

“Beating the news” with EMBERS: Forecasting Civil Unrest using Open Source Indicators

Naren Ramakrishnan, Patrick Butler, Sathappan Muthiah, Nathan Self, Rupinder Khandpur, Parang Saraf, Wei Wang, Jose Cadena, Anil Vullikanti, Gizem Korkmaz, Chris Kuhlman, Achla Marathe, Liang Zhao, Ting Hua, Feng Chen, Chang-Tien Lu, Bert Huang, Aravind Srinivasan, Khoa Trinh, Lise Getoor, Graham Katz, Andy Doyle, Chris Ackermann, Ilya Zavorin, Jim Ford, Kristen Summers, Youssef Fayed, Jaime Arrendondo, Dipak Gupta, David Mares

Presented by Naren Ramakrishnan



The EMBERS Project

- Funded by the IARPA Open Source Indicators (OSI) program
 - aims to develop methods for continuous, automated analysis of **publicly available data** in order to anticipate and/or detect **population-level events** such as mass violence, protests, riots, mass migrations, elections, disease outbreaks, economic instability, resource shortages, and responses to natural disasters.
- OSI program **geographical focus**: Latin America
- Research and Development project: 3yrs, began Apr 2012
- Three initially funded teams
- EMBERS is a deployed system from Nov 2012



The EMBERS team

- 6 universities and 2 companies
- University partners
 - Develop algorithms, test and evaluate locally
- Industrial partners
 - Data ingest, enrichment, algorithm deployment 24x7 in production architecture



This paper

- Forecasting civil unrest
 - Protests, strikes, and occupy events
 - Predict the who, where, when, and why of the protest
- Regional focus on 10 countries
 - Argentina, Brazil, Chile, Colombia, Ecuador, Mexico, Paraguay, El Salvador, Uruguay, and Venezuela



Why forecast protests?

- For the social scientist
 - Insight into how citizens express themselves
- For the traveler
 - Travel alerts
- For law enforcement
 - Design measures to control violence and minimize disruptions
- For the government
 - Prioritizing citizen grievances
- For industries
 - Supply chain management
 - Cascading effects on financial markets, government stability

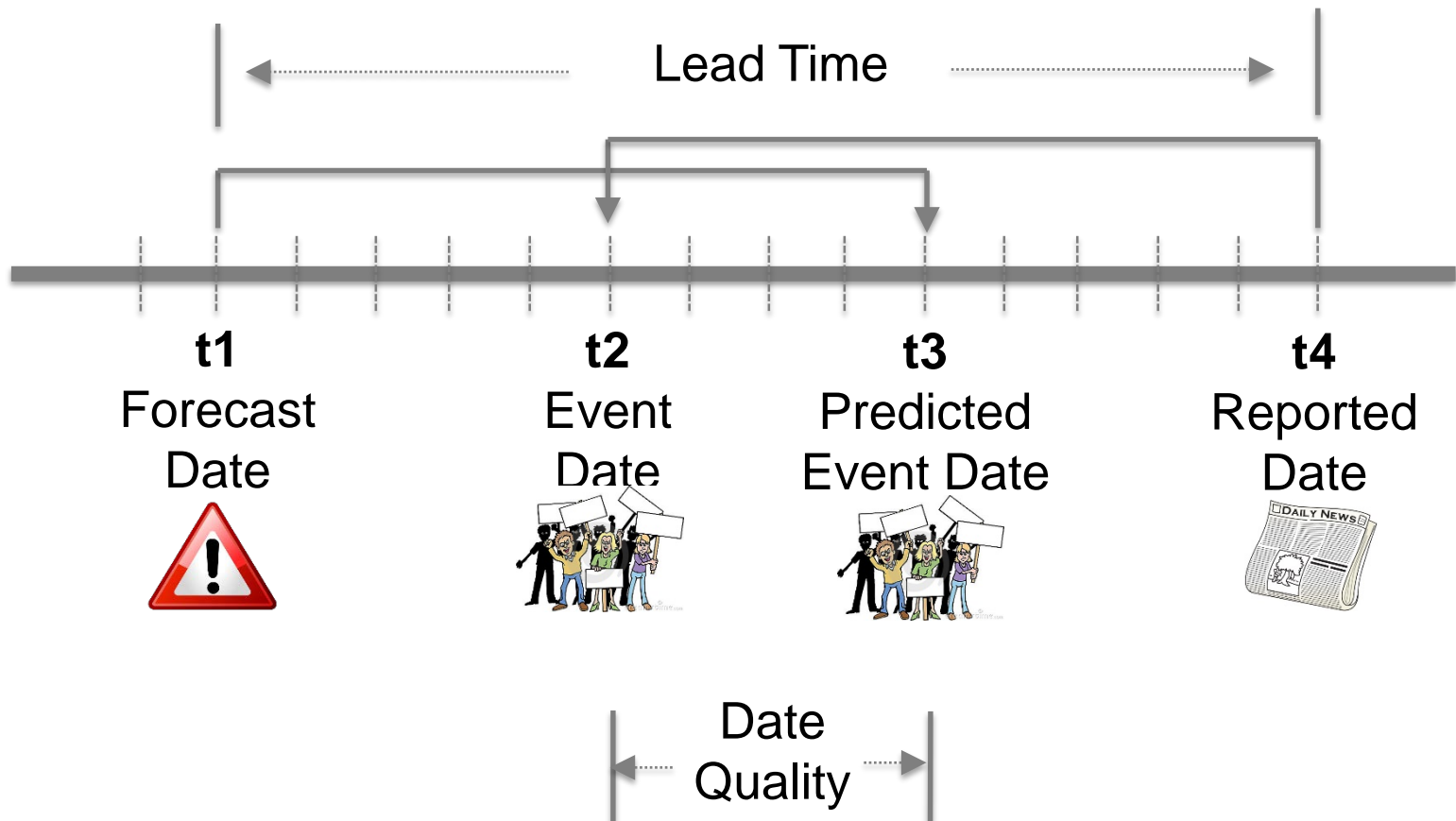
How we get evaluated

- Forecasts automatically emailed for evaluation without human-in-the-loop

```
{8691, [Labor, 0111, 10/03/13, (Brazil, Paraná, Curitiba)], 1.00}  
{8693, [Education, 0161, 10/17/13, (Chile, Coquimbo, Coquimbo)], 1.00}
```

- Evaluation done independently of authors
 - by **MITRE**
- Quantitative metrics for forecasting
 - *Quality* (How good is the warning?; graded on a 0-4 scale)
 - *Lead Time*/Timeliness (How far in advance?)
 - *Recall*, i.e., Completeness (How many events were there warnings for?)
 - *Precision*, i.e., Accuracy (How many warnings matched an event?)
 - *Probability*, i.e., Reliability (How good a likelihood estimate is made?)

Lead time



Other aspects of quality

Alert

```
{
  8691,
  [
    03/10/13,

    Education,

    Civil unrest-
    Employment and Wages
    Non-Violent,

    (
      Brazil,
      Paraná,
      Curitiba
    )],
  1.00
}
```

Date of Delivery
03/03/13

GSR

```
{
  GSR-13891,
  [
    03/08/13,

    Education,

    Civil unrest-
    Housing
    Non-Violent,

    (
      Brazil,
      Paraná,
      Ângulo
    )],
}
```

Earliest Reported Date
03/09/13

Other aspects of quality

Alert

```
{
  8691,
  [
    03/10/13,
    Education,
    Civil unrest-
    Employment and Wages
    Non-Violent,
    (
      Brazil,
      Paraná,
      Curitiba
    )],
  1.00
}
```

Date of Delivery
03/03/13

Date Score
 $1 - \min(7, 2) / 7 = 0.71$

GSR

```
{
  GSR-13891,
  [
    03/08/13,
    Education,
    Civil unrest-
    Housing
    Non-Violent,
    (
      Brazil,
      Paraná,
      Ângulo
    )],
}
```

Earliest Reported Date
03/09/13

Other aspects of quality

Alert

```
{
  8691,
  [
    03/10/13,
    Education,
    Civil unrest-
    Employment and Wages
    Non-Violent,
    (
      Brazil,
      Paraná,
      Curitiba
    )],
  1.00
}
```

Date of Delivery
03/03/13

Date Score
 $1 - \min(7, 2) / 7 = 0.71$

Population Score
1.0

GSR

```
{
  GSR-13891,
  [
    03/08/13,
    Education,
    Civil unrest-
    Housing
    Non-Violent,
    (
      Brazil,
      Paraná,
      Ângulo
    )],
}
```

Earliest Reported Date
03/09/13

Other aspects of quality

Alert

```
{  
  8691,  
  [  
    03/10/13,  
  
    Education,  
  
    Civil unrest-  
    Employment and Wages  
    Non-Violent,  
  
    (  
      Brazil,  
      Paraná,  
      Curitiba  
    )],  
  1.00  
}
```

Date of Delivery
03/03/13

Date Score
 $1 - \min(7, 2) / 7 = 0.71$

Population Score
1.0

Event-Type Score
 $0.33 + 0.0 + 0.33 = 0.66$

GSR

```
{  
  GSR-13891,  
  [  
    03/08/13,  
  
    Education,  
  
    Civil unrest-  
    Housing  
    Non-Violent,  
  
    (  
      Brazil,  
      Paraná,  
      Ângulo  
    )],  
}
```

Earliest Reported Date
03/09/13

Other aspects of quality

Alert

```
{
  8691,
  [
    03/10/13,
    Education,
    Civil unrest-
    Employment and Wages
    Non-Violent,
    (
      Brazil,
      Paraná,
      Curitiba
    )],
  1.00
}
```

Date of Delivery
03/03/13

Date Score
 $1 - \min(7, 2) / 7 = 0.71$

Population Score
1.0

Event-Type Score
 $0.33 + 0.0 + 0.33 = 0.66$

Location Score
0.33
+0.33
+0.0
=0.66

GSR

```
{
  GSR-13891,
  [
    03/08/13,
    Education,
    Civil unrest-
    Housing
    Non-Violent,
    (
      Brazil,
      Paraná,
      Ângulo
    )],
}
```

Earliest Reported Date
03/09/13

Other aspects of quality

Alert

```
{
  8691,
  [
    03/10/13,
    Education,
    Civil unrest-
    Employment and Wages
    Non-Violent,
    (
      Brazil,
      Paraná,
      Curitiba
    )],
  1.00
}
```

Date of Delivery
03/03/13

Date Score
 $1 - \min(7, 2) / 7 = 0.71$

Population Score
1.0

Event-Type Score
 $0.33 + 0.0 + 0.33 = 0.66$

Location Score
0.33
+0.33
+0.0
=0.66

Total Quality Score
 $= 1 + 0.66 + 0.71 + 0.66 = 3.03$

GSR

```
{
  GSR-13891,
  [
    03/08/13,
    Education,
    Civil unrest-
    Housing
    Non-Violent,
    (
      Brazil,
      Paraná,
      Ângulo
    )],
}
```

Earliest Reported Date
03/09/13

Other aspects of quality

Alert

```
{
  8691,
  [
    03/10/13,
    Education,
    Civil unrest-
    Employment and Wages
    Non-Violent,
    (
      Brazil,
      Paraná,
      Curitiba
    )],
  1.00
}
```

Date Score
 $1 - \min(7, 2) / 7 = 0.71$

Population Score
1.0

Event-Type Score
 $0.33 + 0.0 + 0.33 = 0.66$

Location Score
0.33
+0.33
+0.0
=0.66

Total Quality Score
 $= 1 + 0.66 + 0.71 + 0.66 = 3.03$

GSR

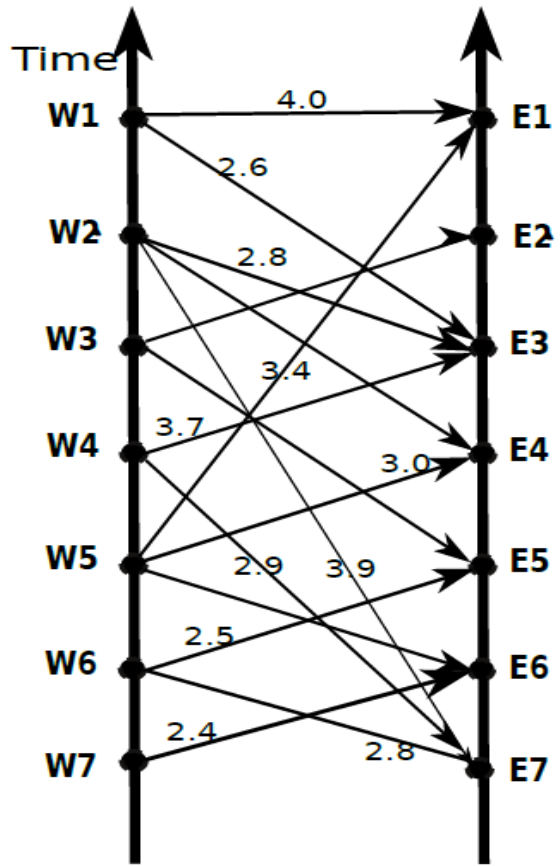
```
{
  GSR-13891,
  [
    03/08/13,
    Education,
    Civil unrest-
    Housing
    Non-Violent,
    (
      Brazil,
      Paraná,
      Ângulo
    )],
}
```

Date of Delivery
03/03/13

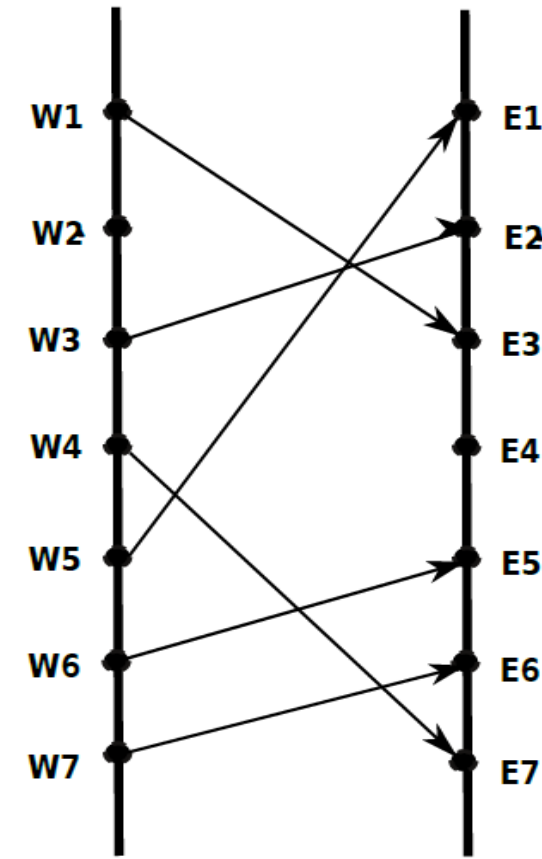
Lead-Time = 6

Earliest Reported Date
03/09/13

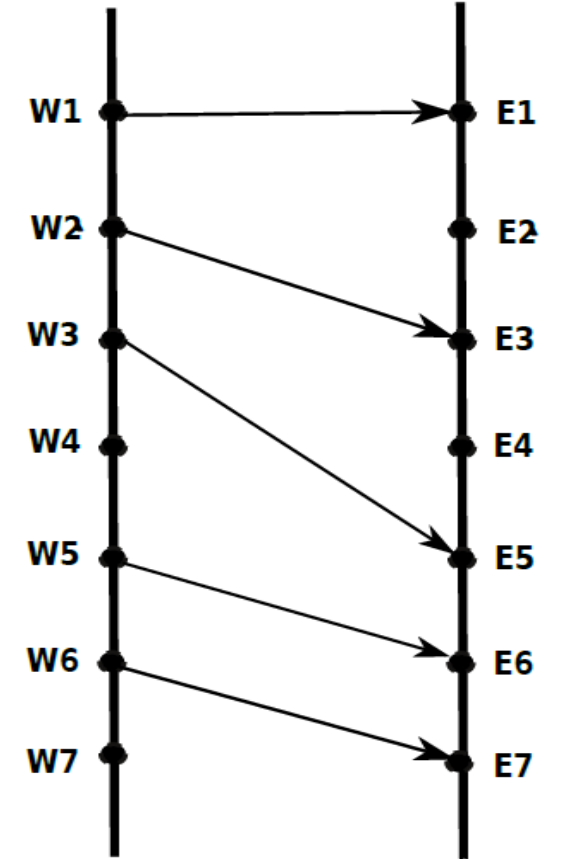
Matching alerts to events



Warnings **Events**
Candidate Matches



Warnings **Events**
Bipartite Matching



Warnings **Events**
Non-Crossing Matching

OSI Program Metrics

Targets

Metric	Month 12 4 months of warnings	Month 24 12 months of warnings	Month 36 12 months of warnings
Mean Lead-Time	1 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

Actual Results after 24 months

Metric		Month 24 12 months of warnings	
Mean Lead-Time		7.54 days	
Mean Probability Score		0.89	
Mean Quality Score		3.1	
Recall		0.65	
Precision		0.94	

EMBERS Architecture

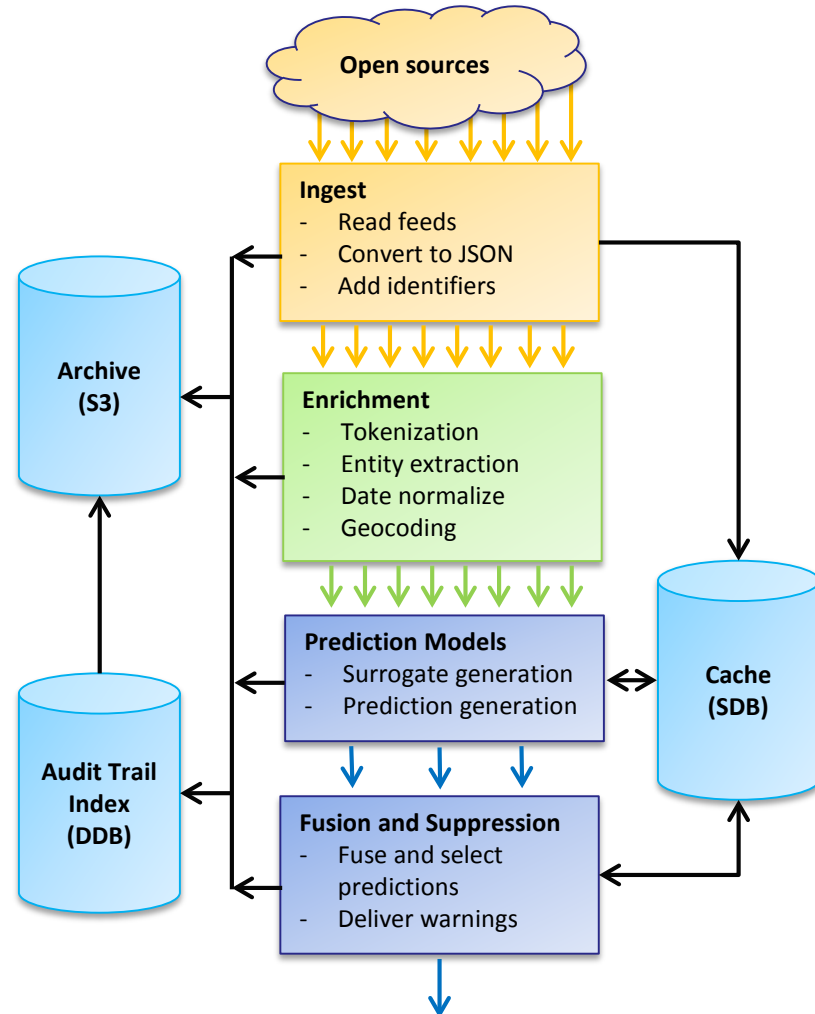
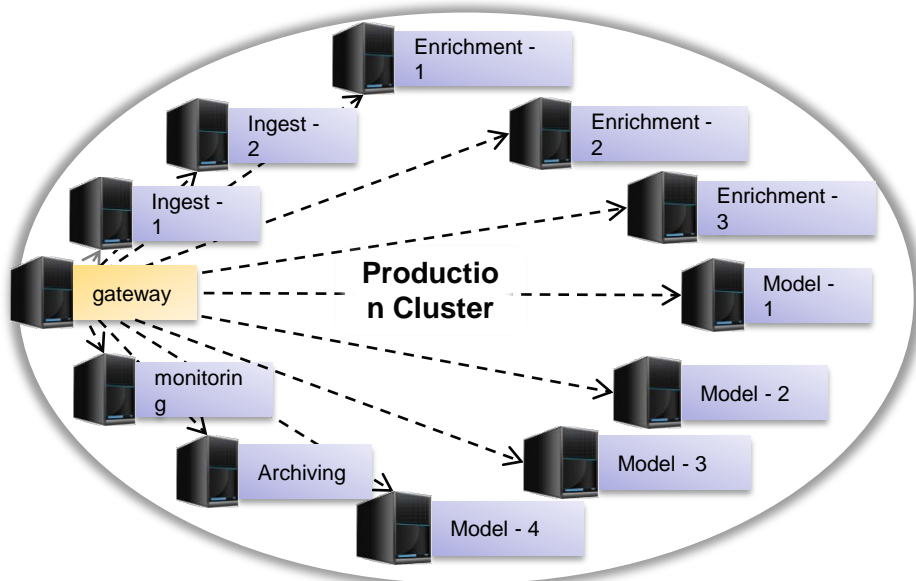


Over 12,000 warnings delivered

Average 40 warnings/day

12 EC2 instances used for production

16 vCPUs, 50 GB Virtual Memory



How we forecast civil unrest

Multiple models “chip away” at different portions of the protest modeling space, so their fusion yields high recall

How we forecast civil unrest

Multiple models “chip away” at different portions of the protest modeling space, so their fusion yields high recall

Data Sources



The GDELT Project

How we forecast civil unrest

Multiple models “chip away” at different portions of the protest modeling space, so their fusion yields high recall

Data Sources



The GDELT Project

Planned protest detection

Trabajadores de Southern Perú
amenazan con huelga indefinida desde
enero

Martes, 25 de diciembre de 2012 | 11:28 am



Mina no para. A pesar de huelga, minera continúa labores en refinería con personal contratado.

How we forecast civil unrest

Multiple models “chip away” at different portions of the protest modeling space, so their fusion yields high recall

Data Sources



The GDELT Project

Planned protest detection

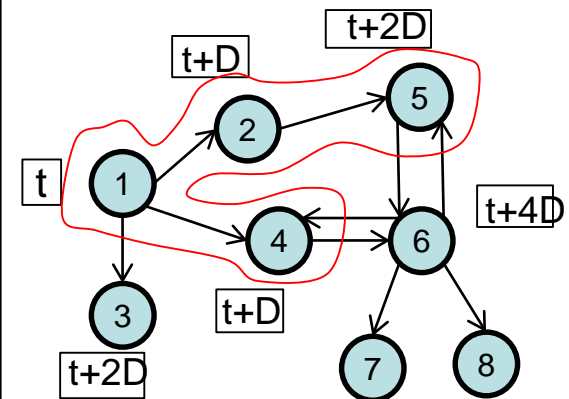
Trabajadores de Southern Perú
amenazan con huelga indefinida desde
enero

Martes, 25 de diciembre de 2012 | 11:28 am



Mina no para. A pesar de huelga, minera continúa labores en refinería con personal contratado.

Cascade regression (tracks online recruitment and viral spread)



How we forecast civil unrest

Multiple models “chip away” at different portions of the protest modeling space, so their fusion yields high recall

Data Sources



The GDELT Project

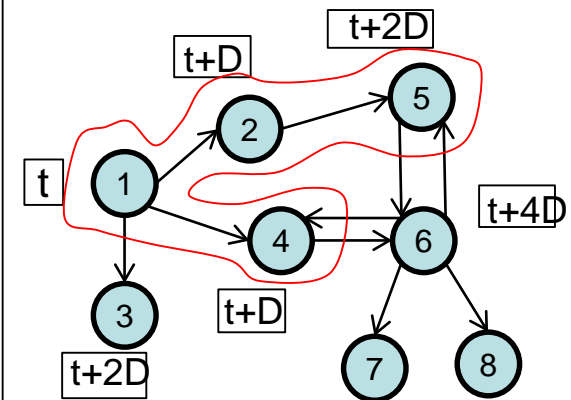
Planned protest detection

Trabajadores de Southern Perú amenazan con huelga indefinida desde enero



Miex no para. A pesar de huelga, minera continúa labores en refinería con personal contratado.

Cascade regression (tracks online recruitment and viral spread)



Dynamic query expansion

(automatically detects emerging keyword groups)



How we forecast civil unrest

Multiple models “chip away” at different portions of the protest modeling space, so their fusion yields high recall

Data Sources



The GDELT Project

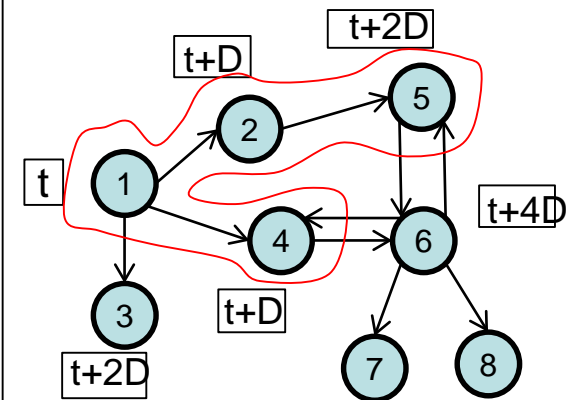
Planned protest detection

Trabajadores de Southern Perú amenazan con huelga indefinida desde enero



Mina no para. A pesar de huelga, minera continúa labores en refinería con personal contratado.

Cascade regression (tracks online recruitment and viral spread)



Dynamic query expansion

(automatically detects emerging keyword groups)



Volume-based model (LASSO approach)

Venezuela			
advertencia	asamble	boicotear	caminata
colegiatura	contingente	contrato	decreto
deforestacion	disturbio	empresarios	
enemigo	excepcional	fiscales	
gasolinera	huelga	implementar	
incinerar	malestar social	negociar	
paralizar	persecucion	pedras	
profesores	prohibicion		
proponemos	propuesta	refugiados	
relacion laboral	transporte de mercancias	vivienda	

How we forecast civil unrest

Multiple models “chip away” at different portions of the protest modeling space, so their fusion yields high recall

Data Sources



The GDELT Project

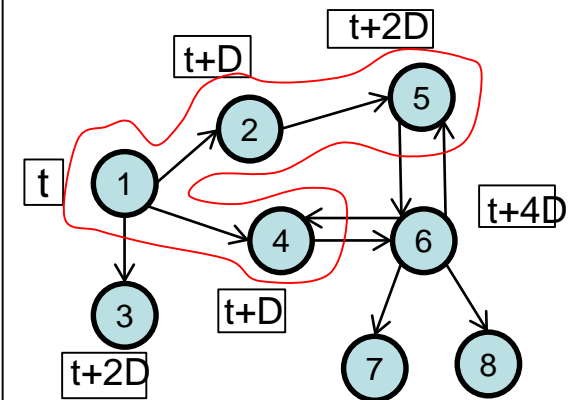
Planned protest detection

Trabajadores de Southern Perú amenazan con huelga indefinida desde enero



Miña no para. A pesar de huelga, minera continúa labores en refinería con personal contratado.

Cascade regression (tracks online recruitment and viral spread)



Dynamic query expansion

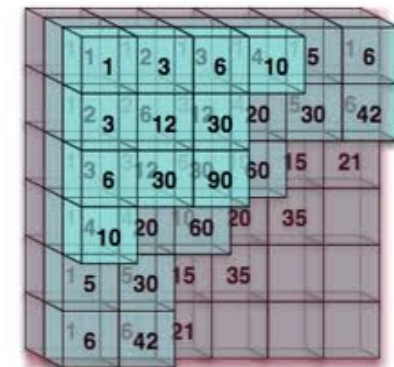
(automatically detects emerging keyword groups)



Volume-based model (LASSO approach)

Venezuela			
advertencia	asamble	boicotear	caminata
colegiatura	contingente	contrato	decreto
deforestacion	disturbio	empresarios	
enemigo	excepcional	fiscales	
gasolinera	huelga	implementar	
incinerar	malestar social	negociar	
paralizar	persecucion	pedras	
profesores	prohibicion		
proponemos	propuesta	refugiados	
relacion laboral	transporte de mercancias	vivienda	

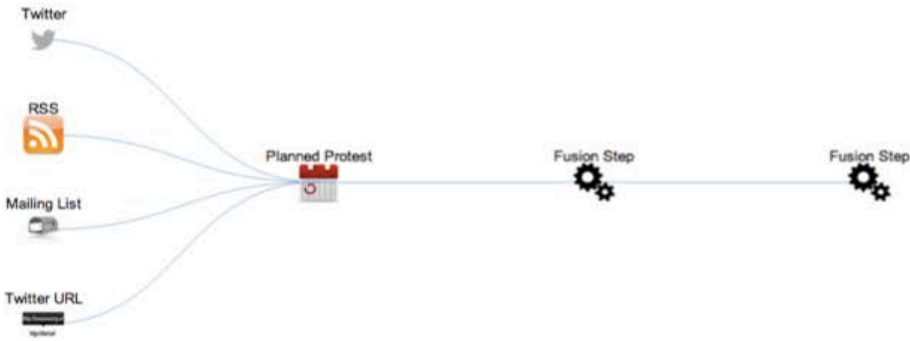
Baseline model (GSR-based)



Audit Trail Interface

Audit Trail: Prediction Pipeline Schematic

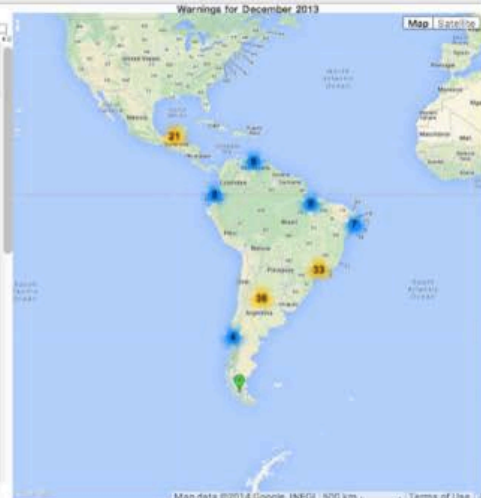
EMBERS Civil Unrest Visualizer IJI Visualizer Rare Disease Visualizer



Month: December 2013

Showing 1 to 25 of 126 records

Generated On	Event Date	Location	Population	Event Type	Violent	Model	Mile ID
2013-12-27	2013-12-30	Colombia, Antioquia, Medellin	General: Population	Employment & Wages	Yes	Social Scan	11048
2013-12-10	2013-12-22	Argentina, Entre Rios, Paysandú	Labor	Employment & Wages	No	Planned Protest	10857
2013-12-13	2013-12-30	Brazil, Sao Paulo, Sao Paulo	Labor	Other	No	Planned Protest	10992
2013-12-16	2013-12-20	Colombia, Bogota, Bogota	General: Population	Other	No	Planned Protest	10973
2013-12-17	2013-12-20	Brazil, Rio Grande Do Sul, Porto Alegre	Labor	Employment & Wages	No	Planned Protest	10987
2013-12-14	2013-12-17	Brazil, Santa Catarina, Florianopolis	Labor	Employment & Wages	No	Planned Protest	10959
2013-12-15	2013-12-17	Brazil, Rio Grande Do Sul, Porto Alegre	General: Population	Energy & Resources	No	LASSO	10961
2013-11-29	2013-12-15	Argentina, Buenos Aires, De la Plata	Labor	Employment & Wages	No	Planned Protest	10817
2013-11-28	2013-12-15	Argentina, Buenos Aires, De la Plata	General: Population	Other	No	Planned Protest	10481
2013-11-30	2013-12-15	Argentina, Buenos Aires, De la Plata	Business	Employment & Wages	No	Planned Protest	10550
2013-12-11	2013-12-15	Paraguay, Asuncion	General: Population	Energy & Resources	No	Dynamic Query	10960



Map

RSS Item Content

Trigger Phrase: anunciar huelga,decidan huelga
Tota Phrase: 20 de diciembre

¿Dónde estoy? | Diario El Mundo » Última Hora » Brasil | Pilotos y azafatas anuncian huelga para 20 de diciembre
 Pilotos y azafatas anuncian huelga para 20 de diciembre
 Sao Paulo (EFE) - Viernes 13 de diciembre 2013, 3:35 pm
 El Sindicato Nacional de Aeronautas (SNA), gremio que reúne a los pilotos y azafatas de Brasil, decretó hoy estado de huelga y anunció una paralización de sus actividades el próximo 20 de diciembre.
 El sindicato adoptó la decisión tras una asamblea realizada hoy con los bases gremiales en las cinco principales ciudades brasileñas, tras rechazar la propuesta de reajuste salarial presentada por las compañías aéreas, apuntó el SNA en un comunicado.
 Las compañías ofrecieron un reajuste salarial del 5,6 %, mientras que los pilotos y azafatas piden un aumento del 8 %, además de otras reivindicaciones en sus condiciones laborales.
 El miércoles 12 de diciembre, está prevista una nueva ronda de negociaciones entre el SNA y el Sindicato Nacional de las Empresas Aéreas (SNEA), la patronal del sector, con el objetivo de reanudar medidas económicas y sociales.
 El Sindicato Nacional de Aeronautas representa e alrededor de 18.000 pilotos y azafatas en todo el país.
 Diario El Mundo abre este espacio de opinión para que se pueda debatir, construir ideas y fomentar la reflexión. Por eso, pedimos que se evite hacer uso de ataques ofensivos, que incluyan malas palabras, o lo contrario nos reservamos el derecho de publicación.
 Recuerde que este es un medio que está para generar opinión constructiva. Deja un comentario

Pizza Hut

EL MUNDO .COM.SU

Inicio Enfoque Editorial & Opinión Nacionales Política Internacionales Economía Entretenimiento

¿Dónde estoy? | Diario El Mundo » Noticias de El Salvador » Última Hora » Brasil | Pilotos y azafatas anuncian huelga para 20 de diciembre

Brasil | Pilotos y azafatas anuncian huelga para 20 de diciembre

Sao Paulo (EFE)

Viernes 13, diciembre 2013 | 3:35 pm

El Sindicato Nacional de Aeronautas (SNA), gremio que reúne a los pilotos y azafatas de Brasil, decretó hoy estado de huelga y anunció una paralización de sus actividades el próximo 20 de diciembre.

El sindicato adoptó la decisión tras una asamblea realizada hoy con las bases gremiales de las cinco principales ciudades brasileñas, tras rechazar la propuesta de reajuste salarial presentada por las compañías aéreas, apuntó el SNA en un comunicado.

Las compañías ofrecieron un reajuste salarial del 5,6 %, mientras que los pilotos y azafatas piden un aumento del 8 %, además de otras reivindicaciones en sus condiciones laborales.

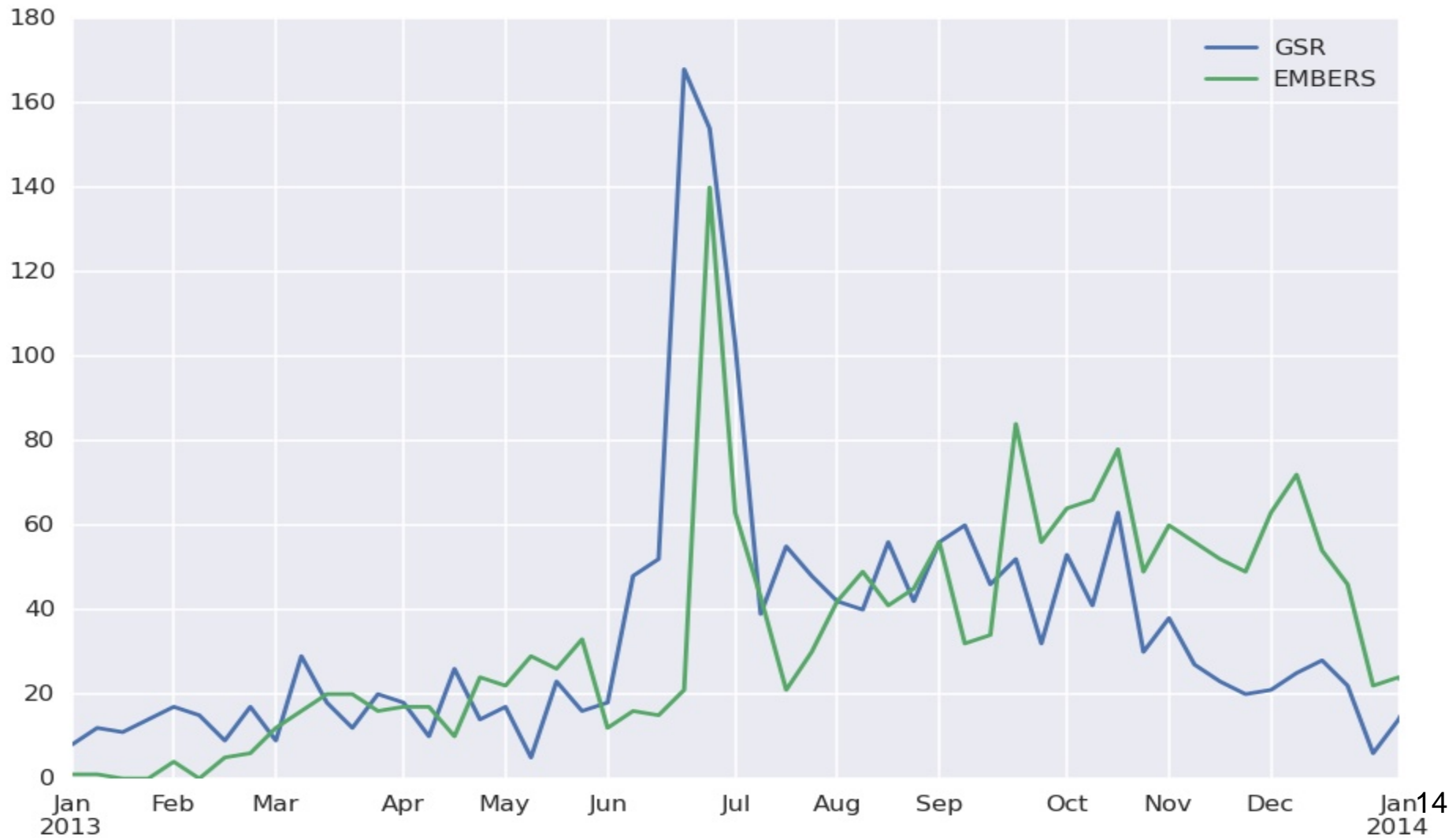
Tweeta el artículo

Diario El Mundo @Diario_ElMundo

Diario El Mundo @Diario_ElMundo

Diario El Mundo @Diario_ElMundo

How we did on the Brazilian Spring



How we did in Venezuela'14



Spread of protests (Venezuela'14)



Future Work

- Social science theory-based approaches
- Narrative generation
- Statistical theory of tradeoffs

- For more info
 - Contact naren@vt.edu