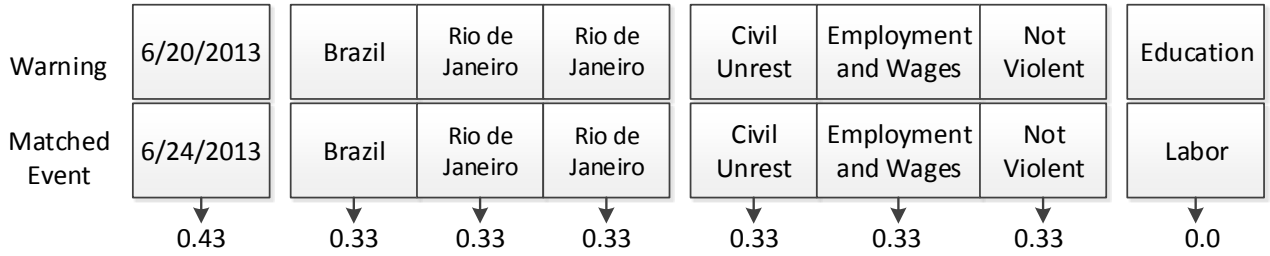


Figure 12. Traditional Scorer Civil Unrest QS Example

Vote QS Calculation

If the population fields of the warning-event pair (W-E) match, the QS is 4. Otherwise, the QS is 0.

Rare Disease QS Calculation

$$\frac{4}{3} * \left\{\left(1 - \frac{ABS(warning\ date - event\ date)}{7}\right) + Location\ Score + Event\ Score\right\}$$

The following figure shows how the Rare Disease QS is calculated.

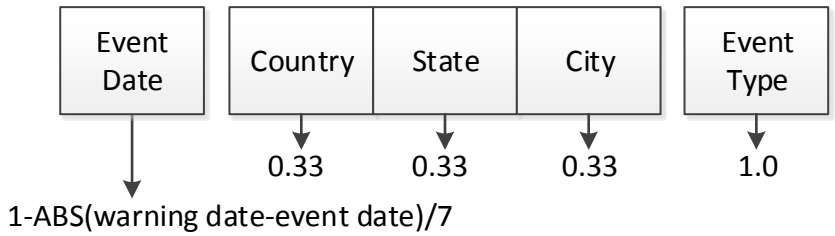


Figure 13. Rare Disease QS Calculation Example

An example of a Rare Disease scoring follows.

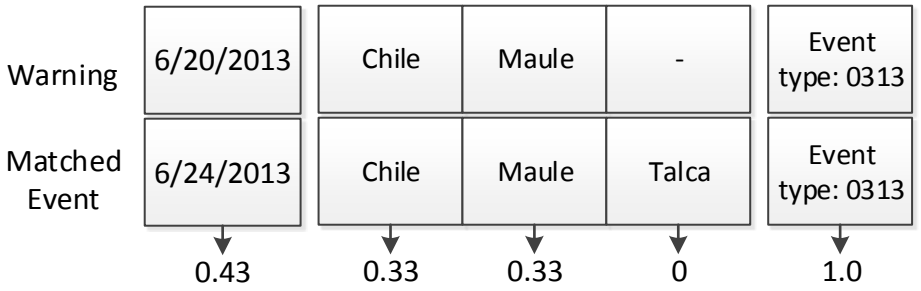


Figure 14. Rare Disease QS Scoring Example

Weekly Case Count Disease QS Calculation

The QS for Case Count Warnings is calculated as follows:

(1-(ABS (Predicted Case Count-Actual Case Count)/MAX (Predicted Case Count, Actual Case Count)))) * 4

Polio QS Calculation

Note that although the number of confirmed case counts is encoded in the GSR, for the calculation of QS, a non-zero number is treated as a one and zero is treated as zero. The QS for Polio W-E pairs calculated as follows:

$$\text{QS is: } 4 * (1.0 - ((Actual\ DPC\ probability - Predicted\ DPC\ probability) ** 2.0))$$

Disease Season Date Warnings Performance Measures

The QS for Date Warnings (03x1, 03x2, 03x3) is calculated as follows:

$\left(\text{MAX}\ \left(0, 1 - \text{ABS}\frac{PredictedEW - ActualEW}{NMax}\right)\right) * 4$, where NMax is the allowable match window measured in weeks, to be determined by the PM. For example, if Nmax = 4 and PredictedEW-ActualEW = 5, then there is no match. If NMax > End Date - Start Date, then NMax = End Date - Start Date.

Influenza-like Illness QS Calculation

$$\left\{1 - \frac{ABS(Predicted\ Case\ Count - Actual\ Case\ Count)}{MAX(Predicted\ Case\ Count, Actual\ Case\ Count)}\right\} * 4$$

ICEWS 05 QS Calculation (Domestic Political Events (DPC))

$\mathbf{4} * (\mathbf{1.0} - ((\boldsymbol{Actual\ DPC\ probability} - \boldsymbol{Predicted\ DPC\ probability})^{\mathbf{2.0}}))$, where DPC represents Domestic Political Events, an ICEWS-provided statistic.

ICEWS 06 QS Calculation

QS is determined as follows:

- Actual = $MAX(actual\ event\ counts, 0)$
- Predicted = $MAX(predicted\ event\ counts, 0)$
- Occurrence_Score = $\begin{cases} 1\ if\ Actual\ and\ Predicted\ are\ both\ zero\ or\ both\ nonzero \\ 0\ if\ otherwise \end{cases}$
- Accuracy_score = $1 - ABS(actual - predicted)/MAX(actual, predicted, 4)$
- QS = $0.5 * Occurrence_Score\ +\ 3.5 * Accuracy_Score$

Appendix E **Base Rate Model**

There is no known OSI-like system that predicts discrete events. A simple benchmark for the Performers' systems is a base rate warning generation model that uses historical event frequencies to generate forecasts. The philosophy of the Base Rate model is that forecasts are generated using only the time series properties of the GSR data and no other inputs. Beginning in Year 2 of the Program, the Government Team provided base rate warnings derived from historical GSR data. As event classes and methods for scoring have changed throughout the Program, the Base Rate model suite has been modified and extended. This appendix describes the Base Rate model suite as of February 2015.

Civil Unrest, Widespread Civil Unrest, and Rare Disease

The Government's simple base rate warning general model is presented below in Figure 15 for event classes 01xx, 0311 through 0315, 0317, 0319, and 07xx.

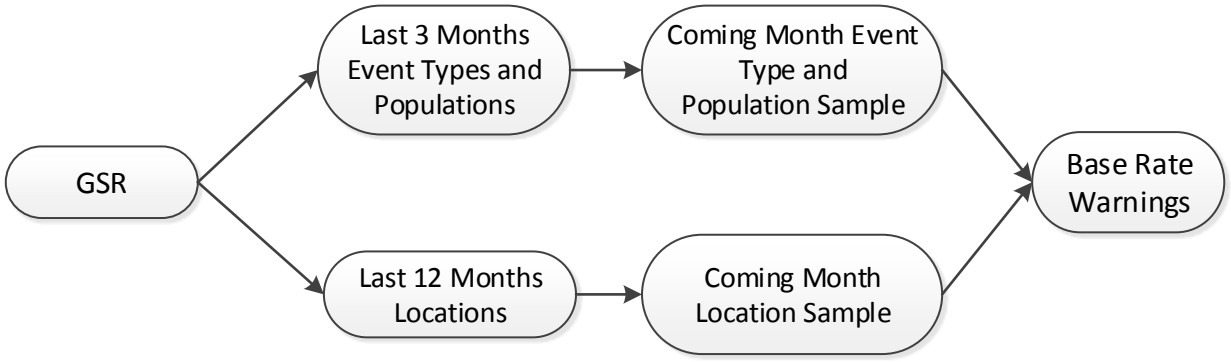


Figure 15. Base Rate Warning Development Process

The Base Rate model for these event classes is a sample from the historical GSR, starting with May 2013. For each country, the previous three-month history is used to draw a sample of event classes, populations, and dates for the coming month.[1] Locations are assigned at the city level using a sample from the previous 12 months of history.

Disease Season Events

The Base Rate model for disease season events (starting, peak week, peak count, ending week, and total count) uses the historical average for each of these parameters for the given country. This is used to generate warnings for event classes 033x, 034x, and 035x.

Disease Case Count Events

Latin American Chikungunya events (event class 0316) are the cumulative count of cases for a given country. They are forecast using a linear extrapolation based on the last two GSR points.

MERS (event class 036) and Polio (event class 0318) events are the incremental weekly count of diseases. The Base Rate model uses the exponentially weighted moving average of the last 4 reports to compute warnings for the coming month.

[1] Analysis by the Government team showed that the overall Quality Score for the Base Rate Model is fairly insensitive to the particular random draw of events.

ICEWS Domestic Political Crisis (DPC) Event

The Base Rate model for event class 05 is very simple. Base Rate uses a status quo model; if a country was in DPC, it remains in DPC and if it was not in DPC, then it stays out of DPC.

ICEWS 06xx Events

The Base Rate model for ICEWS events (event class 06xx) fits an autoregressive moving average process to the observed history of country and capital values. The parameters for the process are country-dependent. This process is used to forecast the weekly counts for the coming week. Additionally, if the forecast event count is less than 0.5 it is reassigned to be 0.

Base Rate Warning Submission and Ingest

The warnings from this Base Rate model are sent to MITRE in batch form prior to the beginning of the month to which these warnings refer. For example, October warnings from the Base Rate model will be sent to MITRE by September 30. These warnings are then scored at the same time as the Teams' warnings and results are reported.

When MITRE receives the Base Rate warning file, it is first checked to make sure that the warning file is comma separated and that the text is in ISO 65001 UNICODE (UTF-8) format. Further, the warning file is checked by a tabular ingest process to ensure that it does not contain any duplicate warning records, including duplicate warnings with different warning IDs. The tabular ingest process then creates entries in the OSI records and warnings database tables, which are similar to the entries generated by the normal OSI ingest process. The separate day, month and year columns are combined to form valid event date and date of receipt entries. The location information is checked against the OGDB for validity. Base Rate warning IDs have the letter "R" prepended to them, but otherwise a Base Rate warning is identical in appearance to a normal OSI warning in the warnings and records database tables.

Future Plans

In the future and depending on new Program requirements, this basic Base Rate model may be upgraded to a more sophisticated model (e.g., by using the base rate distribution of geographic locations at the city level where events occur, or by using base rate distribution of days of the week when events occur).

Appendix F GSR Quality Control (QC)

F.1 Civil Unrest and Disease Quality Review

MITRE conducts a 100% review of all analyst-encoded Civil Unrest events using the common set of encoding guidance (see Appendix A) to correct encoding errors as well as identify and remove duplicate events. As part of this review, MITRE collects and assesses performance measures for each analyst. This allows MITRE to evaluate recurring errors and to: 1) improve/clarify encoding guidance, 2) improve analyst-training materials, or 3) retrain analysts in those areas where confusion may exist.

MITRE disease analysts use two methods for quality review. First, MITRE analysts cross-check one another's encoded events on a monthly basis and more frequently if needed. Also, MITRE uses supplemental ground truth sources (see Section 2 for more details) to ensure the continuing quality of the primary ground truth source.

F.2 Semi-Automated Verification and Validation (V&V)

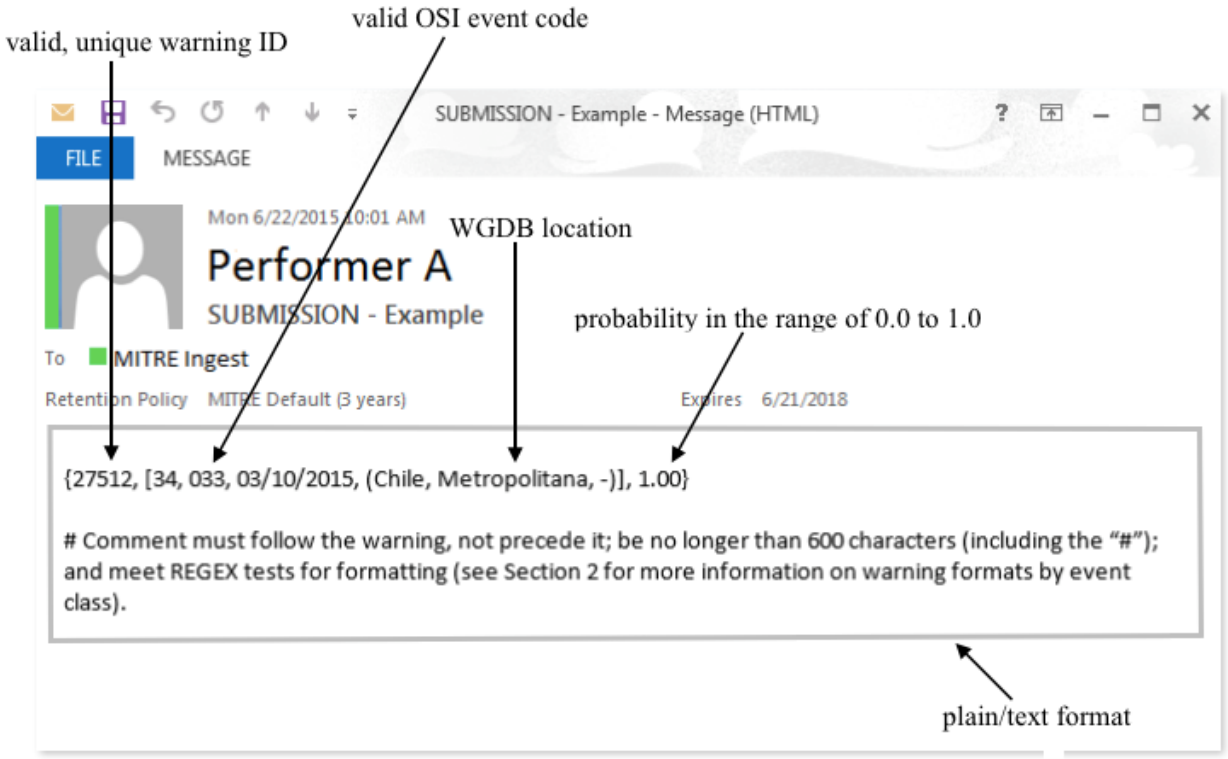
Once the monthly GSR is assembled, MITRE conducts a semi-automated verification and validation process (V&V). The GSR is automatically scanned and inconsistencies are flagged according to the following criteria. Results of the automated check are manually assessed and corrected. This process is repeated until the automated scan finds no inconsistencies. Note the V&V process is expanded as new requirements are identified.

- All location names conform to the spelling and diacritic markings in the OGDB.
- All dates must be encoded in mm/dd/yyyy format.
- All date fields are encoded.
- Earliest Reported Date (ERD) may not precede the event date.
- Location and population fields may not contain leading and trailing blanks.
- Population field must contain a valid OSI population.
- Event code must contain a valid OSI event code.
- There are no duplicate event IDs.
- There are no occurrences of "=CLEAN(#REF!)" or "TRIM." (These functions are used to assemble the GSR and may appear in the GSR.)
- There is no data in GSR columns to the right of Column S.
- No dates are "out of range" dates (before 1/3/2011 or greater than a year past the date run).
- Highlights empty First Reported Link, Other Links (GSS), and Non-GSS Other Links GSR columns (Note only the Non-GSS Other links column may be empty).
- Ensures that all location triplets (Country, State, City) have valid latitude-longitude data in the OGDB.

F.3 Warning Ingest Evaluation Criteria

All Performers send warnings via email message to the MITRE ingest process. The ingest process records the date and time the message is received at the MITRE Corporate Enterprise Mail server and assigns a unique Performer identifier to each warning contained in a warning message. Following receipt, the Ingest Process assesses each message according to the following criteria. Non-compliant warnings are rejected and the Performer informed by email message. Performers may resubmit rejected warnings, but a new date/time-stamp will be assigned.

An example of a compliant warning submission follows:



F.4 Scoring Quality Control (QC)

In additional to a manual check of monthly scores, MITRE developed and uses automated Civil Unrest (CU) and Rare Disease (RD) QC code to check the validity of CU and RD scorer outputs. The automated code takes the intermediate outputs of the scoring code, applies the scoring algorithm, and generates two reports: an assessment of top-level measures and an assessment of bottom-level measures.

The following table is a truncated extract for the automated CU QC top level measures report for Brazil. The top row is the CU scorer output. The bottom row shows the results of the QC rescore. The center row is a logical test to show if the two results are equivalent. For example, the "Ave. QS" column shows a FALSE. This result would require MITRE QC analyst intervention to determine the nature of the error and its resolution.

Table 44. Sample Automated CU QC Output—Top Level Measures

Country	#Events	#Warnings	#W-E Pairs	Ave QS	Ave Prob-M	Ave LT	Precision	Recall
Brazil	124	135	76	3.36	0.65	7.43	0.56	0.61
	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE
	124	135	76	2.694581	0.654703	7.434211	0.562963	0.612903

The next table shows a truncated extract from the automated CU QC bottom level measures report for Brazil. The top row is the CU scorer output. The bottom row shows the results of the QC rescore. The center row is a logical test to show if the two results are equivalent. For example, the "Ave. QS" column shows a FALSE. This result would require MITRE QC analyst intervention to determine the nature of the error and its resolution.

Table 45. Sample Automated CU QC Output—Top Level Measures

Country	Warning ID	GSR Event ID	Date Score	Location Score	Total Quality	Lead-Time	Probability
Argentina	V19872	22271	1	0.9	3.8	1	0.97
			TRUE	TRUE	FALSE	TRUE	TRUE
			1	0.902666	2.235999	1	0.9744

Appendix G **OSI Gazetteer Database (OGDB)**

Early in the Program, the Team used an open-source Gazetteer called the World Gazetteer as the basis to develop a Program-specific location Gazetteer, the OSI Gazetteer Database (OGDB). Note that in July 2013, the open source World Gazetteer ceased operations. MITRE maintains the OGDB, updates it as required, and shares updated versions with both the PM and Performers on a regular basis.

Initially, the OGDB was used as the standard for spelling and diacritic marks. However, with the introduction of Location Scoring (LS), latitude-longitude is necessary for scoring. The OGDB is currently the standard for latitude-longitude data.

The OGDB is modified on an as-needed basis to: 1) add new location triplets (Country, State, City), 2) add or correct GEO data, and 3) correct incorrect data.

To either add a new location triplet or correct/complete an OGDB entry:

- Find the location in an authoritative backup source. Authoritative sources for OSI include, in order of preference:
  - Getty Thesaurus of Geographic Names—primary source (http://www.getty.edu/research/tools/vocabularies/tgn/index.html)
  - National Geospatial-Intelligence Agency Geo-Names (NGA)—secondary source (http://geonames.nga.mil/namesgaz/)
  - Google Maps—tertiary source (https://www.google.com/maps)
- Encode new/corrected location with spelling and diacritic marks from the authoritative source.
- Add any known city-name nicknames to the OGDB record.
- Add latitude longitude data in decimal degrees from the authoritative source.

Appendix H Instructions for Performers

H.1 Audit Trail

Teams were required to generate an "audit trail" for any warning sent to MITRE, after Month 9. The goal of this requirement was to make the warning transparent to the user by allowing a drill-down from warning to data, and to make it comprehensible by generating an analytic narrative that explains why the warning was generated. The drill-down goes through multiple layers of computation and data processing, mirroring the major components of the Teams' theoretical and technical approaches, and the system architecture and processes. The analytic narrative uses natural language to build an argument to back up the warning.

Ultimately, the audit trail capability was mapped to a series of user interfaces as an integral part of the system; this was accomplished in Years 2 and 3 of the Program. During the Base Year of the Program, Teams were required only to provide insight into the behavior of the system, still allowing the user (the OSI Government Team) to navigate from a warning to the data generating it, through analytic layers determined by the Teams' theoretical and technical approaches, and the system architecture and processes. However, in the Base Year there was no focus on user interfaces or analytic narrative; manual access to the major data processing and computation steps were sufficient. Since the audit trail was so dependent on the Teams' approach and system, the Government Team worked with each Team separately to develop the specific requirements for the audit trail. This was accomplished between Months 5 and 9 of the Program. Starting at Month 9, Teams provided an audit trail for all warnings submitted.

For Year 2 of the Program, Teams made their Year 1 audit trail capability web-enabled so that it could be easily demonstrated "live" to potential OSI Government transition partners, and tested by a few analysts, at their respective sites. This was accomplished by Month 17 of the Program, the end of the first severable period of Year 2. For the remainder of Year 2 of the Program, Teams received feedback from Government analysts about the functionality and user interface of the audit trail, and in coordination with the OSI Government Team made the appropriate changes to the audit trail capability.

In addition, by Month 22 of the Program, Teams developed a detailed technical plan and mock-up of how their audit trail capability would include the analytic narrative functionality. This narrative would use natural language and meet two objectives:

1. Build an argument why the warning was sent and why now. This is basically an extraction and synthesis of the insights that the audit trail capability of Year 1 provides. In Year 1, by navigating from warning to data through analytic layers, the analyst is expected to extract these insights and develop his/her own understanding of why the warning was sent. The new requirement for the audit trail aims to make these insights explicit and articulate them into a natural language, logical narrative that provides an explanation and understanding of why the warning was sent.
2. Provide context that might explain why the warning was sent. While the first objective focuses on the data and data processing used by the system to generate the warning, this objective focuses on the societal context that generated that data. For example, the context might include information about what is going on on-the-ground (e.g., known on-going labor negotiations that led to a strike, or tax policy just adopted by the Parliament

that led to protests), or connection to other warnings and events (e.g., protests that spread from an initial location to new locations), or cultural/historical knowledge that might explain the event (e.g., protest to mark the anniversary of a coup, or protest related to indigenous people).

Between Month 22 and the end of Year 2 of the Program, Teams worked with the Government Team to finalize the detailed technical plan and mock-up to achieve these objectives. Teams then implemented the approved plan in Year 3.

H.2 Replication and Ablation Tests

During site visits, Teams were asked to replicate certain warnings they submitted and build that capability into their system. Teams employed version control and proper data management to be able to rerun a specific version of the system, on the specific portion of the data that was used to generate a particular warning. This capability was exercised at the same time as the audit trail, so by Month 9, Teams were able to replicate their warnings. During Years 2 and 3 of the Program, Teams were asked to perform ablation tests. These tests were intended to promote the use of multiple feeds, and to gauge how performance degrades as a function of the availability of any particular feed.

H.3 Trade-offs: Precision v. Recall and Lead-time v. Quality Score

Understanding the trade-offs between precision and recall and between lead-time and quality score is very important for analysts; how the systems set these trade-offs vary greatly across analytic requirements. Some analysts favor recall over precision while others favor precision over recall. Likewise, some analysts favor lead-time over quality score, while others favor quality score over lead-time. Thus, it was important that the systems formalize and quantify these trade-offs and ultimately provide the analyst a "dial" to manage explicitly these trade-offs. This was a requirement for Year 2 and 3 of the Program.

By Month 17 of the Program, the end of the first severable period of Year 2, Teams developed a detailed technical plan and mock-up of how they would formalize and quantify these trade-offs and make them visible to the analysts in the audit trail user interface. Teams worked with the Government Team and the approved plan was implemented during the remainder of Year 2.

By Month 22 of the Program, Teams developed a detailed technical plan and mock-up of how the system would include the "dial" functionality for analysts to manage explicitly the trade-offs. Teams then implemented the approved plan in Year 3 of the Program.

H.4 Automated Warning Generation

The technology developed under OSI has no "human in the loop." The Teams' systems generate warnings without the help of subject matter experts (SMEs), either to guide the system or to filter warnings before they are submitted. SMEs help develop, train, and improve the systems, but the warnings are machine-generated and all warnings generated are submitted automatically.

Appendix I **Program Milestones**

OSI's year-end milestones are presented in the following table:

Table 46. OSI Year-End Program Milestones

Measure	**Month 12** 4 months of warnings	**Month 24** 12 months of warnings	**Month 36** 12 months of warnings
Mean Lead-Time	one day	3 days	7 days
Mean Probability Score	0.60	0.70	0.85
Mean Quality Score	3.0	3.25	3.5
Recall	0.50	0.65	0.80
Precision	0.50	0.65	0.80

Other assessments, qualitative and quantitative, will be performed by the Government Team to evaluate each Team's approach (e.g., the scientific merit of the approach and the research findings). Most importantly, the program will be evaluated by the utility of the warnings as judged by potential users.

Appendix J **Acronyms**

CU	Civil Unrest
DARPA	Defense Advanced Research Projects Agency
DPC	Domestic Political Crisis
ERD	Earliest Reporting Date
EW	Epidemiological Week
Flu	Influenza
GSR	Gold Standard Reference (a monthly report containing an encoded-event data set).
GSS	Gold Standard Source
IARPA	Intelligence Advanced Research Projects Activity
ICEWS	Integrated Conflict Early Warning System
ILI	Influenza-Like Illness
LA	Latin America
Lat-long	Latitude and longitude data
LS	Location Score
MENA	Middle East and North Africa
MERS	Middle East Respiratory Syndrome
MoH	Ministry of Health
NGA	National Geospatial-Intelligence Agency Geo-Names
ODNI	Office of the Director of National Intelligence
QC	Quality Control
QS	Quality Score
PAHO	Pan American Health Organization
PM	Program Manager
ProMED	The Program for Monitoring Emerging Diseases
RD	Rare disease
RSV	Respiratory Syncytial Virus
SME	Subject Matter Expert
T&E	Test and Evaluation
V&V	Verification and Validation
VT	Virginia Tech
OGDB	OSI Gazetteer Database

WHO	World Health Organization
WS	Widespread

This page intentionally left blank.