

Systeme de veille automatique pour la détection de maladies animales émergentes

Mathieu Roche

Cirad – TETIS – Montpellier, France



Web : <http://www.textmining.biz>

Email : mathieu.roche@cirad.fr



Data Science

Volume

Velocity

Variety

3V of Big Data



Variability, Veracity, Value,
Visualisation, Valorization

→ *Epidemiology and Big Data ?*



Information extraction in documents

Animal disease surveillance



Why the need of Epidemic Intelligence?

More than **60% of the initial outbreak reports** come from unofficial informal and **heterogeneous sources**, including sources other than the electronic media, which **require verification** [Arsevska *et al.* ISVEE'2015]



Industries | Wed Jul 23, 2014 9:54am EDT

Related: NON-CYCLICAL CONSUMER GOODS

Poland investigates suspected case of African swine fever in farm pigs

WARSAW, JULY 23

Polish local authorities said on Wednesday that preliminary tests have pointed to a case of African swine fever (ASF) among farm pigs in eastern Poland near the city of Bialystok.

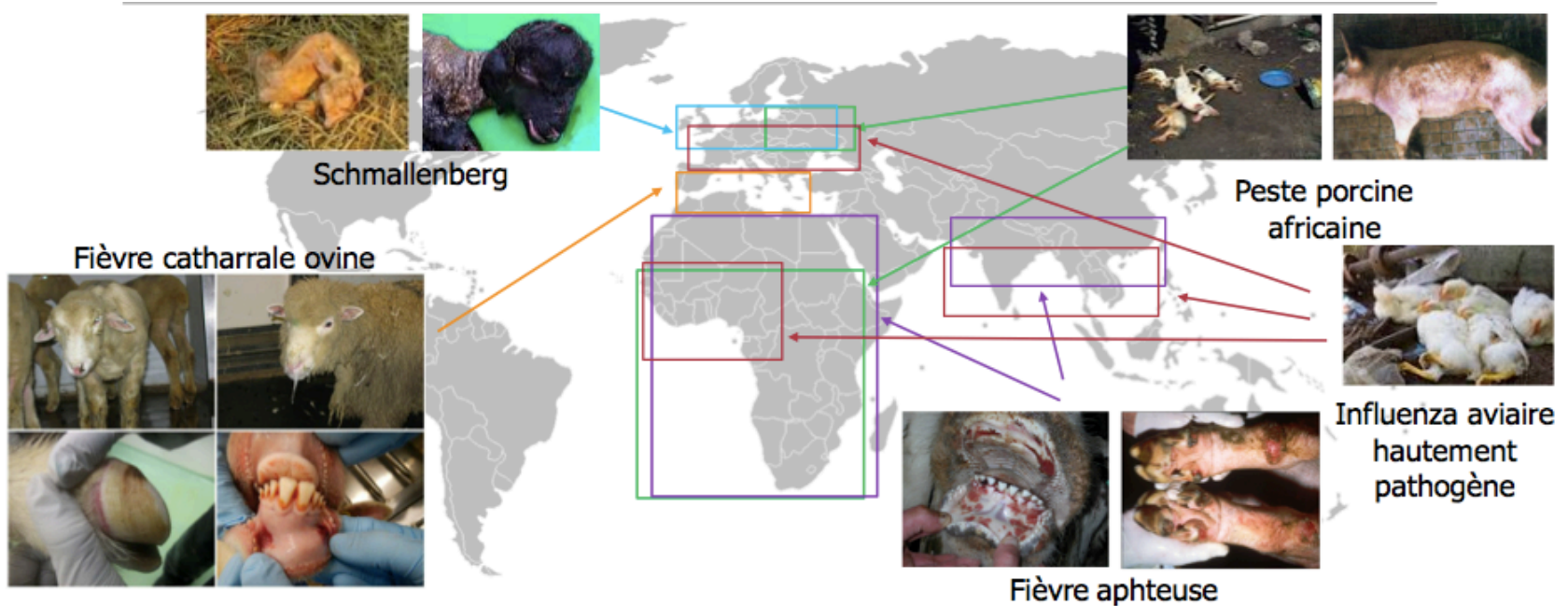
The head of the Grodek county, Wieslaw Kulesza, told Reuters that preliminary results of tests showed that ASF was the cause of death of two-three farm pigs in the county.

"We are marking the area," Kulesza said, adding that further steps, such as laying special mats, were being taken.

Poland's chief veterinary officer was unavailable for comment, while the county veterinary officer said a statement on the issue will be published later on Wednesday. (Reporting by Anna Wlodarczak-Semczuk; Writing by Marcin Goettig)



- **Five animal disease models:** African swine fever (ASF), Foot-and-mouth disease (FMD), Bluetongue (BTV), Schmallenberg virus (SBV), and Influenza



African swine fever cases recorded in Mykolaiv and Odesa regions - Interfax

▲ African swine fever cases recorded in Mykolaiv and Odesa regions - Interfax

▲ African swine fever cases recorded in Mykolaiv and Odesa regions

2016 - # 08 -09 T10 :14 +02:00

10:14

09.08.2016

Cases of African swine fever (ASF) have been recorded in Odesa and Mykolaiv regions, according to a Monday posting on the State Service for Food Safety and Consumers Rights Protection of Ukraine.

According to the chief department of the authority in Odesa region, on August # 5 pigs died in private households of Ukrainians living in Bilyne and Hetmanivka of Odesa region. The examination of the selected samples of biological material diagnosed ASF.

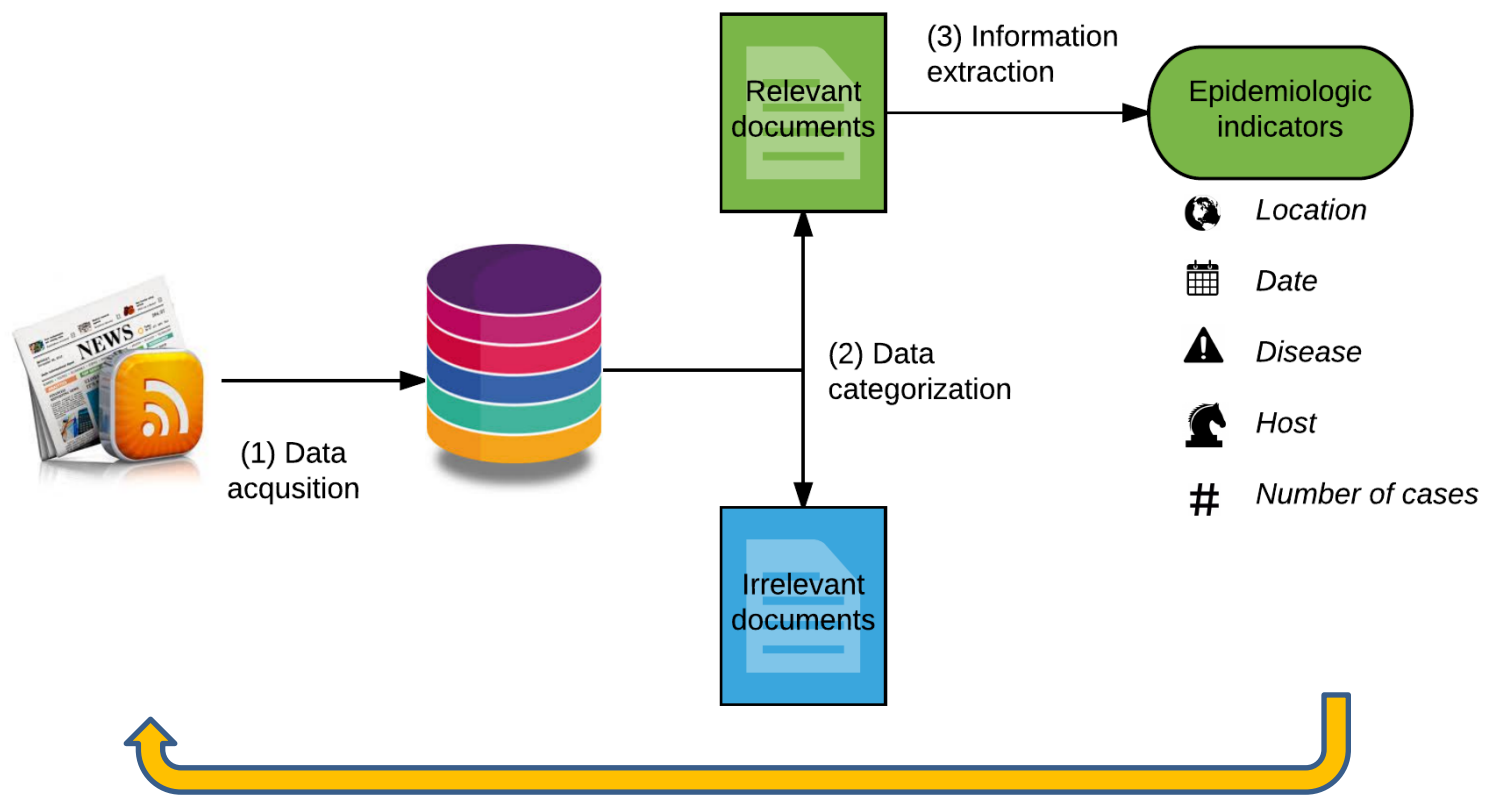
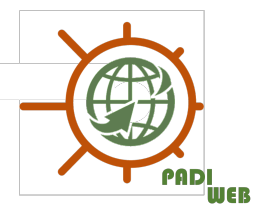
On August # 4, 2016, 11 samples of biological material taken from died pigs found on a landfill in Nova Odesa (Mykolaiv region) proved to have ASF virus DNA.

Measures are being taken to localize the disease on the site of the outbreak and prevent the spread of ASF pathogens.

The ASF virus is the causative agent of ASF. The virus causes a hemorrhagic fever with high mortality rates in pigs, but persistently infects its natural hosts, warhogs, bush pigs and soft ticks of the Ornithodoros genus, without any visible symptoms. There is currently no vaccine against ASF.



article	publication_date	location	latitude	longitude	admin1	country	continent	confidence	user_label
118e93ee4e	13/06/2016	Mamusa	-27.22308	25.27706	North-West	South Africa	Africa	0.72	correct
118e93ee4e	13/06/2016	Free State	-29	26	Orange Free State	South Africa	Africa	0.6	correct
118e93ee4e	13/06/2016	Letsemeng	-29.35811	25.014931	Orange Free State	South Africa	Africa	0.58	correct
148cc73f57	18/03/2016	Kirovohrad	48.5132	32.259701	Kirovohrad	Ukraine	Europe	0.875	correct
3034773164	10/06/2016	Province of North West	-26.5	26	North-West	South Africa	Africa	0.645083	correct
3034773164	10/06/2016	Free State	-29	26	Orange Free State	South Africa	Africa	0.582167	correct
30eace930b	08/03/2016	Ngozi Province	-2.875	29.924999	Ngozi	Burundi	Africa	0.585217	correct
4198a3e4do	26/05/2016	Sumy	50.9216	34.800289	Sumy	Ukraine	Europe	0.99	correct
4198a3e4do	26/05/2016	Sumskaja Oblast	51	34	Sumy	Ukraine	Europe	0.98	correct
4316d582ea	09/06/2016	Chernivtsi	48.29149	25.94034	Chernivtsi	Ukraine	Europe	0.84	correct
6cb2a7dec8	13/06/2016	Free State	-29	26	Orange Free State	South Africa	Africa	0.620091	correct



- **Step 1: Data acquisition - (Social) media data**

→ Web news (*current work*)

Industries | Wed Jul 23, 2014 9:54am EDT Related: NON-CYCLICAL CONSUMER GOODS

Poland investigates suspected case of African swine fever in farm pigs

WARSAW, JULY 23

Polish local authorities said on Wednesday that preliminary tests have pointed to a case of African swine fever (ASF) among farm pigs in eastern Poland near the city of Bialystok.

The head of the Grodek county, Wieslaw Kulesza, told Reuters that preliminary results of tests showed that ASF was the cause of death of two-three farm pigs in the county.

"We are marking the area," Kulesza said, adding that further steps, such as laying special mats, were being taken.

Poland's chief veterinary officer was unavailable for comment, while the county veterinary officer said a statement on the issue will be published later on Wednesday. (Reporting by Anna Wodarczszak-Semczuk; Writing by Marcin Goettig)



→ Blogs, social networks, etc. (*future work*)

First Cases Of African Swine Fever Detected in Czech Republic

By Dave Patterson - June 26, 2017

[Share on Facebook](#) [Tweet on Twitter](#)

The disease, which does not affect humans, was found in two dead wild boars that may have crossed the border into the Zlin region.

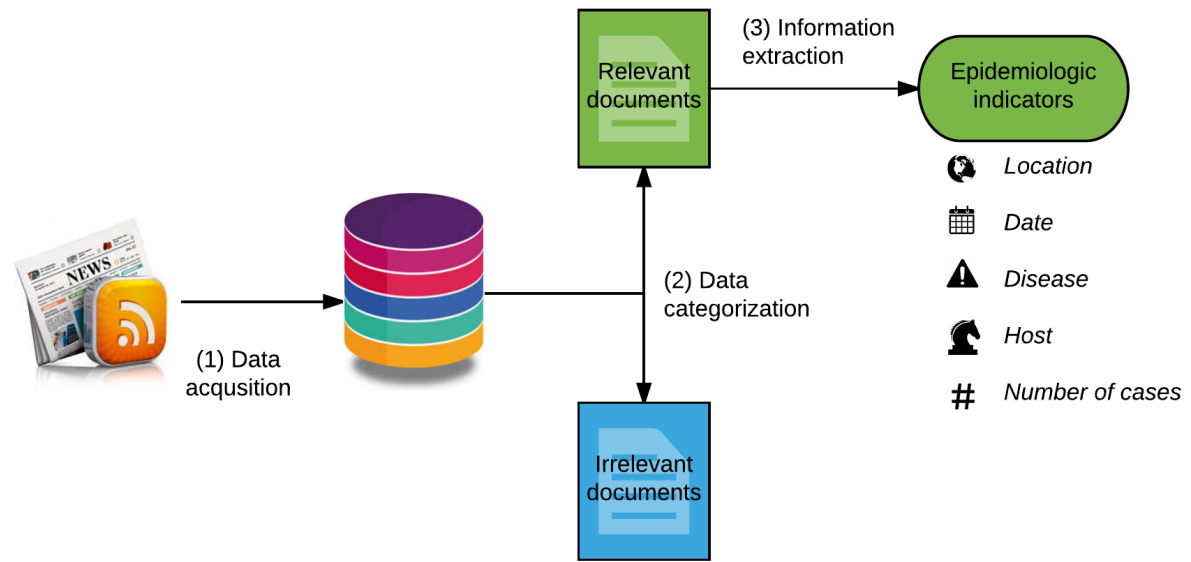
Special monitoring of the wild boar population has been introduced, Agriculture Ministry said on Tuesday.

Ad closed by Google
[Stop seeing this ad](#) [Ad Choices](#)

No pig farm was affected so far.

A 10-km sanitary perimeter has been established including a farm with around 5,000 pigs that are being inspected.

• Step 1: Data acquisition - Process

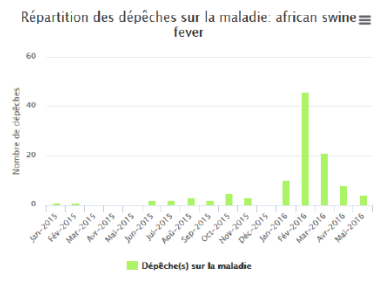


<https://news.google.com/news/feeds?pz=1&cf=all&ned=en&q=Blue+tongue&output=rss>

Statistique

Maladie:

Début période: Fin période:



[Falala et al. IC-Demo'2016]

• Step 1: Data acquisition

Industries | Wed Jul 23, 2014 9:54am EDT

Related: NON-CYCLICAL CONSUMER GOODS

Poland investigates suspected case of African swine fever in farm pigs

WARSAW, JULY 23

Polish local authorities said on Wednesday that preliminary tests have pointed to a case of African swine fever (ASF) among farm pigs in eastern Poland near the city of Bialystok.

The head of the Grodek county, Wieslaw Kulesza, told Reuters that preliminary results of tests showed that ASF was the cause of death of two-three farm pigs in the county.

"We are marking the area," Kulesza said, adding that further steps, such as laying special mats, were being taken.

Poland's chief veterinary officer was unavailable for comment, while the county veterinary officer said a statement on the issue will be published later on Wednesday. (Reporting by Anna Wlodarczyk-Semczuk; Writing by Marcin Goettig)



Visite du ministère de l'élevage et de la pêche dans le cercle de Tominian : La peste porcine africaine, une menace pour le développement économique

Par maliweb.net - il y a 4 mois

252 0

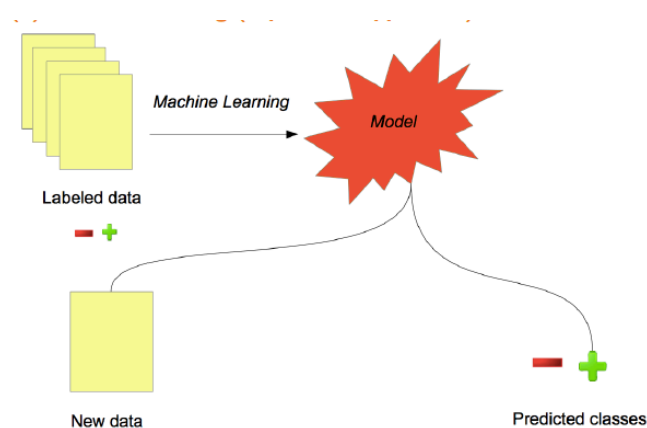
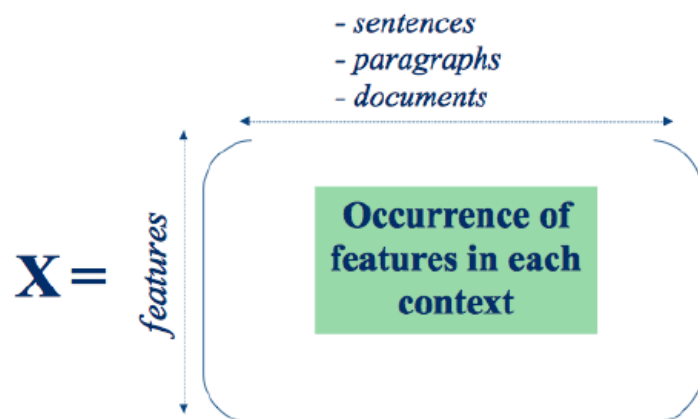
[Portager sur Facebook](#)
[Tweeter sur twitter](#)
[G+](#)
[p](#)
[J'aime](#)
[Tweet](#)



action très résistante

Un communiqué rendu public le jeudi 24 juin 2010, le préfet du département du Diamaré dans la région de l'Extrême-Nord du Cameroun et dont la principale ville est Maroua, a demandé à la population de ne plus manger la viande de porc. Une décision qui fait suite à une épidémie sans précédent de peste porcine africaine (PPA) dans cette partie du pays. Le 02 juin, Nassari Paul Réa avait officiellement

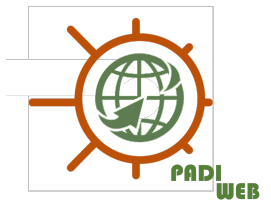
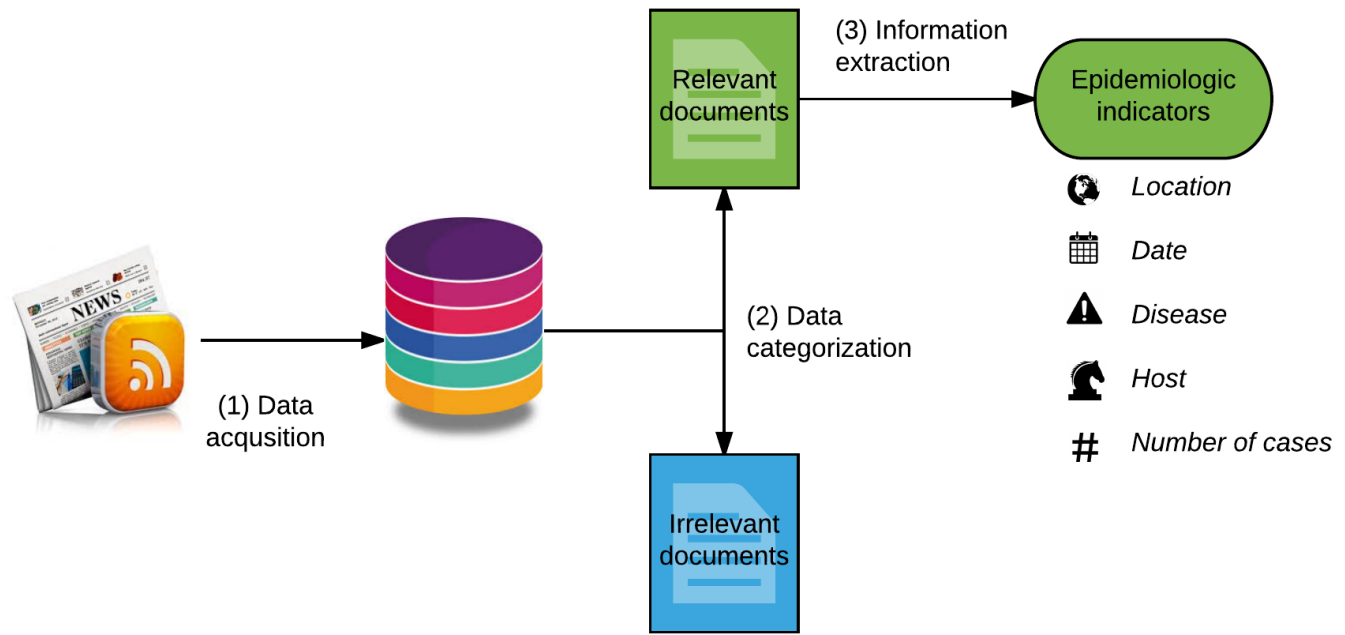
• Step 2: Data classification



Classification algorithm		Naïve Bayes			Support Vector Machine		
Performance		Recall	Precision	F-score	Recall	Precision	F-score
Class	<i>disease</i>	0.724	0.766	0.744	0.657	0.68	0.669
	<i>economy</i>	0.478	0.530	0.503	0.489	0.726	0.584
	<i>general</i>	0.860	0.804	0.831	0.864	0.763	0.810
Weighted average		0.750	0.745	0.747	0.732	0.729	0.725



• Step 3: Information extraction and management



- **Step 3: Information extraction (I)** [Arsevaska *et al.* LTC'2017]

Aim: Automatic detection of **key information** from **Web** news articles (country, species, diseases, number of cases, dates, ...)

“Since its initial appearance in **Poland** in **February 2014**, **72** cases of **African Swine Fever** have been detected in **wild boars** and there have been three outbreaks in **pigs**.” - <http://www.thenews.pl>

- Use **dictionaries** (Geonames, HeidelTime, disease names, species names, etc.), and data mining techniques in order to learn rules

Rules associated with ‘case numbers’:

(number)(species_name,1-3) with frequency **26%** and confidence **83%**

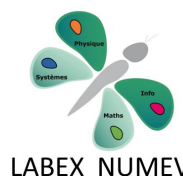
(number)(species_name,1-2) with frequency **21%** and confidence **100%**



• Step 3: Information extraction (I)

- Results for the rule-based approach on the annotated corpus
- Classification based on SVM (*features are rules + specific features for disambiguation of locations*)
- **2 classes:** correct, incorrect
- 10-fold cross validation

Type	Accuracy
Locations	0.80
Dates	0.88
Diseases	0.96
Cases	0.85
Species	0.96



LABEX NUMEV

Julien Rabatel,
LIRMM, Numev, France



• Step 3: Information extraction (I)

Veille Sanitaire Internationale

Projet VSI Accueil Consultation Paramétrage

Annotation

Vous êtes connecté en tant que **Elena** [Déconnexion](#)

Jeux de données

annotations

532 article(s)
< Article précédent
Article suivant >

1. Rivers govt. eliminates chickens infected by flu 1 / 532

✔ Tout a été annoté
 → Nouveau cas
Bilan
Non-pertinent

Date: 2015-01-20

The 🇳🇮 Rivers State Government said on 📅 Tuesday that it had killed # hundreds of 🐔 fowls infected by the 🦠 Avian Flu in a privately owned farm in 📍 **Port Harcourt** .
 The Commissioner for Agriculture , Emma 🇳🇮 Chinda , said that the farm had been quarantined and decontaminated .
 He also said no human infection had been recorded .
 `` On 📅 January # 14 , we got a report from a farm that was worrisome . The report we got suggested that the farm may have been infected by the 🦠 highly pathogenic 🦠 avian influenza .
 According to the commissioner , samples of the flu were taken to the Veterinary Research Institute in Vom , 🇳🇮 Plateau State .
 `` The result came out on 📅 January # 17 and it read positive of 🦠 highly pathogenic 🦠 avian influenza .
 `` On the basis of that , we had to take necessary steps . Apart from quarantining the farm , we had to depopulate the 🐔 birds in the farm to stop further spread `` .
 `` Thereafter , we decontaminated the farm . We are containing the situation because officials of government and experts are on ground monitoring the situation `` , he added .
 Mr. 🇳🇮 Chinda said there was no need to panic because government was well equipped to handle the situation .
 He said before the outbreak , they received information from the 🇳🇮 Federal Ministry of Agriculture on 🦠 avian influenza in 🇳🇮 Kano and a 🐔 bird market in 🇳🇮 Lagos .
 `` We were very much on alert and when it happened here , we handled the situation `` , he said .
 (NAN)

29 candidat(s)


< Candidat précédent
Candidat suivant >

6 / 29

LOCATION

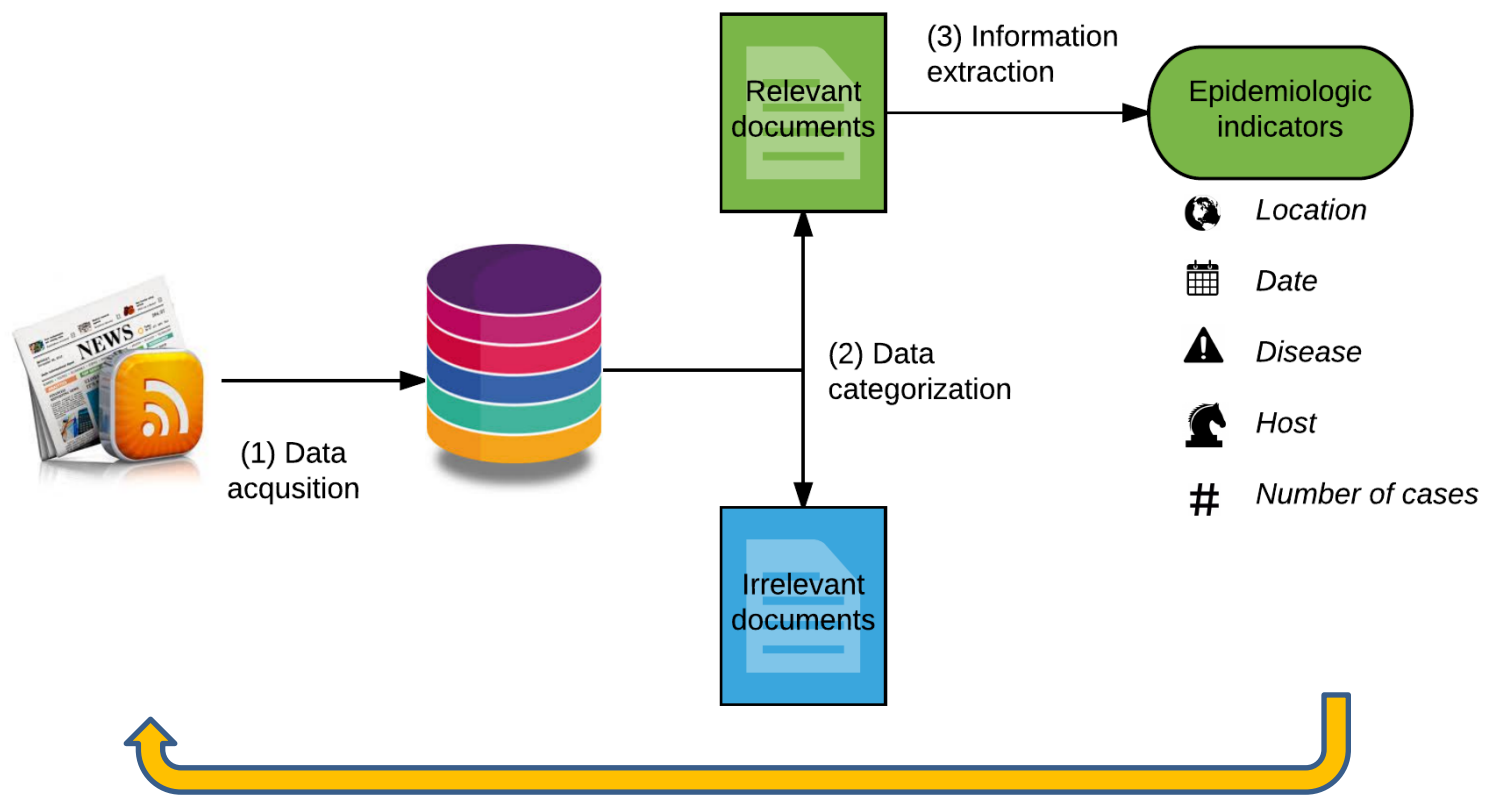
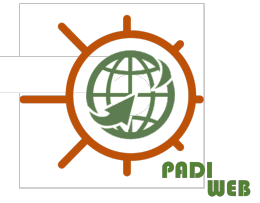
📍 Port Harcourt 📍

✔ Correct
⦿ Partiel
✘ Incorrect
Non annoté



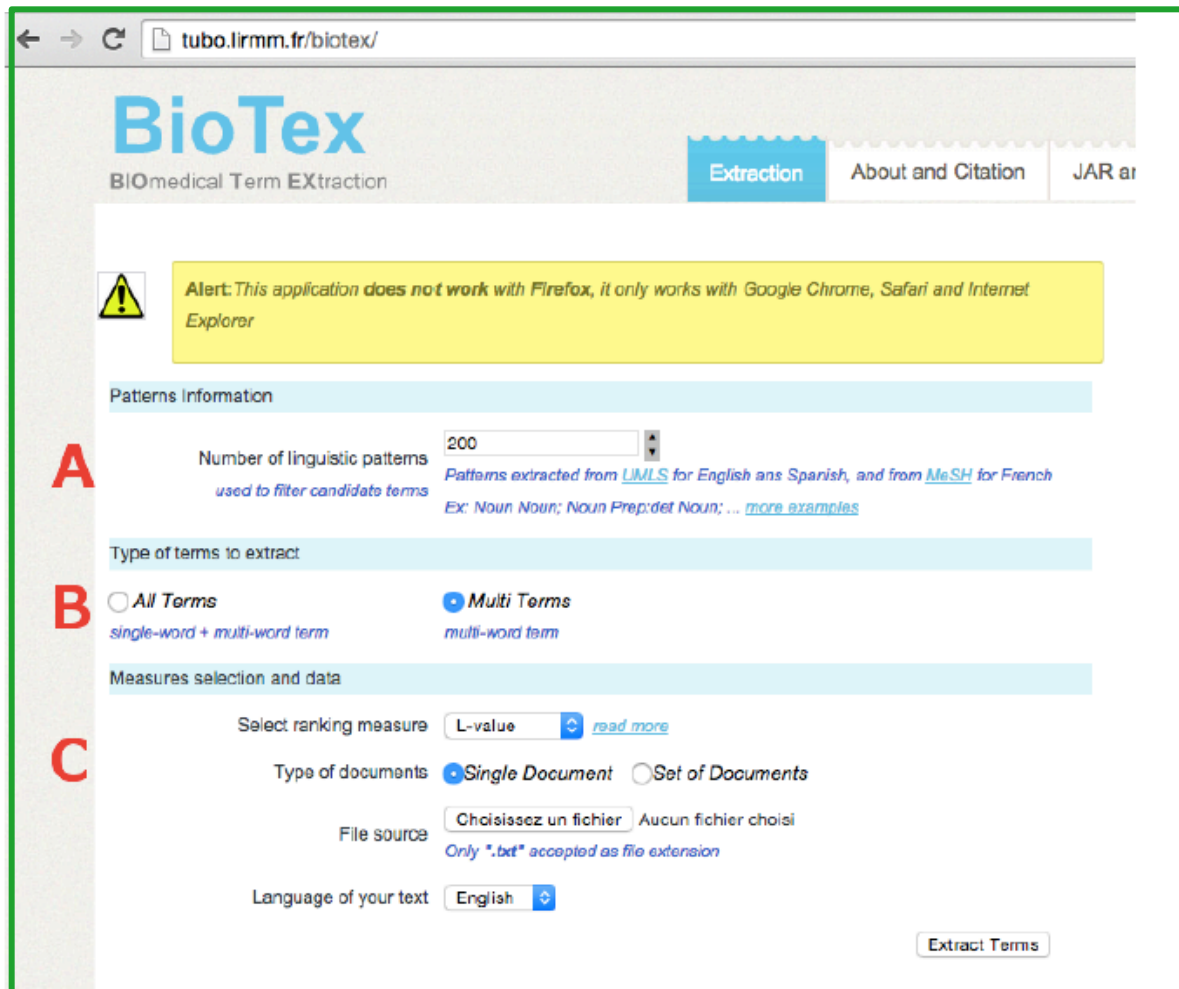
< Article précédent
Article suivant >

Max Devaud-Thomas Filol Copyright 2015 All Right Reserved.



- Step 3: Querying the Web - (a) **Terminology extraction** [Lossio Ventura *et al.* ISWC-Demo'2014]

SIFR
project

- **Step 3: Querying the Web - (b) *Terminology ranking***

- BioTex Ranking [Lossio Ventura *et al.* IRJ'2016]:

$$LIDF\text{-value}(t) = P(t_{dom}) \times IDF(t) \times C\text{-value}(t)$$

- A new ranking function to take into account the **heterogeneity** of the sources (S_i) [Arsevska *et al.* CEA'2016]:

$$w(t) = \sum \alpha_i \times \frac{1}{rank_{S_i}(t)}$$

$$\text{with } \alpha_i \in [0,1] \text{ and } \sum \alpha_i = 1$$



- **Step 3: Querying the Web - (c) *Terminology validation***

List of Extracted Terms

See the terms:

Number of terms : 1200 Page 1 of 10

1	2	3	4	5	6	7	8	9	10
6	lymph node	Validated by : UMLS	<input checked="" type="checkbox"/> YES	<input type="checkbox"/> NO	<input type="checkbox"/> DON'T KNOW				
7	cancer screening	Validated by : UMLS	<input checked="" type="checkbox"/> YES	<input type="checkbox"/> NO	<input type="checkbox"/> DON'T KNOW				
8	radiation therapy	Validated by : UMLS	<input checked="" type="checkbox"/> YES	<input type="checkbox"/> NO	<input type="checkbox"/> DON'T KNOW				
9	cancer patients		<input checked="" type="checkbox"/> YES	<input type="checkbox"/> NO	<input type="checkbox"/> DON'T KNOW				
10	endocrine tumors		<input checked="" type="checkbox"/> YES	<input type="checkbox"/> NO	<input type="checkbox"/> DON'T KNOW				
11	cutaneous t-cell lymphomas	Validated by : UMLS	<input checked="" type="checkbox"/> YES	<input type="checkbox"/> NO	<input type="checkbox"/> DON'T KNOW				
12	renal cell carcinoma	Validated by : UMLS	<input checked="" type="checkbox"/> YES	<input type="checkbox"/> NO	<input type="checkbox"/> DON'T KNOW				

Using of **Delphi method** [Arsevaska *et al.* LREC'2016]

→ *consensus with experts (5 to 7 experts for each disease)*

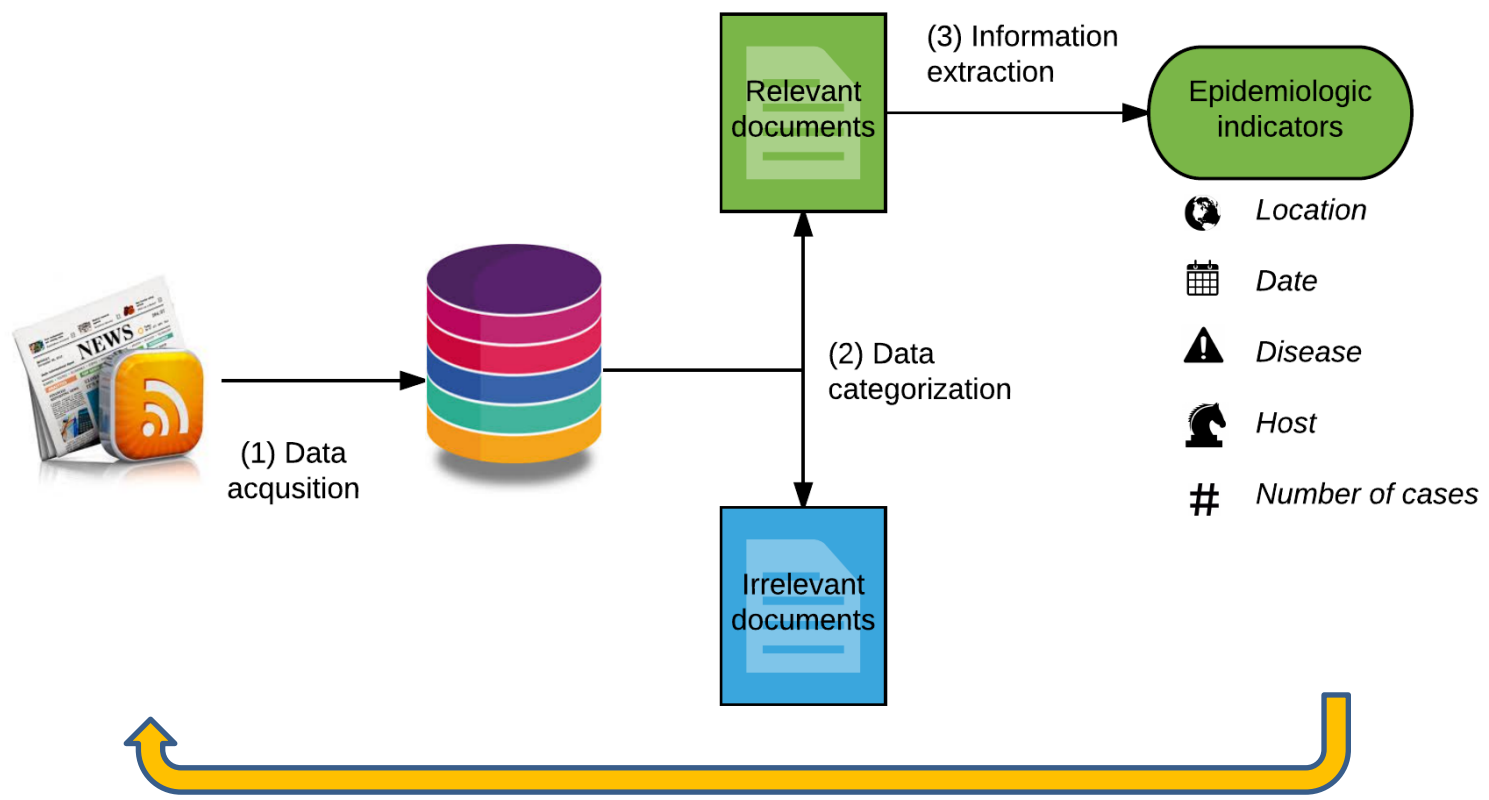
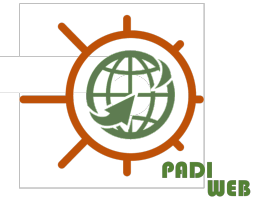


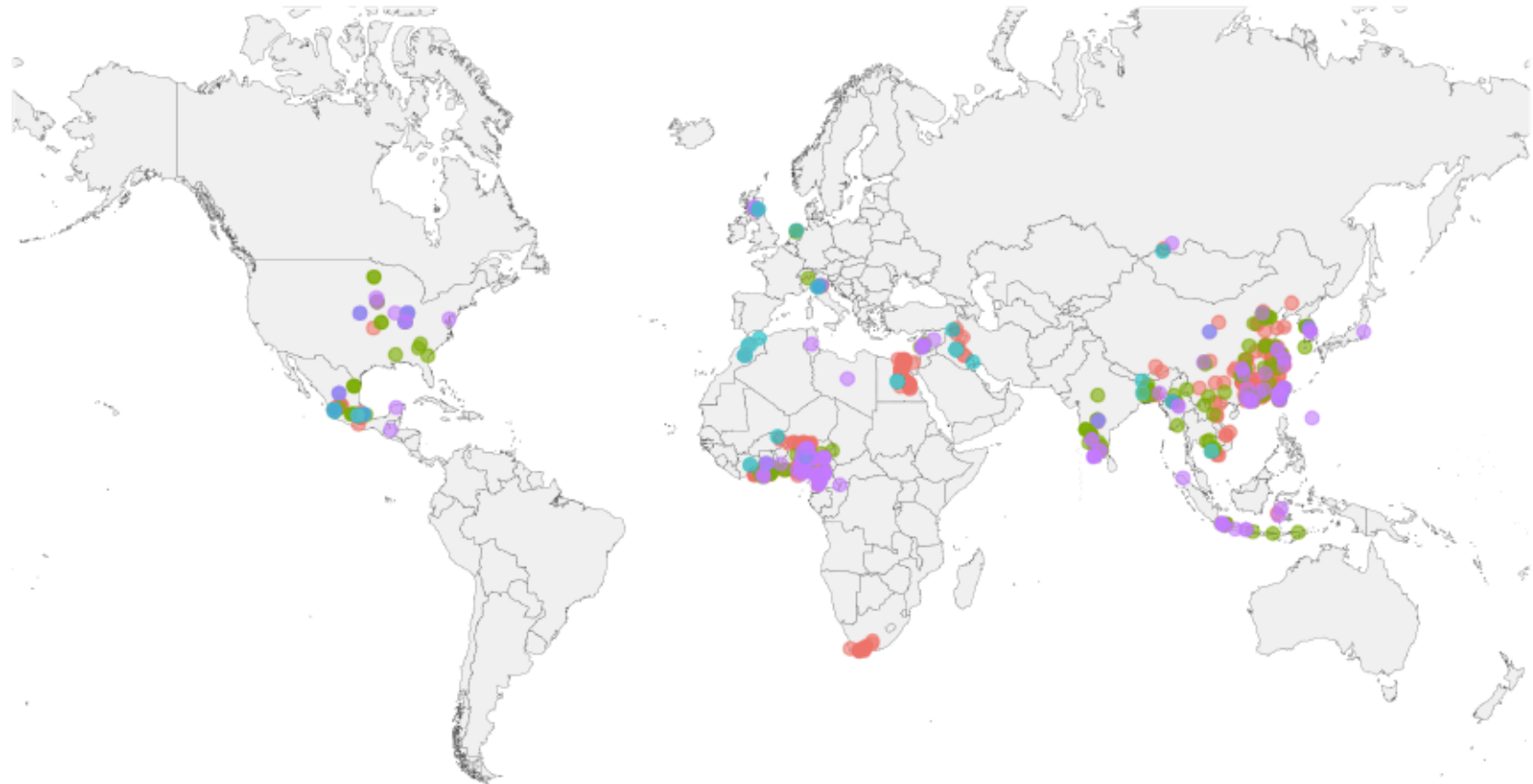
- **Step 3: Querying the Web - (d) *Association of terms***

$$D_{web}^{AND} = \frac{2 \times \text{hit}(h \text{ AND } cs)}{\text{hit}(h) + \text{hit}(cs)}$$

[Roche and Prince Informatica'2010 ; Arsevska *et al.* IJAEIS'2016]

Rank	Bluetongue <i>hosts / clinical signs</i>	Schmallenberg virus infection <i>hosts/ clinical signs</i>
1	general clinical signs / pregnant ewes	stillborn bovine foetuses / camels
2	livestock deaths / sheep	stillborn bovine foetuses / bison
3	embryonic death / cow	aborted foetuses / sheep
4	general clinical signs / sheep	deformed offspring / sheep
5	livestock deaths / cow	stillborn bovine foetuses / deer
6	livestock deaths / deer	aborted foetuses / cattle
7	fever outbreak / sheep	deformed offspring / cattle
8	embryonic death / sheep	stillborn bovine foetuses / calves
9	fever outbreak / cow	deformed offspring / lambs
10	embryonic death / pregnant ewes	acute bronchopneumonia / bison





Source ● Empres-i ● VSI-web ● HealthMap|ProMED ● HealthMap|Autre



Scientific results in Epidemiology:

- **Spatial detection** more efficient than official surveillance (e.g. PPA detected in Ouganda in February and March 2016)
- Detection of **important information** related to diseases not given by official surveillance systems (e.g. preventive vaccinations)
- Detection of information about **other diseases** (e.g. SBV)

Scientific results in Computer Science:

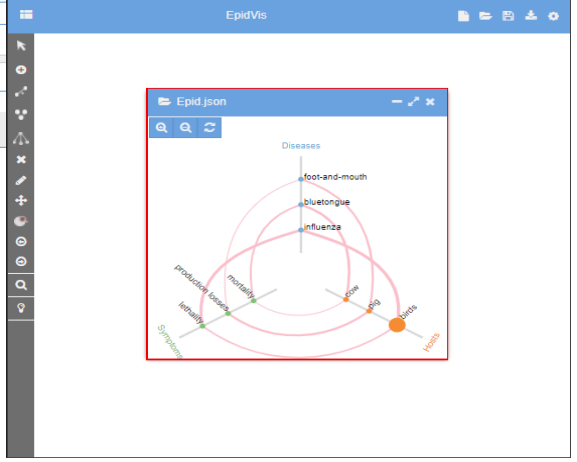
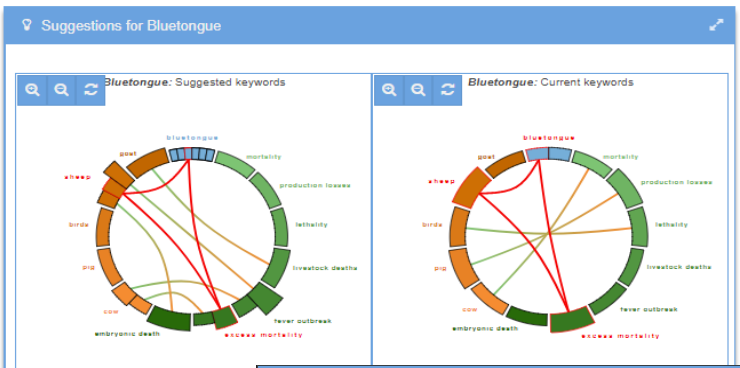
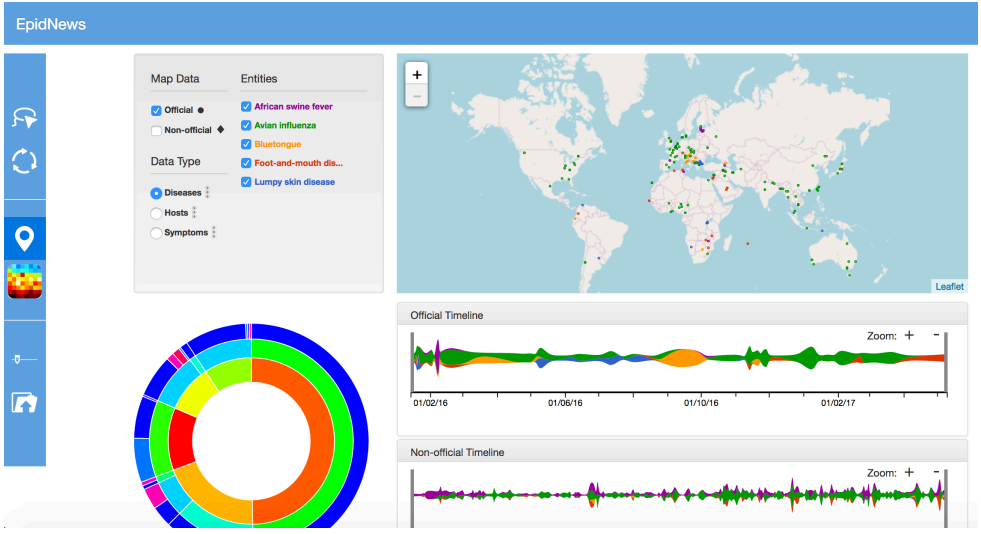
- New **visualisation** systems
- New **ranking measures** to extract keywords from texts
- New **methods to extract information** in textual documents (i.e. new features based on sequential patterns)
- Implemented **system** in collaboration between LIRMM, TETIS, and ASTRE



- **Data Mining vs. Information Retrieval?**



- **Integration of visual analytics skills [Fadloun et al., IV'2017]**





Mathieu Roche



Renaud Lancelot



Sylvain Falala



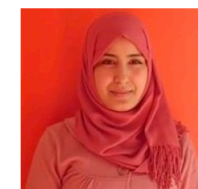
Elena Arsevka



Alizé Mercier



Sarah Valentin



Samiha Fadloun



Jocelyn De Goër De Herve



Julien Rabatel



Pascal Poncelet



Arnaud Sallaberry

Stagiaires : Baptiste Belot, Clément Hemeury, Max Devaud, Thomas Filol, Ibrahim Bompoko Ebengo, Rohan Goel

Autres collaborateurs : Barbara Dufour, Pascal Hendriks, David Chavernac

