

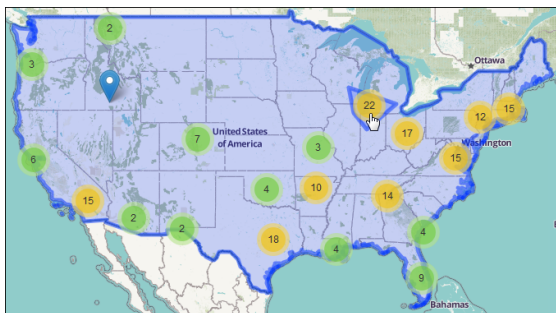
## Corrélation géographique des sujets sur Twitter

### Contexte et domaines de recherche

Les domaines de recherche de ce projet se situent à l'intersection des technologies émergentes du *data mining* (analyse de données), des réseaux sociaux, ainsi que du développement web, en permettant d'analyser avec une application web, GeoTwit, des sujets variés parmi la cascade de données sociales que regroupe et propose la plateforme Twitter.

### But de l'application

GeoTwit permet à l'utilisateur d'entrer un ou deux sujets de son choix et fournit divers outils de filtrage des Tweets, afin de visualiser en temps réel ou de manière statique les activités liées à ces sujets, à l'aide de cartes géographiques, de graphes, et de nombreuses autres données utiles. Il est par exemple possible de comparer la variation de perception des sujets "trump" et "clinton" aux USA.

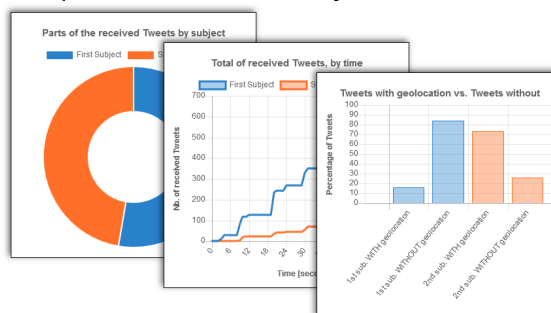


### Des outils de recherche avancés

Afin d'offrir à l'utilisateur un choix vaste et étoffé de possibilités de recherche, les outils de recherche ont été particulièrement soignés. Il est en effet possible de trier les Tweets reçus selon leur contenu à l'aide de filtres "ET" et "OU", permettant d'être plus précis ("*donne-moi les résultats contenant ('scala' OU 'java') ET ('play' ET 'framework')*"). Il est en outre aussi possible de filtrer les résultats par zones géographiques (soit en sélectionnant manuellement un rectangle sur la carte, soit en sélectionnant un pays parmi une liste de pays proposés), par date et par langue.

### Des résultats parlants et interactifs...

En plus de la carte géographique interactive affichant les Tweets reçus sous forme de marqueurs de couleurs (une par sujet), groupés dans des grappes (*clusters*), il est possible de visualiser divers graphes intuitifs et interactifs en temps réel, liés aux données de réception de chacun des sujets mentionnés.



Finalement, il est possible d'accéder à diverses autres informations en temps réel, comme la vitesse de réception des Tweets, le contenu des Tweets reçus ou encore le temps écoulé depuis le début du processus.

### ...mais aussi réutilisables

Une fois le processus d'analyse terminé, il est possible d'exporter les résultats obtenus dans un fichier texte externe, afin de pouvoir les réimporter et donc d'y accéder à tout moment dans GeoTwit, ou de pouvoir les parcourir avec un autre programme.

### Des technologies modernes

Du côté serveur, une architecture MVC est implémentée avec le langage Scala couplé au framework Play, ainsi qu'aux APIs REST et Streaming de Twitter.

Du côté client, les technologies web basiques qui sont le HTML5, le CSS3 et le JavaScript (ainsi que des bibliothèques en découlant, comme jQuery) sont utilisées.

Le processus de réception de Tweets en temps réel utilise une communication WebSockets entre le serveur avec Akka Streams et le client avec JavaScript.

Auteur: Miguel Santamaria  
Répondant externe: -  
Prof. responsable: Nastaran Fatemi  
Sujet proposé par: Nastaran Fatemi