

# Piattaforme bioinformatiche di sorveglianza: l'esempio di MedISys

Centro comune di ricerca (CCR)

Commissione Europea

Text and Data Mining Unit

Jens Linge, Marco Verile, Eleonora Mantica, Erik van der Goot



# We all do media monitoring

Favourite newspapers (browser, app, rss, print)

Google news alerts

https://www.google.com/alerts

Email alerts from expert sites (e.g. ProMED-mail)

http://www.promedmail.org

Twitter (App, Tweetdeck)

https://tweetdeck.twitter.com

Browsing blogs and dedicated websites, reddit

Scientific literature



# **Epidemic intelligence**

**Event-based surveillance:** 

Soft data (news reports, news wires, social media)

Indicator-based surveillance:

Hard data (indicators such as mortality rates, notifications from health care providers)



### **MEDISYS**

Reads approx. 300k news items & 55M tweets per day

### Analyze:

categorize, filter, geolocate, statistics entity and event metadata extraction

### Present results:

website, email alerts, rss, moderation tools

http://medisys.newsbrief.eu



#### advanced search

Wednesday, September 14, 2016

2

3:29:00 PM CEST

subscribe | manage

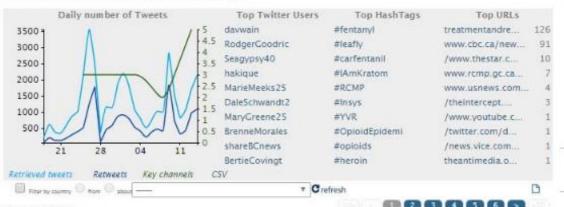
info I definition

current page | all pages

RSS I MAP



#### Latest News About - Fentanyl



Latest articles

RCMP officers tell their story of being exposed to fentanyl 
| story | globalnews Wednesday, September 14, 2016 3:29:00 PM CEST | info (other)

By Online News Producer Global News Amy Judd ABOVE: Two RCMP officers from Kamloops are speaking out and sharing their story of accidental exposure to fentanyl. Catherine Urguhart reports. Some B.C. RCMP officers are speaking out about the dangers of fentanyl exposure after nearly overdosing on the Job....

Articles published more than 10 minutes ago

Eleven SWAT officers treated for exposure to fentanyl and heroin in drug raid 🗹

WashingtonPost Wednesday, September 14, 2016 3:10:00 PM CEST | info (other)

HPD busts major Heroin op. 50,000 bags of heroin, 2 hand guns, 350gr raw heroin. Fentanyl, Packaging Op. 3 arrested. — D/C Foley (@LtFoley) September 14, 2016 Eleven police officers in Hartford, Conn., were taken to a hospital for possible exposure to heroin and fentanyl after drug powder blew......

Articles published more than 20 minutes ago

The teacher thought it was candy: New Jersey boy, 5, came to school with 30 packets of heroin in his lunchbox [6]

dailymail Wednesday, September 14, 2016 3:09:00 PM CEST I info Jother!

The boy, 5, is enrolled in kindergarten at the International Academy of Trenton Charter School; When his teacher saw him playing with what looked like candy, the boy was told to put it away; Moments later, the boy took out the packet again, prompting the teacher to give a closer look; She was......

Articles published more than 1 hour ago

Accused in Kelowna fentanyl bust was former Hells Angels 'middleman' 

| CSC Wednesday, September 14, 2016 2:19:00 PM CEST | Info [other]

A Kelowna man charged in connection with a massive fentanyl bust was a former 'middleman' for the Helis Angels, according to parole documents. Leslie John McCulloch was charged with possession for the purpose of trafficking this summer following a raid in which RCMP seized two industrial pill....... Languages 🖸 🖸

Select your languages

U5

Most reported countries (24h)





### **EMCDDA News**

14 Sep 2016 - Edition No. 1

### INDEX

Trend topics	1
Potential Serious and Urgent Events	5
Cannabinoid	8
Fentanyl	10
Crystal Meth	12
Diazepam	13
Cocaine seizure	14



## Annex I A I of the Council Directive 2000/29/EC Insects, mites and nematodes

#### Spodoptera frugiperda

#### War on armyworms as ministry declares 'agric state of emergency'

[fr] [de] 08.05.2017

Parliament has been asked to declare an 'agricultural state of emergency' as an armyworm invasion puts a GHC560m agricultural project at 'serious' risk. Describing the invasion as requiring urgent response, the Agriculture Minister Dr Owusu Afriyie Akoto wants to rope in parliament to wage a war... (more)

#### Plan d'action R&D contre le ravageur du maïs

Science and Development Network - Afrique Sub-Saharienne

[en] [de] **08.05.2017** 

#### Researchers call for a strategy against armyworm

Un groupe de scientifiques estiment qu'une recherche sur l'interaction entre les modèles de vie des larves de la chenille légionnaire d'automne et le calendrier de pulvérisation d'insecticide est urgemment nécessaire pour lutter contre ce ravageur qui fait des dégâts dans les champs de maïs en Afrique. (more)

A group of scientists believe that research on the interaction between the life patterns of Fall Armyworm larvae and the insecticide spray schedule is urgently needed to control this damaging pest in maize fields in Africa.



March 2017

### **Monthly Summary of Articles** on Food Fraud and Adulteration

Retrieved mainly from the JRC tool Medisys (http://medisys.newsbrief.eu/)



#### Selected Food Fraud Cases

#### 01/03/17 - 35 000 eggs seized near Parma

About 35 000 eggs were seized in the region of Parma by the anti-fraud unit of the local Carabinieri. The declaration of origin and other labelling information were either missing or incomplete. Ansa

#### 03/03/17 - 5000 bottles of fake Barolo and other quality wines seized near Asti

Fraud inspectors seized 5000 bottles of prestigious wines (Italian DOC and DOCG denominations), including Barolo, in a wine cellar near Govone. They also found false labels and thousands of fake marks, worth tens of thousands of euro. Tests are on-going to determine the origin of the sold wine. La Stampa

#### 05/03/17 - Results of the operation "Sardina Pilchardus"

The operation Sardina Pilchardus took place in February in Calabria to control the local fish market. Notably, 116 000 euro in fines were imposed and 1254 kg of fish seized for lack of or incomplete traceability labelling.

Quotidiano del Sud



Italy





Origin masking



Italy



Wine



Counterfeit



Italy





Mislabelling



# What are we looking at?

### **Human communicable diseases => ECDC, WHO**

**Syndromes** 

CBRN threats => G7

Mass gatherings (EXPO 2015, Olympics, FIFA, UEFA, etc.)

Food & Feed contaminations including GMOs => EFSA

Animal health => EFSA

Plant health => EFSA

**Medical Devices** 

Psychoactive Substances / Drugs => EMCDDA



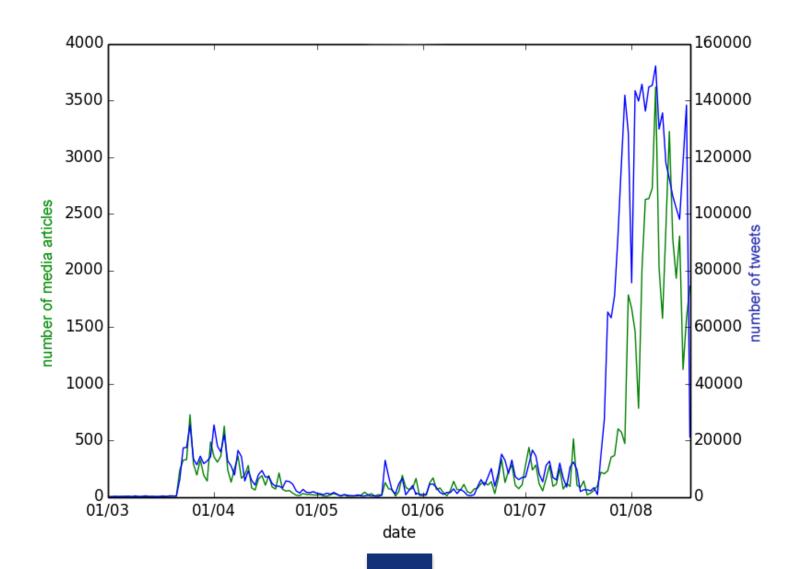
# **Example - Ebola in West-Africa**

Events in March 2014:

14/3 Une étrange fièvre fait 8 morts à Macenta ⇒ triggered UnknownDisease

18/3 Mystery hemorrhagic fever kills 23 in Guinea ⇒ triggered EbolaHemorrhagicFever



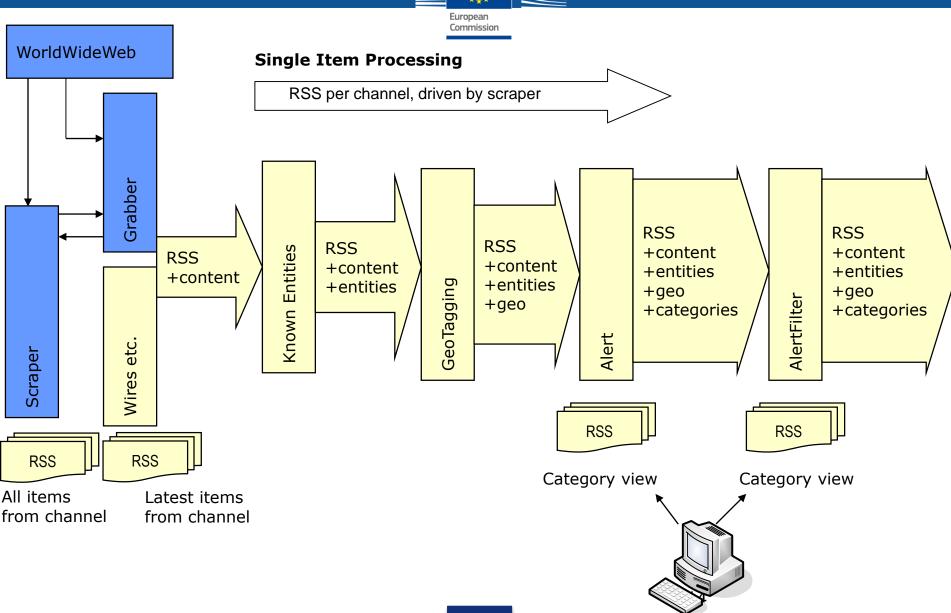




# **Technical background**

Processing chain Multi-lingual approach







# Multilingual keyword patterns

```
Patterns for <u>unknown disease</u> comprise: 
mysterious+disease
unidentified+disease
неизвестн%+инфекц%
новый+смертельный+вирус
непозната+болест
未知的疾病
யூவிப்+குவியிய்
etc.
```

Boolean operators (AND, OR, NOT) and proximity



# How can you profit?

Email alerts, e.g. Coronavirus RSS, e.g. Measles Website can be further customized

Newsletters with human moderation => Newsdesk training

Create categories using your own keywords => CategoryEditor training



**Joint Research Centre (JRC)** 

www.jrc.ec.europa.eu

Jens Linge Text and Data Mining Unit

http://medisys.newsbrief.eu http://emm.newsbrief.eu

Serving society
Stimulating innovation
Supporting legislation



Joint Research Centre

Supporting

legislation



### **Customers**

**European Commission** 

DG COMM, DG HOME, DG GROW,

DG SANTE, DG RESEARCH, etc.

European External Action Service

**European Council** 

European Parliament

EU agencies: ECDC, EFSA, EMCDDA, FRONTEX

UN, WHO

EU member states (ISS, INVS, RKI)

G7 countries





Italian - German





English - French





Spanish - Portuguese





# **Twitter analysis**

Twitter data are noisy, high-volume and mostly (90%) not geo-located

### BUT:

Twitter data can be used to give additional information on events extracted from traditional media:

- additional sources that we don't cover
- news items which are often cited
- blog and forum entries which are often cited



### **Data retrieval**

Converts dirty html with adverts, menus, html tags, 'related stories', etc. into clean and standardised UTF-8 encoded RSS

RSS, XML feeds

Crawling of sites





# **Example - EHEC outbreak in Germany**

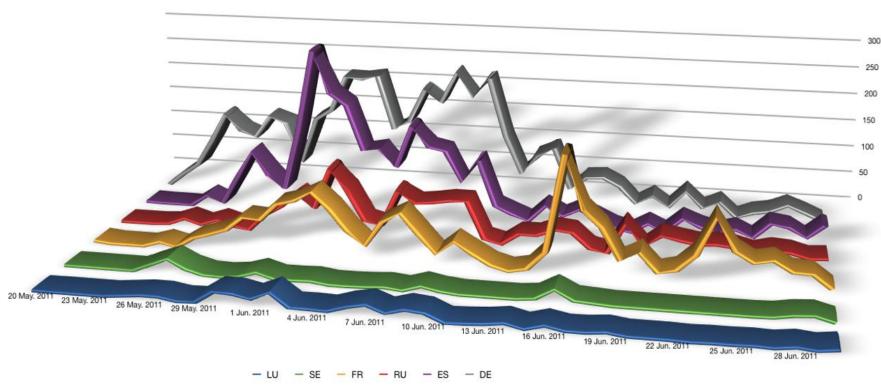
Early alerting via German media (May 2011)

Tracking of a known outbreak with high reporting:

clustering

analysis by country





Number of articles in the *E. coli* category that also mention one of the countries Luxembourg (LU), Sweden (SE), France (FR), Russia (RU), Spain (ES) and Germany (DE)