



Event Registry

Gregor Leban

Event registry

- <http://eventregistry.org/>
- Event registry je sistem za
 - Zbiranje globalnih novic v različnih jezikih
 - Avtomatsko identifikacijo svetovnih dogodkov iz zbranih novic
 - Ekstrakcijo glavnih informacij o dogodkih
- Bogate možnosti iskanja
- Širok spekter možnih vizualizacij podatkov



Kaj je dogodek?

- Dogodek = nekaj pomembnega, kar se je zgodilo v svetu
- Predpostavka: Če je nekaj dovolj pomembno, bodo o tem poročali različni mediji
- Dogodke torej identificiramo kot zbirke člankov, ki govorijo o isti stvari, ki se je zgodila



Flydubai plane crashes in Russia; 62 dead

WHEN: Sat, March 19, 2016

WHERE: Rostov-on-Don, Russia

CATEGORIES: Society→Issues→Transportation , Science→Technology→Aerospace , Business→Transportation and Logistics→Aviation

ENTITIES: Rostov-on-Don, Russia Dubai, United Arab Emirates Russia Boeing Flydubai

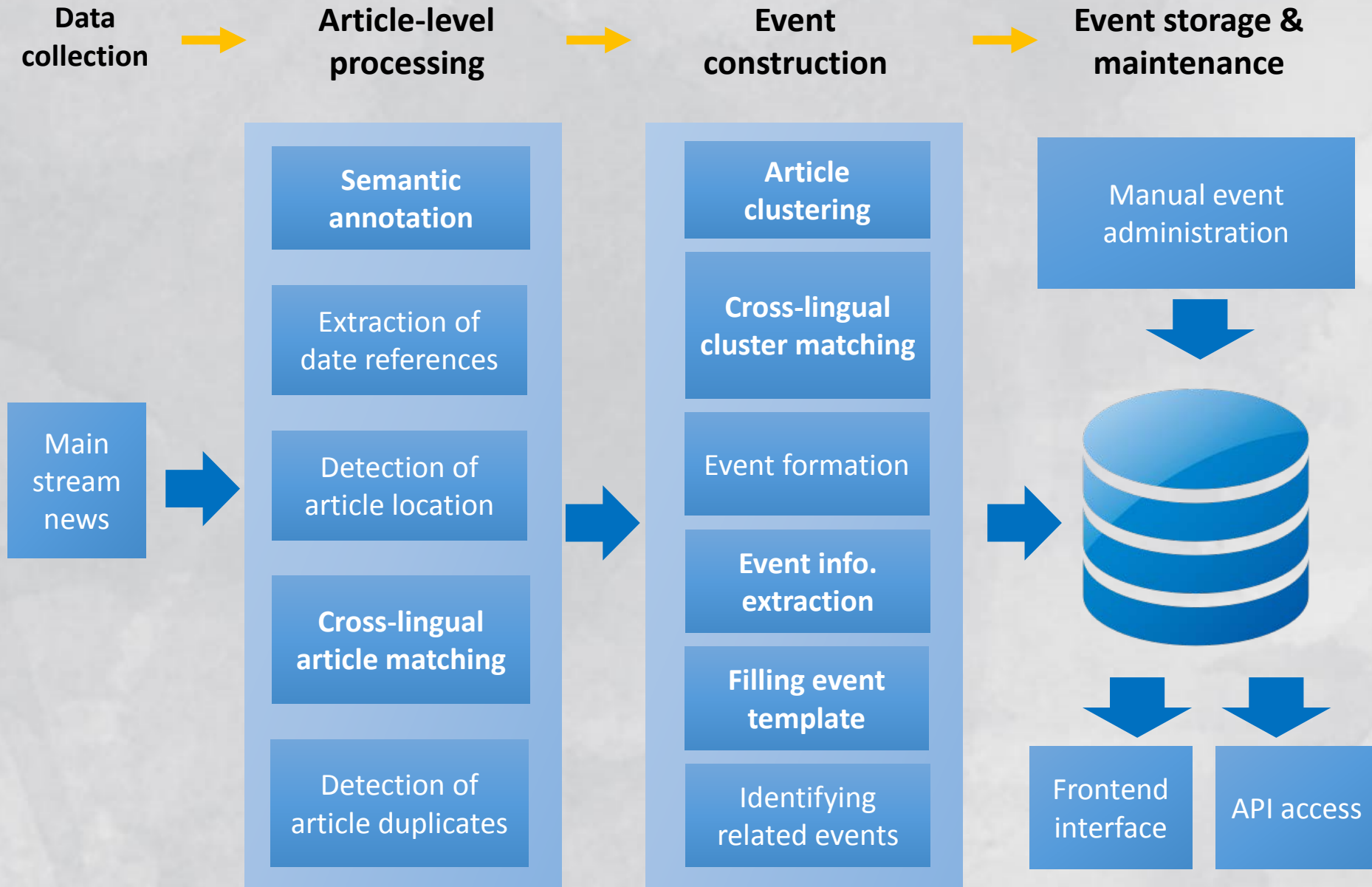
KEYWORDS: Airport Boeing 737 Fixed-wing aircraft Airline Runway

ARTICLES: 815 (362 English, 92 German, 106 Spanish, 21 Portugese, 28 Italian, 87 French, 5 Russian, 32 Turkish, 66 Chinese, 16 Croatian)

A CCTV camera captured the moment the aircraft exploded on the runway A FlyDubai Boeing 737-800 has crashed in a fireball while it attempted to land at a Russian airport. The passenger jet was believed to be making its second attempt to land at Rostov-on-Don airport in poor visibility. A video published on Youtube appears to show a fireball as the aircraft crashes into the ground. This is believed to be the moment the crashed on landing at Rostov-on-Don airport in Russia The aircraft...

🔥 609

Event registry pipeline



Zbiranje novic

- <http://newsfeed.ijs.si/>
- Novice iz ~30.000 medijev po celem svetu
- Na dan zberemo ~300.000 člankov v 15 jezikih:
eng, ger, spa, chi, slo, cat, por, ita,
fra, rus, ara, tur, cro, ser

Semantična anotacija novic

Andrew Kay Dies at 95; Pioneer in Computing

By JOHN MARKOFF

Andrew Kay, the designer of an early portable computer, the Kaypro II, that became a smash hit in the early 1980s before his company fell into bankruptcy in the '90s as the computer industry leapfrogged ahead of him, died Aug. 28 in Vista, Calif. He was 95.

His death was confirmed by his son Allan.

For a time, Mr. Kay's company, Kaypro, was the world's largest portable computer maker, ranked fourth in the PC industry over all behind IBM, Apple Computer and RadioShack.

Early on, the Kaypro II developed a loyal following. Arthur C. Clarke used a Kaypro to write his novel "2010: Odyssey Two"; Stephen M. Case, who would become chief executive of America Online, bought one as his first computer, adding a 300-baud modem to explore the fledgling online world.

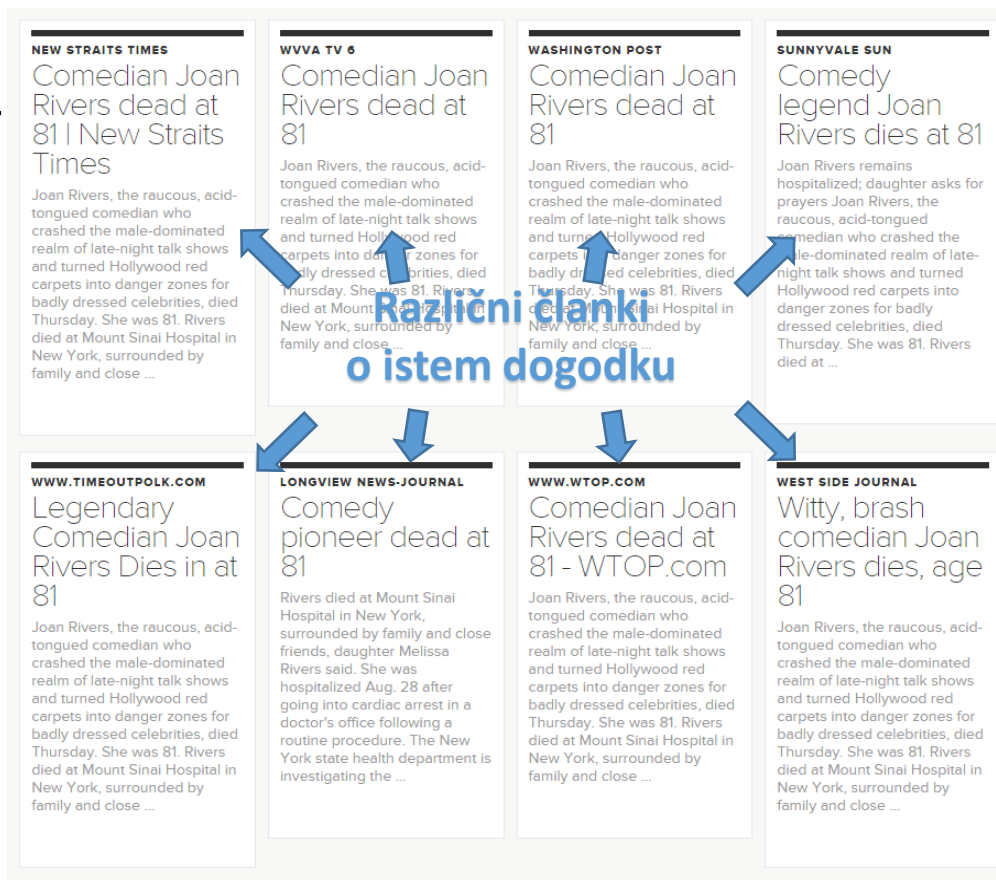
"The Kaypro computer was a necessary step in getting to the iPad," Paul Freiberger, co-author of "Fire in the Valley: The Making of the Personal Computer" said in an interview on Friday. "Back then few thought of making a



Andrew Kay's company Kaypro surged in the 1980s but fell into bankruptcy in the early 1990s. Nancee E. Lewis/U-T San Diego

Iskanje gruč v novicah (clustering)

- Članke v vsakem jeziku združimo v skupine, kjer vsaka skupina novice opisuje isti dogodek
- Pri gručenju primerjamo
 - Naslov in vsebino članka, ter
 - Semantične anotacije



Čezjezično povezovanje gruč

- Gruče v različnih jezikih lahko govorijo o istem dogodku

English cluster



SpaceX lands its third consecutive rocket on ship in Atlantic Ocean

Yes, SpaceX has done it again. For the fourth time, it has managed to have a successful rocket landing. On Friday at 5:40pm Falcon 9 rocket blasted off from Cape Canaveral, Florida, carrying the Thaicom 8 satellite.

The rocket returned to earth around 10 minutes later. It made a landing in an ...

PERFSCIENCE.COM

Sat, 28 May, 08:45

0.847



SpaceX launches Asian satellite, attempts to land rocket

© SPACEX/AFP/FileThe Falcon 9 rocket on December 16, 2015 in Cape Canaveral -- SpaceX has managed to successfully land the first stage of its Falcon 9 rockets three times before

Miami (AFP) - SpaceX on Friday launched an Asian communications satellite before trying for a fourth successful landing ...

BUSINESS INSIDER

Fri, 27 May, 22:33

0.836



SpaceX launches Asian satellite, attempts to land rocket

Miami (AFP) - SpaceX on Friday launched an Asian communications satellite before trying for a fourth successful landing of its Falcon 9 rocket.

Under blue skies dotted with clouds, the shiny white rocket blasted off from Cape Canaveral, Florida at 5:40 pm (21:40 GMT).

The satellite, known as ...

YAHOO

Fri, 27 May, 22:04

0.811



SpaceX makes fourth successful rocket landing

That would make space deliveries - and travel - more economical

Elon Musk's SpaceX company notched up another victory today when it managed to safely land a rocket for the fourth time - after it had deposited an Asian communications satellite into distant orbit.

Under blue skies dotted with ...

MAIL ONLINE

Sat, 28 May, 01:29

0.807

f 14

Spanish cluster



SpaceX lo ha vuelto a hacer

La empresa californiana de transporte aeroespacial logra poner en órbita otro satélite de comunicaciones y hacer aterrizar en medio del mar la primera fase de su cohete Falcon 9.

Ya es la tercera vez y la compañía liderada por Elon Musk vuelve a anotar otro crucial tanto en su carrera ...

HIPERTEXTUAL

Fri, 27 May, 22:12

0.689

f 51



Hacer de lo imposible algo rutinario, nueva misión para SpaceX

El cohete Falcon 9 de la compañía privada SpaceX despegará desde la estación de las fuerzas aéreas de Cabo Cañaveral este jueves con el objetivo de poner en órbita satélites de comunicación. Un paso más para normalizar lo que hace un año se creía imposible: reutilizar cohetes.

SpaceX, ...

HIPERTEXTUAL

Thu, 26 May, 08:32

0.747

f 47

SpaceX consigue recuperar otro cohete tras su lanzamiento

Este es el video (acelerado) del retorno al espaciuuerto autónomo flotante Of Course I Still Love You de la primera etapa del Falcon 9 FT de SpaceX que puso en órbita el satélite de telecomunicaciones Thaicom 8.

El Thaicom 8 tiene como destino final una órbita geoestacionaria, por lo que el ...

MICROSIERVOS.COM

Sat, 28 May, 13:31

0.593

f 8



Se aplaza por 24 horas lanzamiento de satélite tailandés

El director de SpaceX, Elon Musk, anunció que su empresa decidió aplazar por al menos 24 horas el lanzamiento de del satélite tailandés Thaicom 8 como consecuencia de una falla en un motor. El despegue se iba a llevar a cabo desde Cabo Cañaveral (Florida).

"Se trata de una pequeña anomalía ...

ELESPECTADOR

Fri, 27 May, 04:30

0.839

f 4

?

=

Izločanje informacij o posameznih dogodkih

- Z analiziranjem člankov določimo različne Informacije o dogodku:
 - Lokacija
 - Datum
 - Ključne osebe, podjetja, besede
 - Ključne kategorije (politika, naravna katastrofa, ...)
 - ...

DEMO (<http://eventregistry.org/>)

Current articles and events

Any language ▾ See large map ↗

Afghan police killed by US strike - Afghan officials - Courier Mail

Location: Tarinkot, Afghanistan

Articles about event: 60

Edo Guber Watch: INEC, security chiefs meet

Location: Edo State, Nigeria

Mentioned in 11 articles

Sam Bradford leads Vikings over Packers 17-14 in Minnesota debut

Location: Minnesota, United States

Mentioned in 192 articles

Syria ceasefire seen near collapse

Location: Aleppo, Syria

Mentioned in 19 articles

U.S. Marine shot in head in South LA while on military leave

Location: Segovia, Spain

Mentioned in 17 articles

FJSC okays Onnoghen as next QJN - Vanguard News

Location: (unknown)

Mentioned in 7 articles

EVENT REGISTRY STATISTICS

Data info	#
Imported articles:	141,227,827
Identified events:	5,007,526
News sources:	154,160
Unique concepts:	2,692,089
Categories:	5,015

RECENTLY ADDED ARTICLES

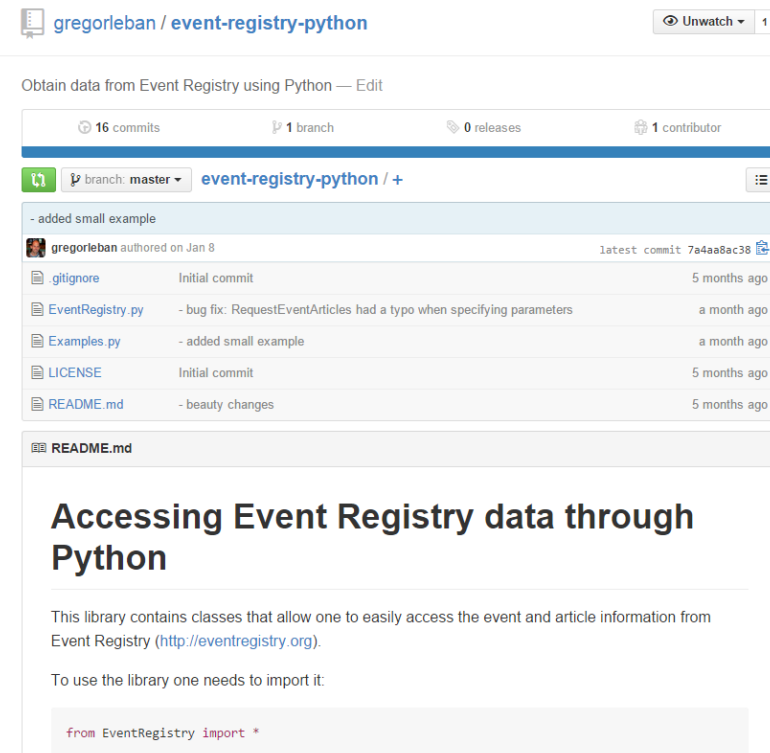
Date	Articles
Aug 21	150,000
Aug 22	250,000
Aug 23	280,000
Aug 24	270,000
Aug 25	260,000
Aug 26	250,000
Aug 27	240,000
Aug 28	230,000
Aug 29	220,000
Aug 30	210,000
Aug 31	200,000
Sep 01	250,000
Sep 02	270,000
Sep 03	280,000
Sep 04	290,000
Sep 05	280,000
Sep 06	270,000
Sep 07	260,000
Sep 08	250,000
Sep 09	240,000
Sep 10	230,000
Sep 11	220,000

RECENTLY IDENTIFIED EVENTS

Date	Events
Aug 21	4,500
Aug 22	8,500
Aug 23	7,500
Aug 24	7,000
Aug 25	6,500
Aug 26	6,000
Aug 27	5,500
Aug 28	5,000
Aug 29	4,500
Aug 30	4,000
Aug 31	3,500
Sep 01	8,500
Sep 02	8,000
Sep 03	7,500
Sep 04	7,000
Sep 05	6,500
Sep 06	6,000
Sep 07	5,500
Sep 08	5,000
Sep 09	4,500
Sep 10	4,000
Sep 11	3,500

API access

- Content in Event Registry can be accessed using the API
- To simplify access, a Python library is available at:
<https://github.com/gregorleban/EventRegistry>



gregorleban / event-registry-python

Obtain data from Event Registry using Python — Edit

16 commits 1 branch 0 releases 1 contributor

branch: master event-registry-python / +

- added small example

gregorleban authored on Jan 8 latest commit 7a4aa8ac38

.gitignore	Initial commit	5 months ago
EventRegistry.py	- bug fix: RequestEventArticles had a typo when specifying parameters	a month ago
Examples.py	- added small example	a month ago
LICENSE	Initial commit	5 months ago
README.md	- beauty changes	5 months ago

README.md

Accessing Event Registry data through Python

This library contains classes that allow one to easily access the event and article information from Event Registry (<http://eventregistry.org>).

To use the library one needs to import it:

```
from EventRegistry import *
```

API access example (1)

```
>>> from EventRegistry import *
>>> er = EventRegistry()
>>> q = QueryEvents()
# get events related to Barack Obama
>>> q.addConcept(er.getConceptUri("Obama"))
# and are related to issues in society
>>> q.addCategory(er.getCategoryUri("society issues"))
# and have been reported by the BBC
>>> q.addNewsSource(er.getNewsSourceUri("bbc"))
# return event details for first 30 events
>>> q.addRequestedResult(RequestEventsInfo(page = 0, count = 30))
# execute query and obtain results
>>> res = er.execQuery(q)
```

API access example (2)

```
'events': { 'resultCount': 122,  
  'results': [  
    {'articleCounts': {'eng': 54.0, 'total': 54.0},  
     'categories': [{...}],  
     'concepts': [{...}, ...],  
     'eventDate': '2014-08-29',  
     'eventDateEnd': '',  
     'multiLingInfo': { 'eng': {  
       'title': ..., 'summary': ...}},  
     'uri': '1211229', 'wgt': 9.0}  
    }, ...  
  ]  
}
```

API access example (3)

```
# get information about event with ID 123
```

```
>>> q = QueryEvent("123");
```

```
# return concept labels in 3 languages
```

```
>>> q.addRequestedResult(RequestEventInfo(["eng", "spa", "slv"]))
```

```
# get 10 most central articles
```

```
>>> q.addRequestedResult(RequestEventArticles(0, 10))
```

```
# get information how articles about the event were trending
```

```
>>> q.addRequestedResult(RequestEventArticleTrend())
```

```
# get top keywords
```

```
>>> q.addRequestedResult (RequestEventKeywordAggr())
```

```
>>> eventRes = er.execQuery(q);
```

API access example (4)

```
>>> q = QueryArticles();  
# articles should be from a particular time period  
>>> q.setDateLimit(datetime.date(2014, 4, 16),  
datetime.date(2014, 4, 28))  
# they should mention apple  
>>> q.addKeyword("apple")  
# they should also mention iphone  
>>> q.addKeyword("iphone")  
# get top 30 articles that match criteria  
>>> q.addRequestedResult(RequestArticlesInfo(page=0, count  
= 30));  
>>> res = er.execQuery(q)
```

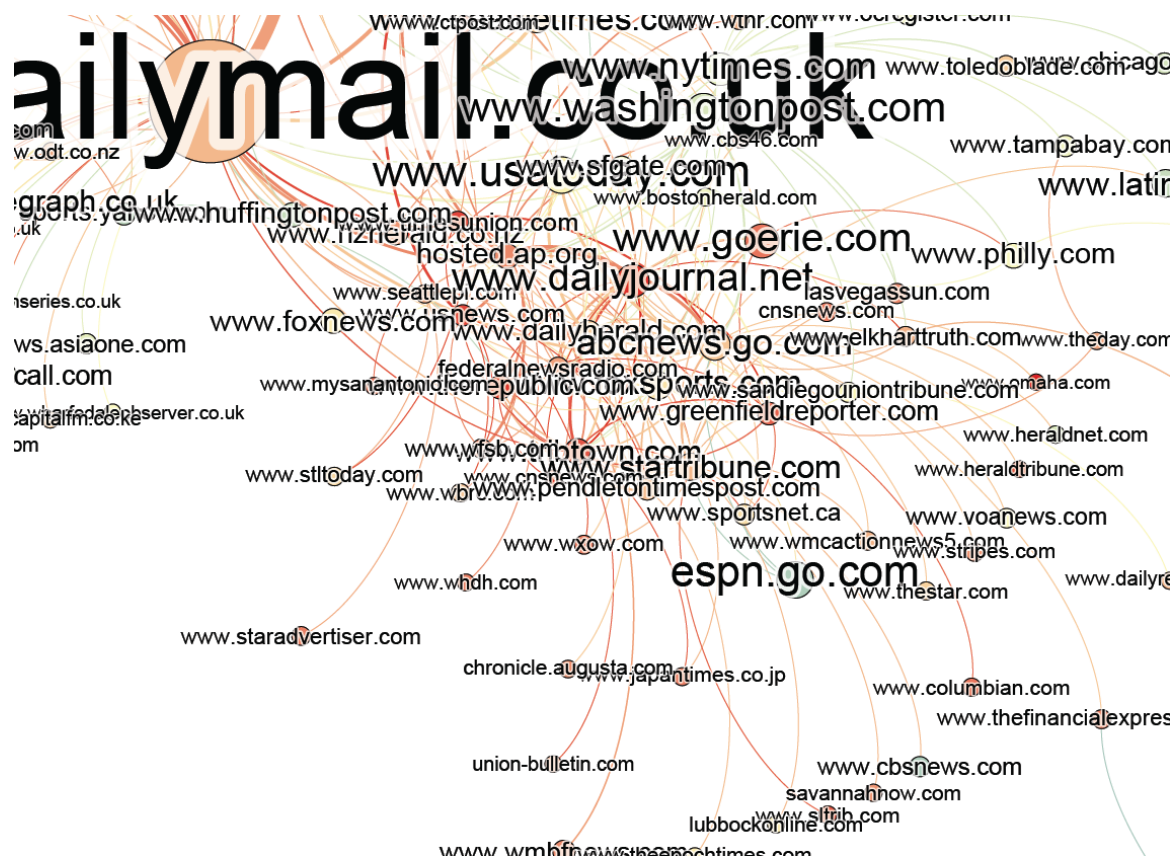
Future / Ongoing work

Detection of news bias

- Various ways in which news publishers can be “biased”
- Examples:
 - Gender bias
 - News source citations
 - Topic bias
 - Geographical bias
 - Vocabulary bias


Information diffusion/propagation

- Analysis of how information propagates from one publisher to the next



Infobox extraction / slot filling

- Events are of different types (e.g. earthquakes, mergers and acquisitions, product releases, ...)
- For each event type we can define a specific set of properties (slots) that should be extracted

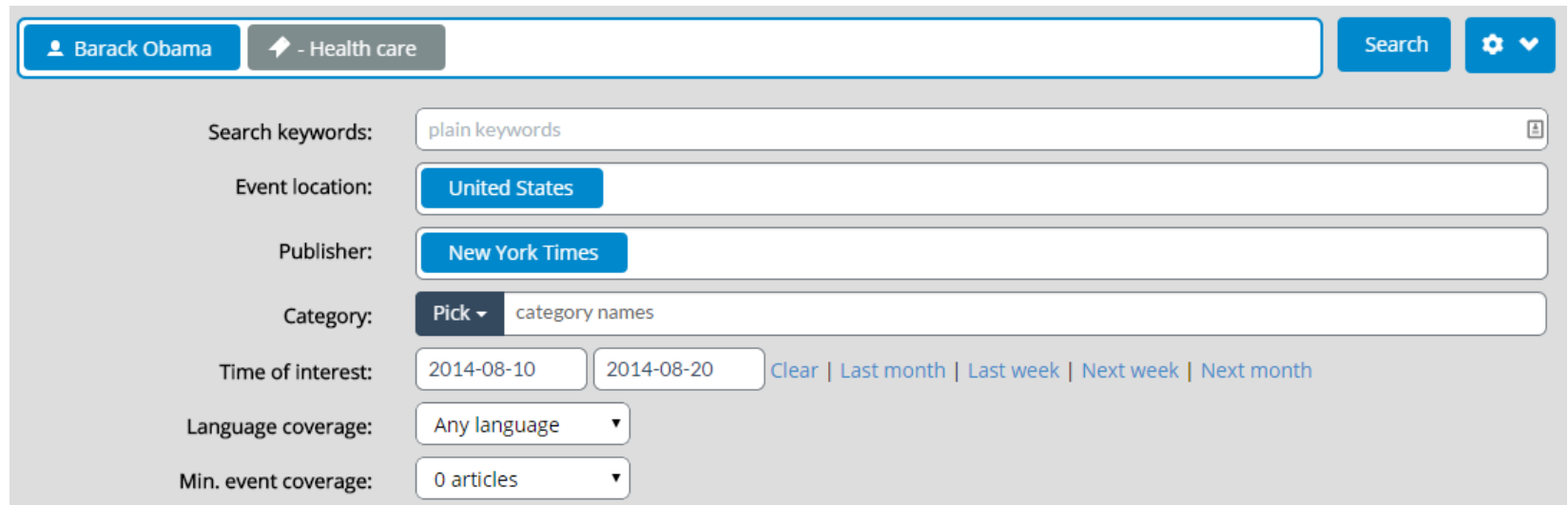
Date	11 March 2011 (3 years, 5 months and 2 weeks)
Origin time	14:46:24 JST (UTC+09:00)
Duration	6 minutes ^[1]
Magnitude	9.0 (M_w) ^{[2][3]}
Depth	30 km (19 mi)
Epicenter	 38.322°N 142.369°E
Type	Megathrust earthquake
Areas affected	Japan (primary) Pacific Rim (tsunami, secondary)
Total damage	Tsunami wave, flooding, landslides, fires, building and infrastructure damage, nuclear incidents including radiation releases
Max. intensity	IX - Violent
Peak acceleration	2.99 g
Tsunami	Up to 40.5 m (133 ft) in Miyako, Iwate, Tōhoku
Landslides	None
Foreshocks	7

Causality

- Goal: **Finding causal templates in event sequences**
- Example 1: campaign events → election day → inauguration
- Example 2: flood → outbreak of a disease
- How to do it:
 - For each event we can identify related events
 - Generalize events to find common patterns

Event registry

- Event registry provides extensive search options...



The screenshot displays a search interface for an event registry. At the top, there is a navigation bar with a blue button labeled "Barack Obama" and a grey button labeled "- Health care". To the right of these buttons is a search bar with the text "Search" and a settings icon. Below the navigation bar, the search criteria are listed on the left, and the corresponding input fields are on the right. The criteria include:

- Search keywords:** A text input field containing "plain keywords".
- Event location:** A dropdown menu with "United States" selected.
- Publisher:** A dropdown menu with "New York Times" selected.
- Category:** A dropdown menu with "Pick" and "category names" visible.
- Time of interest:** Two date input fields showing "2014-08-10" and "2014-08-20", followed by links for "Clear", "Last month", "Last week", "Next week", and "Next month".
- Language coverage:** A dropdown menu with "Any language" selected.
- Min. event coverage:** A dropdown menu with "0 articles" selected.

Content to display:

List of events 

List of events

The list contains the information about events that match the search criteria.

18,833 events match the search criteria



First < 1 2 3 4 5 6 7 8 > Last

RECAP: Doctors investigating first ebola case in Scotland after aid worker returns from West Africa

WHEN: December 30, 2014

WHERE: London, United Kingdom

ENTITIES: Republic of Sierra Leone Glasgow Scotland London West Africa Casablanca British Airways Republic of Liberia Republic of Guinea Kingdom of Morocco

KEYWORDS: Ebola virus disease Hospital United Kingdom Infection Patient Virus Disease London Heathrow Airport Therapy Health

ARTICLES: 721 (575 eng, 87 ger, 45 spa, 14 chi, 0 slo)

The authorities will next be contacting passengers on the two flights connecting Casablanca to Heathrow, and the London airport to Glasgow with all relevant precautionary information. LIVE updates as the first ebola case has been confirmed in an aid worker seeking medical assistance in a Glasgow hospital. 7:44 pm THIS year's Ebola outbreak is the deadliest occurrence of the disease since it was first discovered in 1976. The current outbreak has killed more than all other Ebola outbreaks...

US Withdrawing Most Troops Fighting Ebola in West Africa

WHEN: February 11, 2015

WHERE: Republic of Sierra Leone

ENTITIES: West Africa United States Republic of Liberia Barack Obama Republic of Guinea Republic of Sierra Leone World Health Organization United States Armed Forces Washington, D.C. United States Department of Defense

KEYWORDS: Ebola virus disease Epidemic Disease President of the United States The Pentagon White House Virus Infection Health Viral disease

ARTICLES: 358 (296 eng, 18 ger, 44 spa, 0 chi, 0 slo)

WASHINGTON -- The United States is preparing to withdraw nearly all of its troops fighting the Ebola outbreak in West Africa, the White House said Tuesday, as the global health crisis recedes amid a sharp decline in Ebola cases. Of the 2,800 troops the U.S. deployed, just 100 will remain in West Africa after April 30, officials said. About 1,500 of those troops have already returned home. Those staying in West Africa will work with Liberia's military, regional partners and U.S. civilians to...

↗ Ebola virus disease

Search



Search keywords:

plain keywords

Event location:

city or country names

Publisher:

names of the publishers

Category:

Pick ▾

category names

Time of interest:

start date

end date

[Clear](#) | [Last month](#) | [Last week](#) | [Next week](#) | [Next month](#)

Language coverage:

Any language ▾

Min. event coverage:

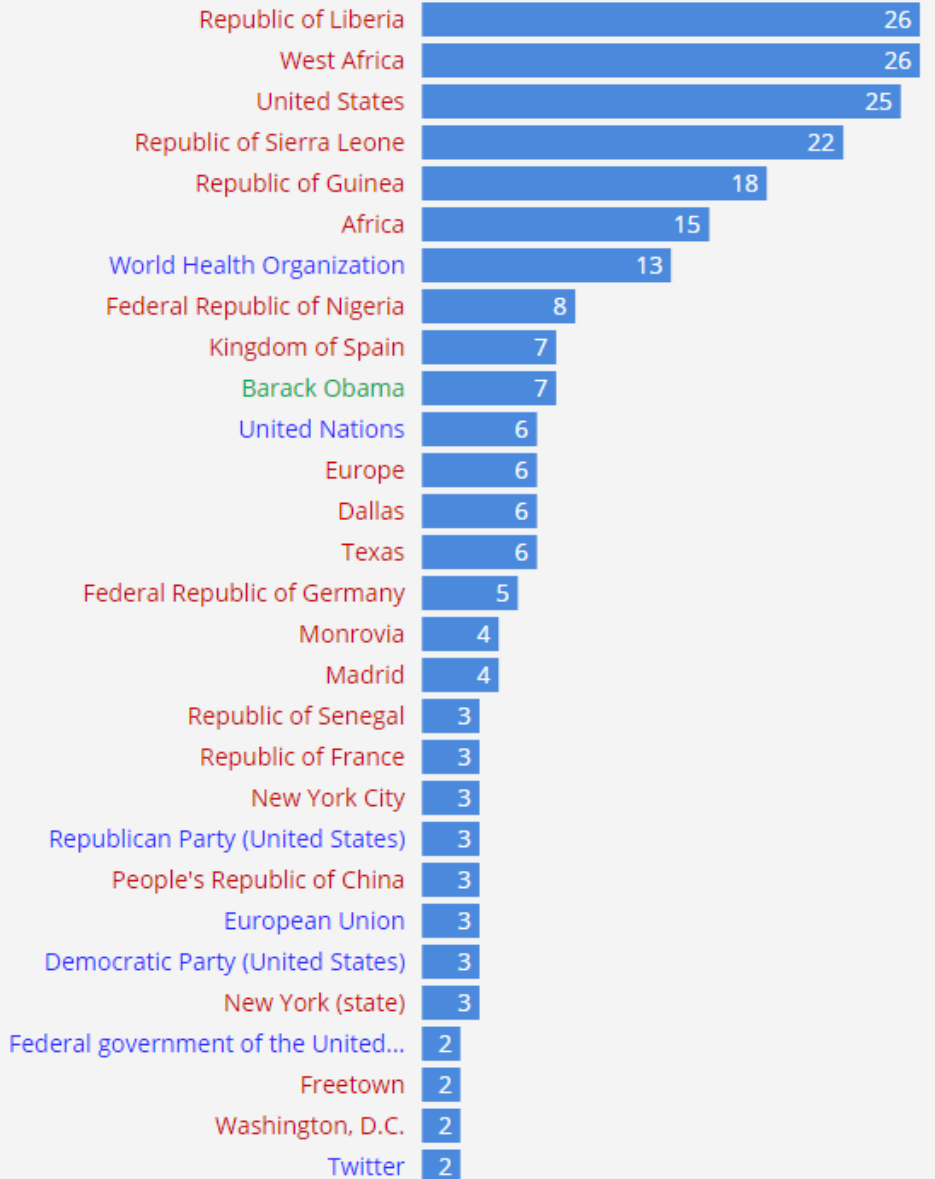
0 articles ▾

Top concepts in events

List of concepts (entities and keywords) that best describe what the events are about

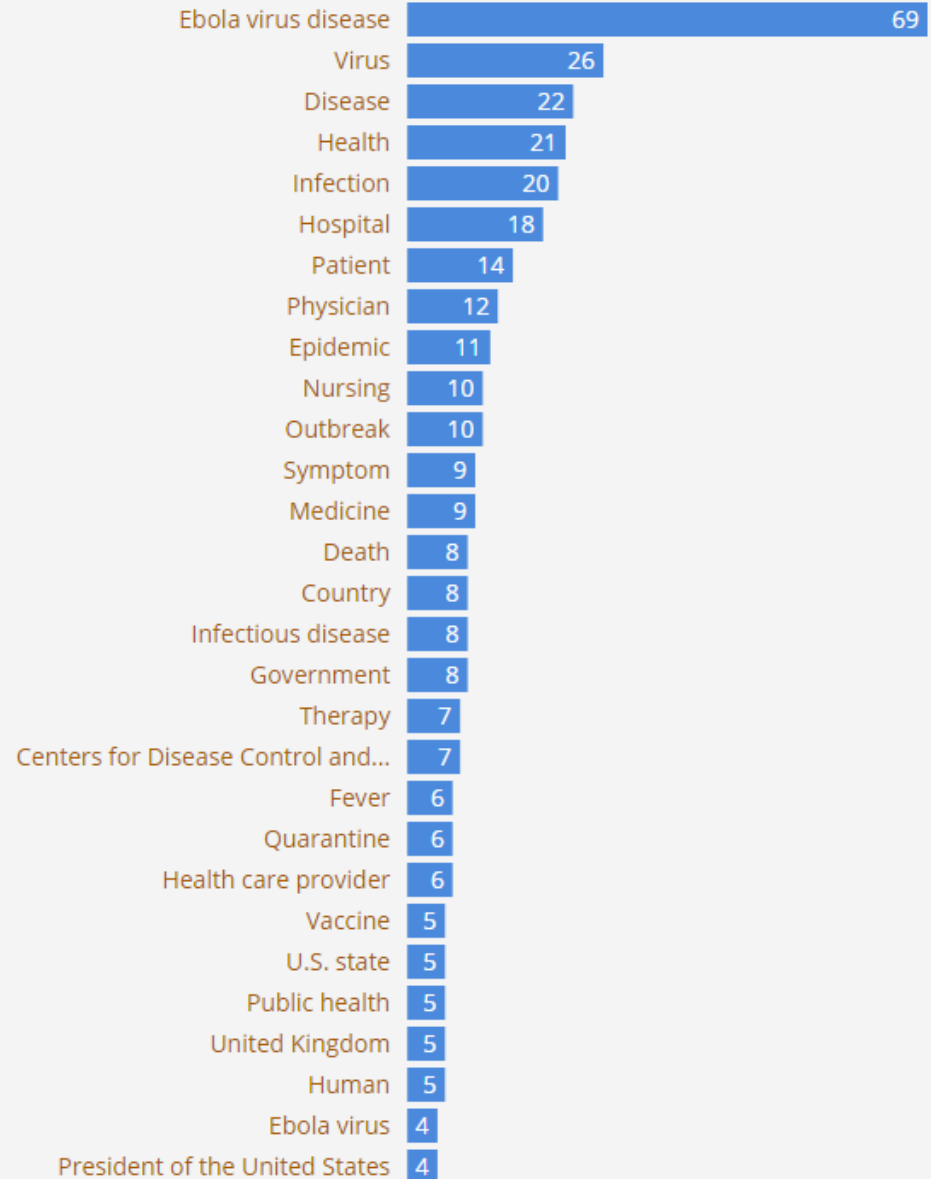
ENTITIES

Concept relevance 0 2 4 6 8 10 12 14 16 18 20 22 24 26



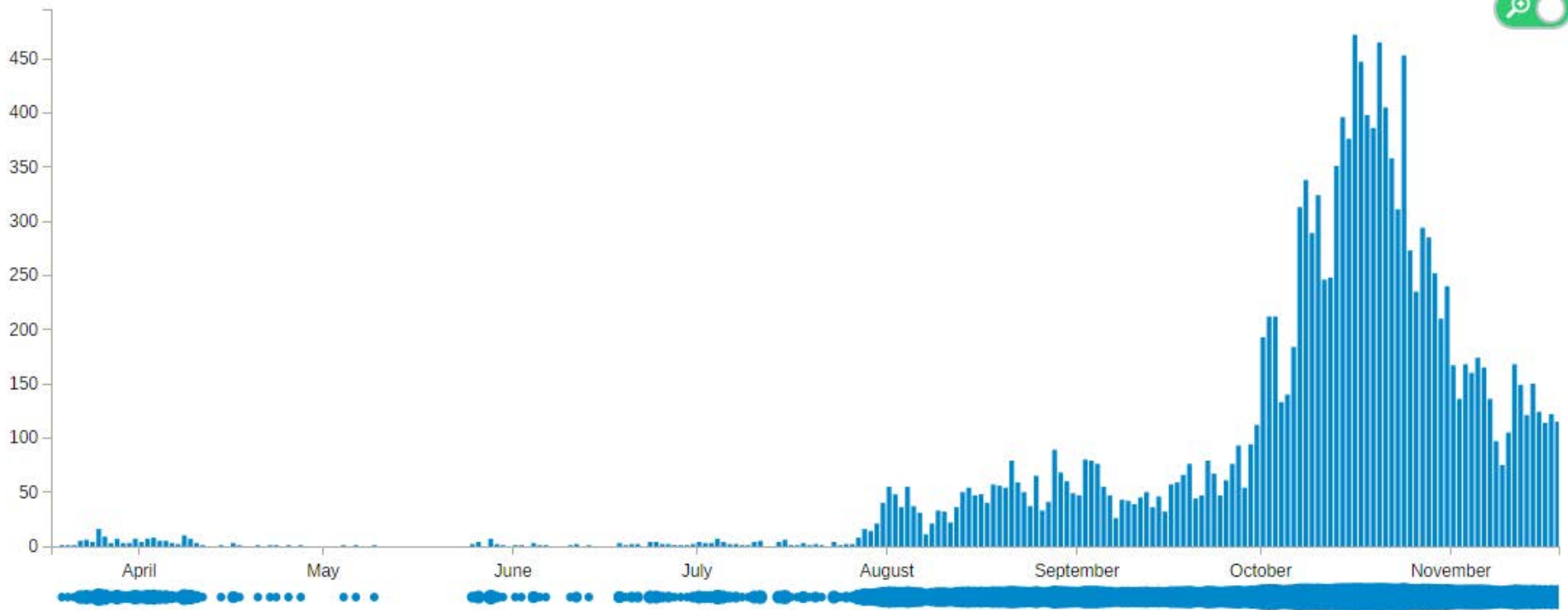
KEYWORDS

Concept relevance 0 10 20 30 40 50 60



Timeline of events

See when the events that match the search criteria happened. Click a particular column to display a list of top events for that date.

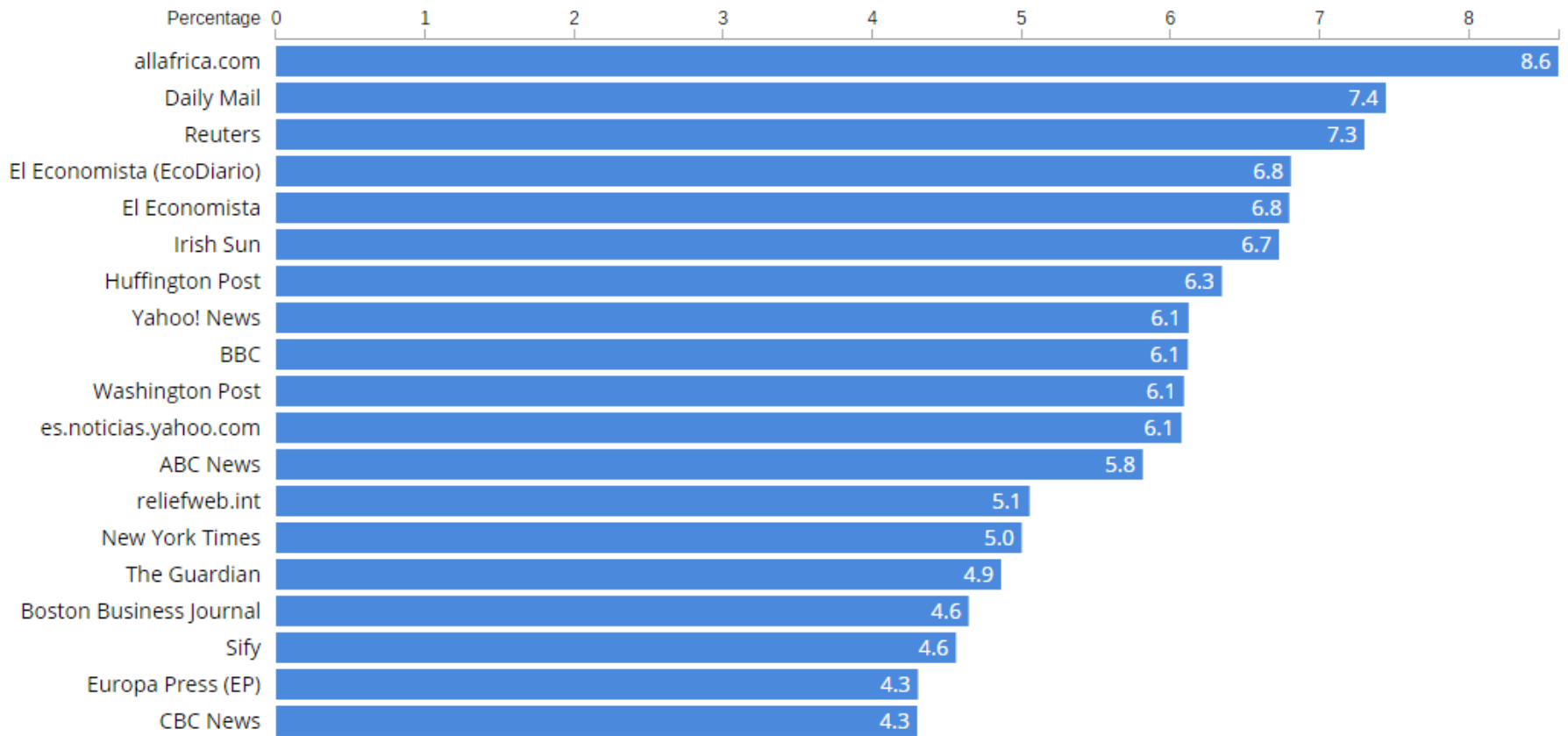


Top publishers

The information displays who are the publishers that report the most about the resulting events. The numerical values represents the percent of the resulting events that are covered by a particular publisher. The geographical map shows the locations of the publishers and the circle sizes correspond to percentages.

Ranked list

Geographical map

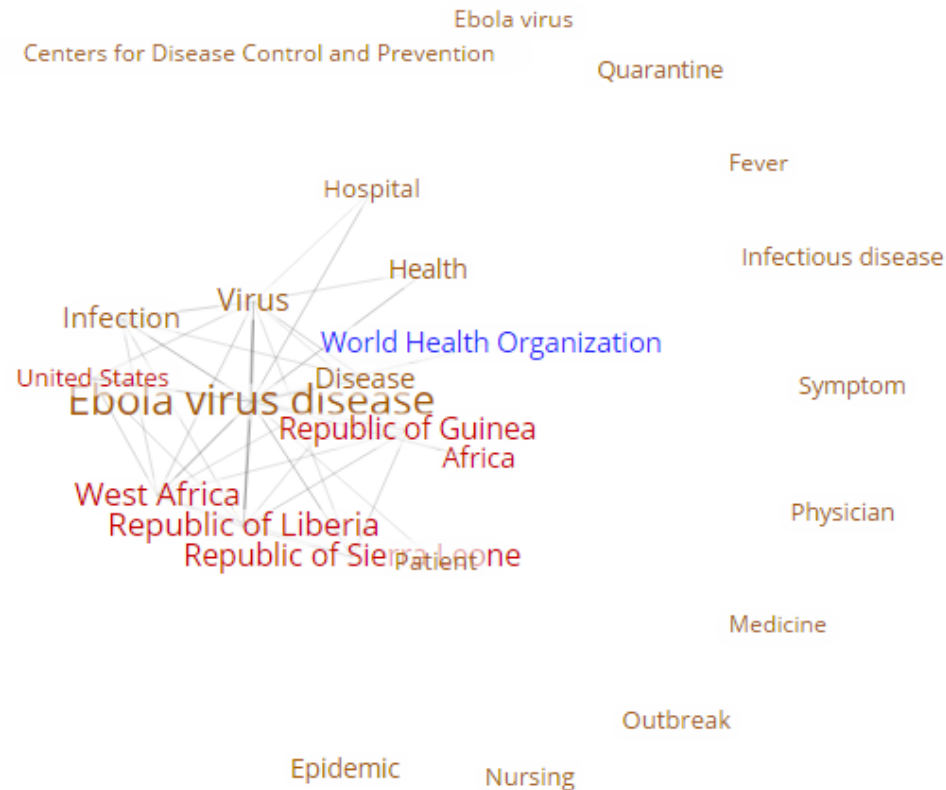


Top entities and relations

The nodes in the graph are the most relevant concepts in the events. The edges are displayed between the concepts that frequently co-occur in the same events.

Concept types to display

- Person
- Location
- Organization
- Nonentities



Trending of concepts in search results

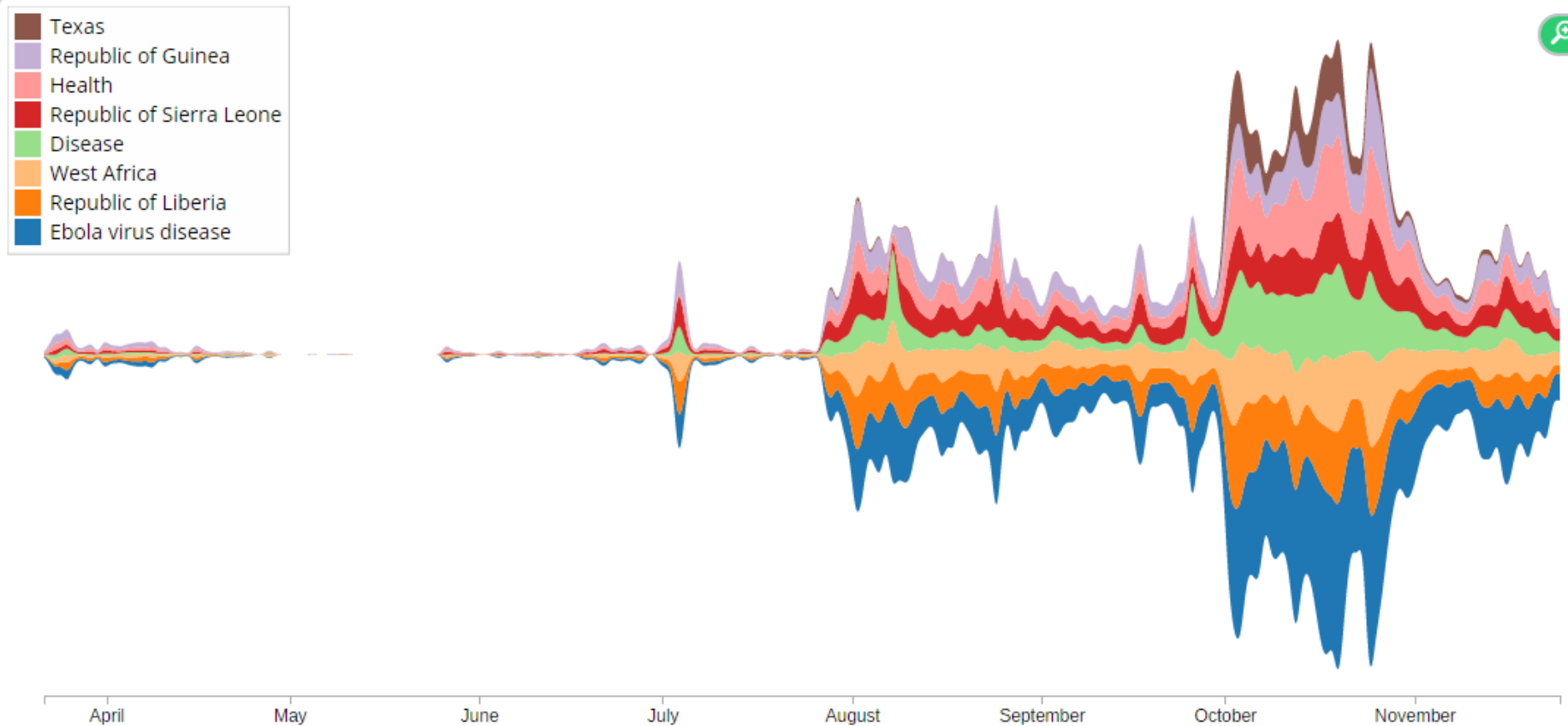
The visualization shows the trending of individual concepts in the events. You can choose the types of concepts to visualize or enter your own.

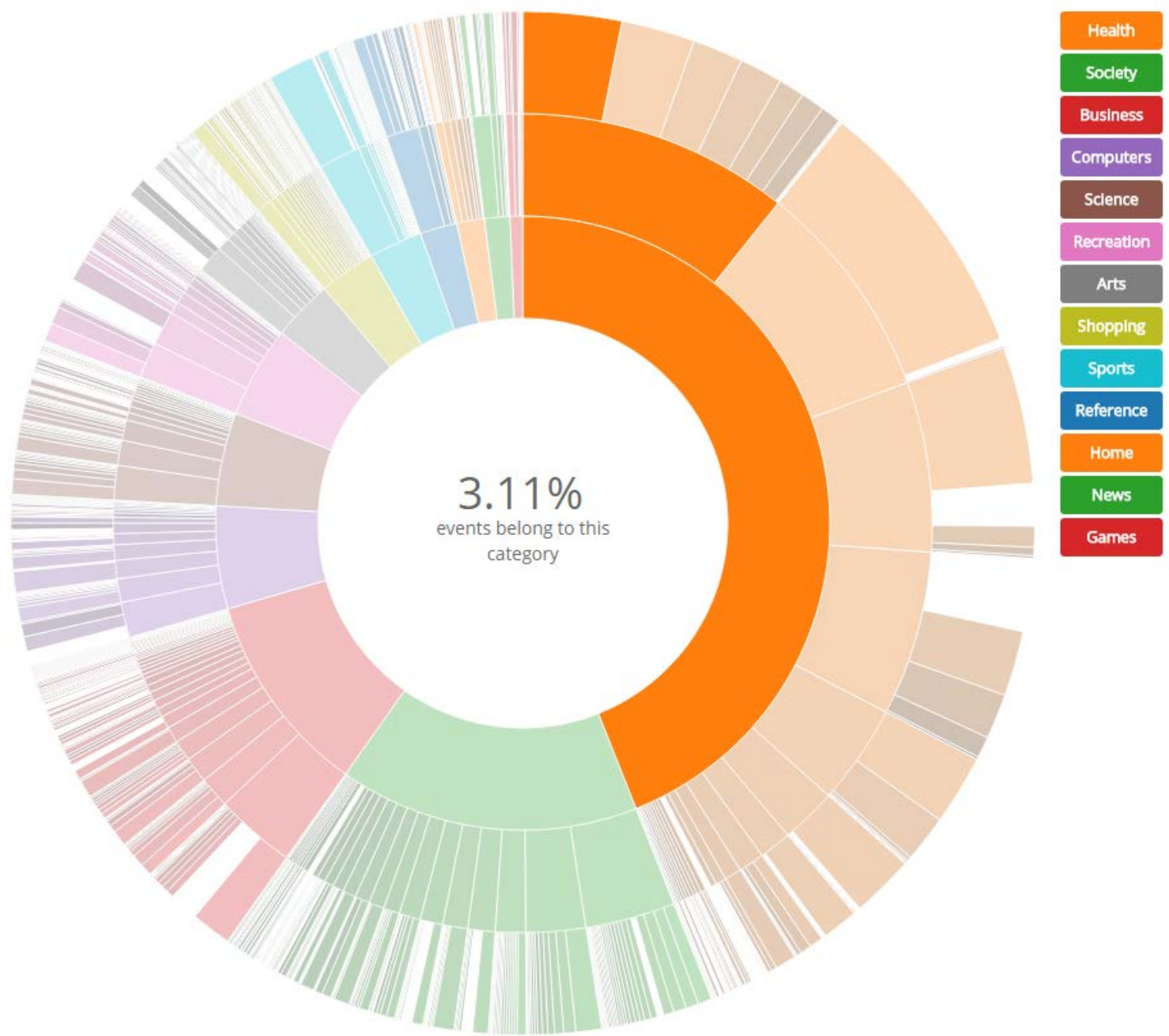
Concept types to display & normalization

Person Location Organization Nonentities

Normalize data by number of imported articles for that day ▾

📍 Ebola virus disease 📍 Republic of Liberia 📍 West Africa 📍 Disease 📍 Republic of Sierra Leone 📍 Health 📍 Republic of Guinea 📍 Texas





Malaysian plane shot down over Ukraine

17 Jul 2014

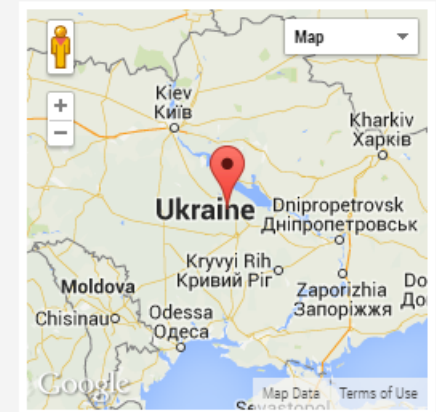
Ukraine

Society→Issues→Warfare and Conflict , Society→Issues→Transportation

KIEV, Ukraine (AP) -- A Ukrainian official said a passenger plane carrying 295 people was shot down Thursday over a town in the east of the country, and Malaysia Airlines tweeted that it lost contact with one of its flights over Ukrainian airspace.

Anton Gerashenko, an adviser to Ukraine's interior minister, said on his Facebook page the plane was flying at an altitude of 10,000 meters (33,000 feet). He also said it was hit by a missile fired from a Buk launcher, which can fire missiles up to an altitude of 22,000 meters (72,000 feet).

The fate of the passengers wasn't immediately...



Content to display:

Articles

Articles

Below is a list of articles describing the event.

eng 773 spa 114 ger 170 chi 30 slo 0

First < 1 2 3 4 5 6 7 8 > Last

Official: Malaysian plane shot down over Ukraine

KIEV, Ukraine (AP) -- A Ukrainian official said a passenger plane carrying 295 people was shot down Thursday over a town in the east of the country, and Malaysian Airlines tweeted that it lost contact with one of its flights over Ukrainian airspace.

Anton Gerashenko, an adviser to Ukraine's ...

GREENWICH TIME

17 Jul, 16:22

0.9952

Official: Malaysian plane shot down over Ukraine - WTOP.com

KIEV, Ukraine (AP) -- A Ukrainian official said a passenger plane carrying 295 people was shot down Thursday over a town in the east of the country, and Malaysian Airlines tweeted that it lost contact with one of its flights over Ukrainian airspace.

Anton Gerashenko, an adviser to Ukraine's ...

WWW.WTOP.COM

17 Jul, 17:04

0.9932

Malaysian Airlines plane shot down over Ukraine; 295 feared dead

NORTHWEST FLORIDA

DAILY NEWS

Event concepts

Top concepts describing what the event is about.

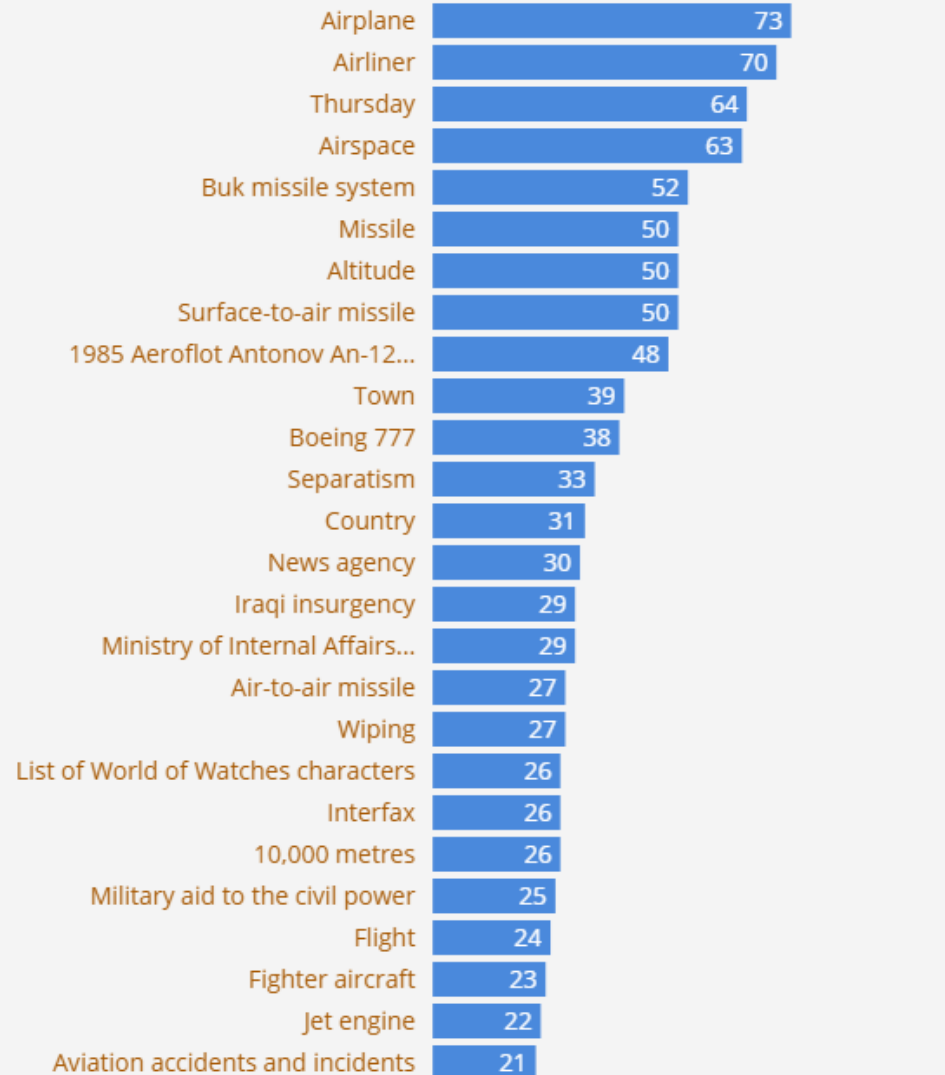
ENTITIES

Concept relevance 0 10 20 30 40 50 60 70 80 90 100



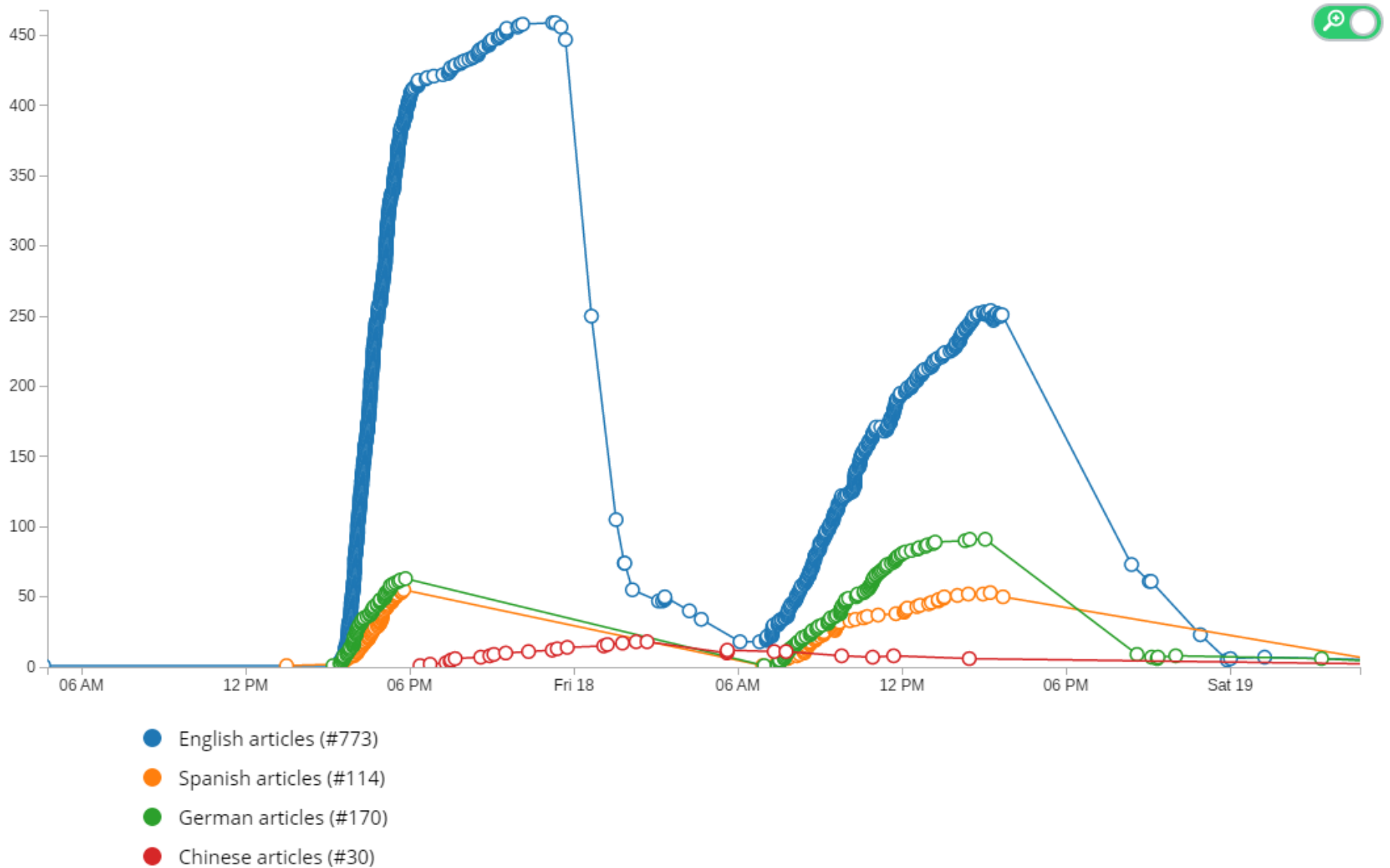
KEYWORDS

Concept relevance 0 10 20 30 40 50 60 70 80 90 100



Timeline of articles

Publishing times of articles describing the event. The height of the curve indicates the number of articles published in last 8 hours.



Sources of the news articles

Map shows the geographic locations of the news sources that published articles about the current event.

Below the map is the time distribution of articles. Click and drag to select a specific time period and highlight only the corresponding subset of sources on the map.



Event storyline



September 11, 2011

Obama to unveil NSA reforms, response to Snowden

WASHINGTON: President Barack Obama presents Americans on Friday with long-awaited reforms of spy agency phone and Internet data collection sweeps, prompted by the damaging torrent of leaks unleashed by Edward Snowden. Caught between civil liberties campaigners and a resistant intelligence...



SEPTEMBER 11, 2001

Presidential panel recommends limits on NSA spying



DECEMBER 20, 2013

Obama Can't Avoid the NSA Report, But He'll Try

- Obama announces changes to data collection
- Obama on NSA Spying Program: Someone Else, Not
- Obama fuels reform on some but not all NSA spying

Obama to unveil NSA reforms, response to Snowden



WWW.ANSAMED.INFO

Syria: HRW, evidence of regime torturing prisoners in Raqqa

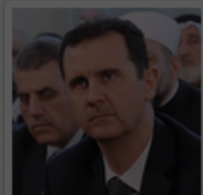
2013-05-17 15:34

(ANSAmEd) - BEIRUT, MAY 17 - Cells and torture chambers where detainees were arbitrarily held and tortured have been found in Raqqa, northern Syria's main city which is under the control of local armed opposition groups, Human Rights Watch (HRW) said. Researchers with the US NGO travelled to Raqqa in April and visited the state security and military intelligence facilities where detainees were reportedly tortured. The HRW team found documents and instruments which perfectly matched torture reports by Syrian political prisoners.

'The documents, prison cells, interrogation rooms and torture devices we saw in the government's security facilities are consistent with the torture former detainees have described to us since the beginning of the uprising in Syria', said Nadim Houry, the Beirut-based deputy Middle East director of Human Rights Watch.

HRW said its researchers found in the state security facility, among other things, a 'bsaat al-reeh' torture device, which former inmates have said has been used to immobilize them and severely stretch or bend their limbs. (ANSAmEd)

- Entities:** Syrian Arab Republic United States Beirut Prison Report Document Research
 Musical instrument Local government Middle East Rebellion Guerrilla warfare Bicameralism
 Hierarchy Non-governmental organization Human Rights Watch Ar-Raqqah
- Keywords:** City Retirement Security Interrogation Sovereign state
 Municipalities with language facilities Room Social group Beginning Zelle (Biologie) Torture
 Arbitrariness Impedance matching Consistency Limb (anatomy) Political prisoner
 Sa'id of Egypt Detainee Military intelligence Ansa Mediterranean Nadim Re'eh



Articles



0 slv [Icons]

May. 17, 15:34 WWW.FOXNEWS.COM

May. WWW.ANSAMED.INFO

May. 17, 18:28 WWW.YNETNEWS.COM

May. 17, 15:34