

Paris Attacks

A pretext to sentiment analysis on social media

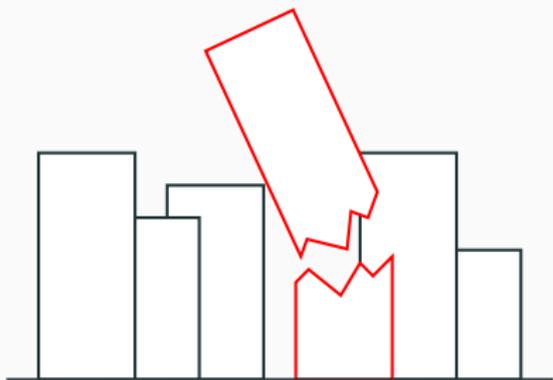
M. GABORIT – F. BLAIN

HAUM Talks – Janvier 2016

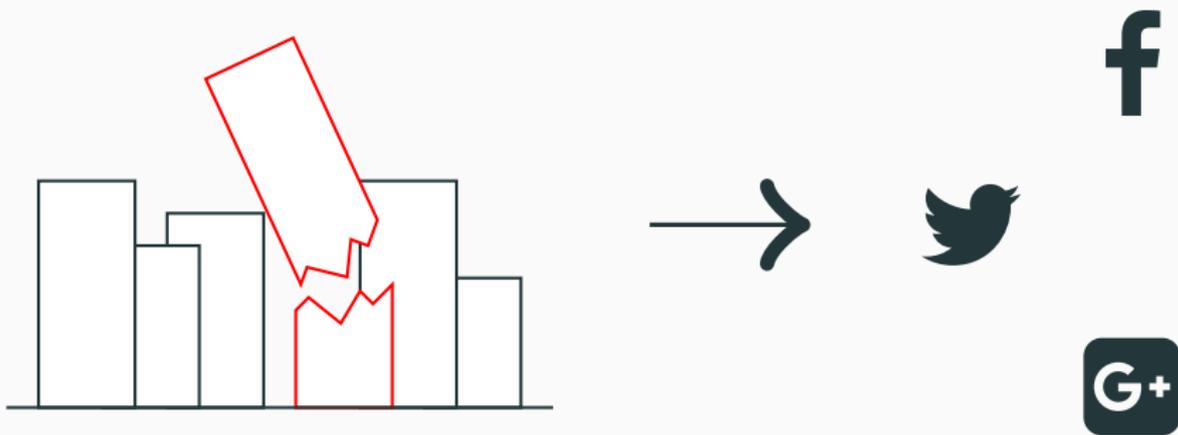
1. L'idée
2. Mise en oeuvre
3. Hypothèses et Pistes
4. Ce que font les autres

L'idée

Nouvelle vision des évènements



Nouvelle vision des évènements

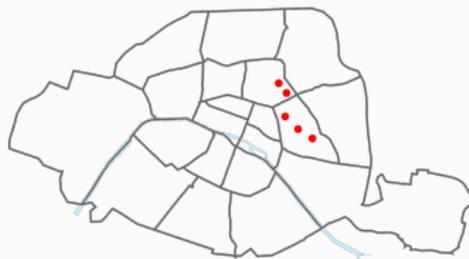


Un évènement entraîne une réaction **publique** et **spontanée**.

Attaques terroristes du 13 novembre

Très particulières pour plusieurs raisons :

- inattendues
- énormément de victimes et de témoins
- touchant une population assez jeune ?
- revendiquées *a posteriori* ▶▶



Attaques terroristes du 13 novembre

Réactions à la hauteur des attaques :

- énormément de messages postés (population connectée)
- réactions suivies sur plusieurs heures/jours [?]
- utilisation de téléphones portables ⇒ données géolocalisées **▶**
- grande variabilité des réactions dans le temps

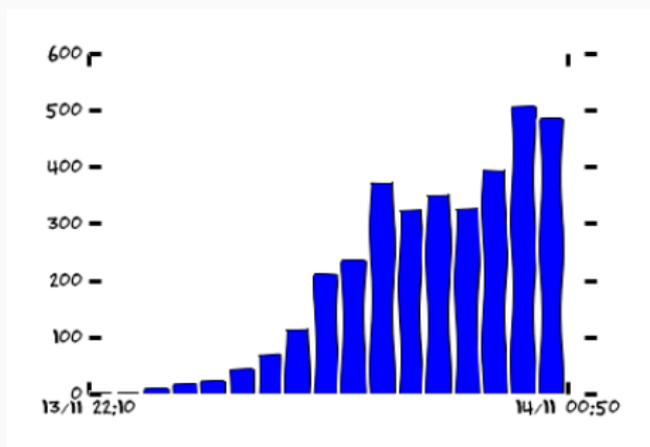


Figure 1: Évolution du nombre de tweets postés dans les premières heures

Est-il possible d'extraire des temps caractéristiques sociaux des tweets postés en rapport avec les attaques ?

Est-il possible d'extraire des temps caractéristiques sociaux des tweets postés en rapport avec les attaques ?

Temps caractéristique : temps permettant de distinguer différents évènements et/ou d'évaluer la vélocité de leur réponse.

Pourquoi les tweets ?

- Format court
 - Facile à analyser
 - Propice à des réactions claires et distinguables [?]
 - Rédaction faiblement réfléchie («à *chaud*»)
- Facilité de récupération
 - Utilisation du format natif de MongoDB
 - Traitement *off-line* efficace
 - Plusieurs passes possibles
- Grande taille du jeu de données (plusieurs milliers de tweets)



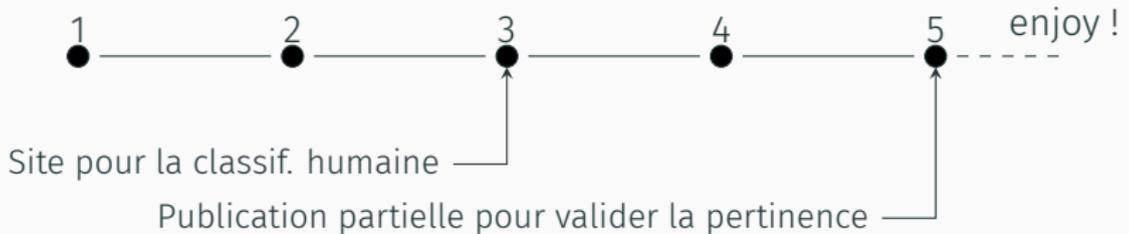
mongoDB®



Mise en oeuvre

Plan de bataille

1. Récupérer les tweets pour une analyse hors ligne (*#ParisAttacks*)
2. Classer les tweets suivant le langage
3. Classer manuellement un sous-ensemble de tweets
4. Utiliser le sous-ensemble classé pour classer automatiquement le reste
5. Proposer des observables afin de déduire les temps caractéristiques sociaux



Hypothèses et Pistes

Réactions suivies

Peu d'ambiguïté

Changement post-revendication

Réactions suivies

A vérifier en suivant l'activité des comptes parmi les tweets récupérés

Peu d'ambiguïté

Hypothèse raisonnable dans un cadre post-traumatique, soutenue par le format court.

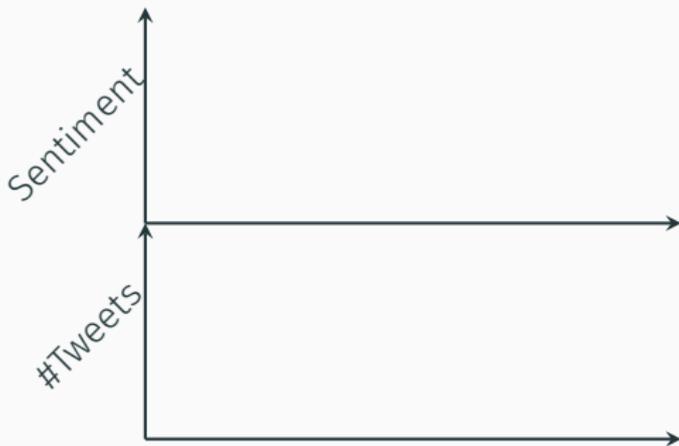
Changement post-revendication

A analyser à vérifier *a posteriori*... peut être faudra-t-il exclure les tweets évoquant Daech directement

- Utiliser la **géolocalisation** pour analyser les données avec une composante géographique
- Corréler les résultats avec la **parution de reportages/d'images/d'informations**
- Comparer avec d'**autres évènements** marquants (sportifs, politiques, etc...)
- Analyse par **canaux** :
 - sources médiatiques
 - politique
 - autres...

Pistes

- Mettre en regard avec les évènements de Charlie Hebdo (hypothèse de **réaction couplée**)
- Explorer les **modes d'auto-organisation** (PorteOuverte🐦, l'm Safe**f**)
- **Auto-émulation** ou **réponse à un stimulus** (tweet d'un politique, d'un média)
- Analyse **bi-variante**



A propos des statistiques

- Panel non-représentatif de la population
- Premières réactions biaisées par les annonces des journaux télévisés (le media n'est *jamais* impartial)
- Biais induit par la classification manuelle
- etc...

Ce que font les autres

Ce que font les autres



Michal Lukasik, Trevor Cohn, and Kalina Bontcheva.
Classifying Tweet Level Judgements of Rumours in Social Media.

In *Proceedings of EMNLP*, volume 796, pages 2590–2595, 2015.



Michal Lukasik, Trevor Cohn, and Kalina Bontcheva.
Modeling Tweet Arrival Times using Log-Gaussian Cox Processes.

In *Proceedings of EMNLP*, volume 796, pages 250–255, 2015.



Laurent Luce.
Twitter sentiment analysis using Python and NLTK.



Oriane Piquer-Louis.
Documenter l'expérience habitante de la ville sur les réseaux sociaux : corpus photographiques et données numériques, le sens des collections.

Merci !

Des questions ?

fred@haum.org – mathieu@haum.org

On accepte aussi les coups de main !