

Stage de M2R, projet IMU : IDENUM Identités des Mondes Urbaines. Auteur : Firas ODEH

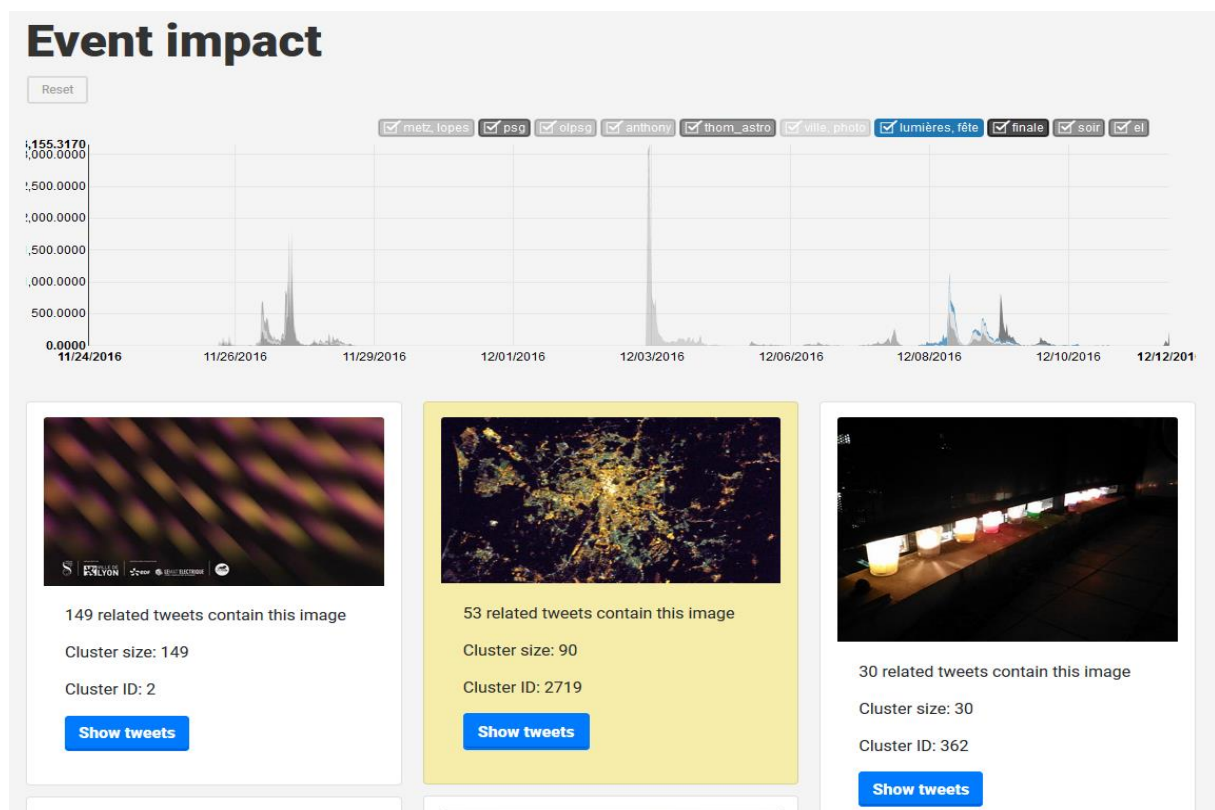
Superviseurs :

Elöd EGYED-ZSIGMOND (LIRIS, elod.egyed-zsigmond@insa-lyon.fr), Adrien GUILLE (ERIC)

Titre : Détection d'évènements dans des flux Twitter

Résumé : Le flux de données Twitter fournit une grande quantité d'informations textuelles et visuelles. De telles données permettent d'identifier des événements significatifs en se basant sur le contenu textuel, les descriptions des utilisateurs, les images jointes et les liens que les utilisateurs utilisent dans les tweets.

Lors de ce stage, nous améliorons un système de détection d'événements basé sur du texte et des mentions pour implémenter les images dans les processus de détection et de classification, après quoi nous développons une nouvelle méthode basée sur un modèle dite « word embedding » afin de générer des mots-clés liés supplémentaires pour les événements pour améliorer la récupération de tweets liés aux évènements.



La figure ci-dessus présente un fragment de l'interface du système en ligne développé qui fournit un moyen assisté de détecter les événements dans le flux de tweets et de les classer en fonction de leur relation avec un événement donné, à l'aide de caractéristiques textuelles et tirées d'images.

Stage financé par le LABEX IMU, Université de Lyon <http://imu.universite-lyon.fr/>

