

Acte de conférences



10^{ème} édition du congrès

28 et 29 mai 2018

Towards an automatic animal diseases surveillance system based on textual media analysis

Sarah Valentin, M. Roche, R. Lancelot

ASTRE, Animal, Santé, Territoires, Risques, Ecosystèmes, UMR INRA/CIRAD F-34398 Montpellier, France

TETIS, Territoires, Environnement, Télédétection et Information Spatiale, UMR AgroParisTech/Irstea/Cirad, Montpellier, France

Mots-clés: Veille du web, données non-officielles, santé animale, text mining

La veille en santé animale, et notamment la détection précoce d'émergence au niveau mondial d'agents pathogènes, est l'un des moyens permettant de prévenir ou d'anticiper l'introduction de dangers sanitaires en France. Dans le cadre de la Veille Sanitaire Internationale (VSI) de la plateforme nationale d'épidémiologie en santé animale (ESA), un outil dédié à la veille automatique des dépêches issues des médias électroniques a été développé (PADI-web : Platform for Automated Extraction of Animal Disease Information from the Web). Son objectif est de compléter voire d'anticiper les données produites par les sources de notification de référence en santé animales telles que l'OIE ou la FAO. Le fonctionnement de PADI-web repose sur une méthode de fouille de texte pour la détection, la collecte, la catégorisation et l'extraction de l'information sanitaire à partir des données textuelles non structurées publiées sur le web. Une évaluation de l'outil a été effectuée à partir d'un cas d'étude sur une maladie animale d'intérêt (la fièvre aphteuse) en Afrique. Cette évaluation a mis en évidence des performances quantitatives de détection d'événements variables selon les pays, et une forte plus-value qualitative en terme de complémentarité d'informations. Les verrous méthodologiques identifiés (ambiguïtés géographiques, extraction d'événements pertinents en particulier) font l'objet des développements de cette thèse.