

## **Extension de PADI-Web<sup>1</sup>**

### **Extension d'un logiciel de veille sanitaire pour analyser l'émergence et la propagation de maladies animales**

Responsables Inra<sup>2</sup> et Cirad<sup>3</sup> : **Sylvain Falala, Mathieu Roche**  
Encadrants liés au projet : **Alizé Mercier, Jocelyn de Goër de Hervé**

Cirad, Campus de Baillarguet, Montpellier

[sylvain.falala@cirad.fr](mailto:sylvain.falala@cirad.fr), [mathieu.roche@cirad.fr](mailto:mathieu.roche@cirad.fr), [alize.mercier@cirad.fr](mailto:alize.mercier@cirad.fr),  
[jocelyn.degoer@inra.fr](mailto:jocelyn.degoer@inra.fr)

#### **1) Contexte**

La veille en santé animale, et notamment la détection précoce d'émergences au niveau mondial d'agents pathogènes, est l'un des moyens permettant de prévenir l'introduction en France de dangers sanitaires.

Dans le cadre de la thématique "Veille sanitaire internationale" de la Plateforme nationale d'épidémiosurveillance en santé animale (Plateforme ESA)<sup>4</sup>, le Cirad, l'Anses<sup>5</sup> et la DGAI<sup>6</sup> développent depuis 2013 un système de veille automatique du Web qui effectue : (1) le recueil quotidien de dépêches épidémiologiques provenant de sources non officielles, incluant les médias électroniques, (2) l'extraction automatique d'informations (nom de maladie ou symptômes, lieu, date et espèce touchée) issues de ces dépêches et (3) une restitution synthétique et agrégée de l'information : cartes, séries spatiotemporelles.

Actuellement, cinq maladies animales exotiques sont ainsi surveillées, mais d'autres pourraient l'être aisément, car l'outil est développé de façon générique. Ce système sera utilisé par la Plateforme ESA pour la France et par le réseau de vétérinaires CaribVet situé dans les Caraïbes.

#### **2) Approche et technologies utilisées**

Le recueil des dépêches s'appuie sur des requêtes constituées de mots-clés de maladies, d'hôtes et de symptômes pour collecter, avec un script PHP, des articles

---

<sup>1</sup> Platform for automated extraction of disease information

<sup>2</sup> Institut national de la recherche agronomique

<sup>3</sup> Centre de coopération internationale en recherche agronomique pour le développement

<sup>4</sup> <http://www.plateforme-esa.fr/>

<sup>5</sup> Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail

<sup>6</sup> Direction générale de l'alimentation - Ministère de l'Agriculture, de l'Agroalimentaire et de la Forêt

issus de Google News. Ces mots-clés ont été définis par des experts et/ou par des méthodes de fouille de textes (Arsevska *et al.*, 2016).

Chaque article est prétraité et normalisé (suppression de balises HTML, reconnaissance de la langue, etc.) avant d'être stocké dans une base de données MySQL.

L'extraction d'information dans les dépêches collectées identifie les éléments clés (noms de maladies, lieux, dates, nombres et espèces d'animaux touchées). Elle repose sur des dictionnaires dédiés et des règles préalablement construites par un processus de fouille de données. La technologie utilisée est Java.

Une interface Web (développée avec PHP, HTML, CSS, JavaScript et Ajax) permet de paramétrer le processus de recueil, de consulter les articles collectés et de récupérer sous forme de tableaux les informations extraites (cf. Figure 1 et 2).

The screenshot displays a web-based search interface with the following components:

- Maladie (Disease):** A list of diseases including African swine fever, Avian influenza, Bluetongue, Schmallenberg virus infection, Peste porcine africaine, and Influenza aviaire. A "réinitialiser" button is located above the list.
- Symptôme (Symptom):** A list of symptoms including Congenital, Digestive, Fever, General, Haemorrhagic, and Mortality. A "réinitialiser" button is located above the list.
- Hôte (Host):** A list of hosts including Avian, Bovine, Other ruminant, Ovine/caprine, Porcine, and Oiseaux. A "réinitialiser" button is located above the list.
- Source:** A list of news sources including NBC2 News, ournal (blog), The Hindu, Times of India, New York Times, and Interfax. A "réinitialiser" button is located above the list.
- Date début période:** A text input field containing "01-07-2016".
- Date fin période:** A text input field containing "05-10-2016".
- Rechercher:** A dark button to execute the search.
- Recherche par titre:** A text input field for searching by title.
- Recherche contenu:** A text input field for searching by content.
- Résultat Recherche:** A section containing search results, with an "Exporter les résultats" button at the top left. Two results are visible:
  - New cases of foot-and-mouth disease detected in Mauritius - News Ghana**  
New cases of foot-and-mouth disease detected in Mauritius  
Mauritius Agro Industry and Food Security Minister Mahen Seeruttun said importation of ...
  - (Infectious) Foot-and-Mouth Disease In Rodrigues - Island Crisis News**  
(Infectious) Foot-and-Mouth Disease In Rodrigues  
Article by Hibz Y.D. Published on August 3 2016. Categorized under General News[1]Several Cows H...

Figure 1 : interface de recherche multicritères (nom de maladie, symptôme, hôte, source, période...) pour consulter les articles recueillis

New cases of foot-and-mouth disease detected in Mauritius - News Ghana  
 New cases of foot-and-mouth disease detected in Mauritius  
 Mauritius Agro Industry and Food Security Minister Mahen Seeruttun said importation of ...

Options

**Titre :** New cases of foot-and-mouth disease detected in Mauritius - News Ghana  
**Date parution :** 14-08-2016 22:39  
**Mots-clés :** Foot-and-mouth disease - sheep - cattle - goat - lamb - goats  
**Source :** News Ghana

**LOCATION**



**New cases of foot-and-mouth disease detected in Mauritius - News Ghana**

New cases of foot-and-mouth disease detected in Mauritius - News Ghana  
 New cases of foot-and-mouth disease detected in Mauritius  
 Mauritius Agro Industry and Food Security Minister Mahen Seeruttun said importation of animals to Rodrigues could have resulted into the virus reaching this tourist island.  
 Foot-and-mouth disease normally affects cattle, sheep and goats.  
 The minister said the disease was detected by Rodrigues veterinary services on July 7, 2016.  
 A total of 579 animals suspected of suffering from foot-and-mouth disease have been slaughtered in Rodrigues island, while 300 others have been slaughtered in other parts of Mauritius.  
 The exact cause of the transmission of the disease in Rodrigues is not yet known, Seeruttun said, adding that "we suspect the virus was carried to the island through infected meat carried by travellers using yachts."  
 The minister said the results of analyses carried out in France and South Africa will enable authorities to know the source of the disease.  
 "It may take about three years to completely eradicate foot-and-mouth disease from Rodrigues island. But currently there is a temporary ban on importation of beef, pork, goat meat as well as lamb meat on the island," the minister said. Enditem  
 Source : Xinhua/News Ghana [ 1 ]  
 References ^ News Ghana ( www.newsghana.com.gh )

Annotation automatique ▾  
 Classe prédite pour l'annotation : Correcte  
 Confiance : 94.50%  
 Correct 94.50% Inc

Figure 2 : interface de consultation d'une dépêche avec identification automatique des informations clés (lieu, date, maladie, espèce, nombre de cas...)

### 3) Travail à réaliser

Plusieurs tâches de différentes natures sont à effectuer, par ordre de priorité :

#### 1-Gestion des langues (Technologies à utiliser : PHP, Java)

Intégration du français et de l'espagnol au niveau des requêtes et de l'extraction d'information.

#### 2-Extension de l'interface (Framework Bootstrap, PHP, HTML, CSS, JavaScript, Ajax)

Ajout d'outils visuels, notamment une carte mondiale dynamique indiquant les foyers émergents en temps réel.

#### 3-Classification automatique des documents (Java)

Intégration de briques logicielles de méthodes de classification automatique afin d'identifier les documents pertinents à traiter.

#### 4-Optimisation du recueil des documents (PHP)

- Parallélisation du moteur de webscraping.
- Extensions de la collecte via des réseaux sociaux, en particulier Twitter.

#### **4) Cadre et environnement de travail**

Le stage se déroulera au Cirad, sur le campus de Baillarguet, à Montferrier-sur-Lez, dans l'Unité Animal, Santé, Territoires, Risques et Ecosystèmes (ASTRE). Le site est accessible depuis Montpellier par 2 lignes de bus.

Le (la) stagiaire sera encadré(e) par des informaticiens du Cirad et de l'Inra, ainsi que des épidémiologistes du Cirad et de l'Anses.

Une gratification mensuelle sera attribuée au stagiaire. Un restaurant d'entreprise sera à sa disposition.

#### **Référence sur la plateforme de veille PADI**

<http://www.cirad.fr/nos-recherches/resultats-de-recherche/2016/veille-sanitaire-sur-le-web-un-outil-pour-prevenir-la-propagation-des-maladies-animales>

#### **Références**

ARSEVSKA E., ROCHE M., HENDRIKX P., CHAVERNAC D., FALALA S., LANCELOT R. & DUFOUR B. (2016). Identification of terms for detecting early signals of emerging infectious disease outbreaks on the web. *Computers and Electronics in Agriculture*, **123**, 104 – 115.

FALALA S., DE GOER DE HERVE J., ARSEVSKA E., ROCHE M., RABATEL J., CHAVERNAC D., HENDRIKX P., DUFOUR B., LANCELOT R., LEFRANCOIS T. (2016). Système de veille sanitaire pour analyser l'émergence et la propagation de maladies animales. *Atelier IN-OVIVE 4<sup>ème</sup> édition, Conférence IC2016*, 7 juin 2016, Montpellier.