

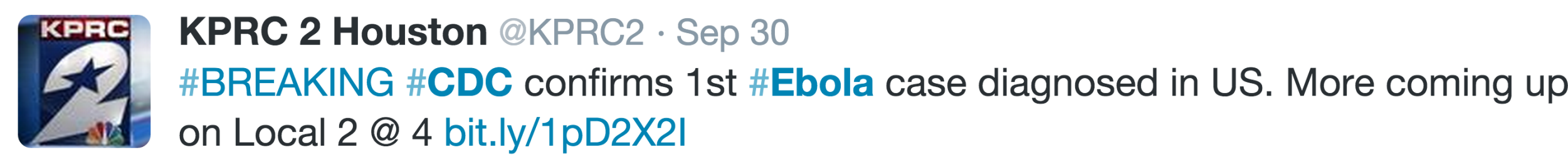
NEWSFLASH



Motivation @newsflash

Twitter promised to be the heartbeat of the world. How can we realize that promise?

Journalists are in the dark when it comes to finding stories evolving in real time. For large scale events, traditional sources may work. Micro-events still have to be manually discovered, like the first case of Ebola in the US.



Ebola Is Diagnosed in Texas, First Case Found in the U.S.

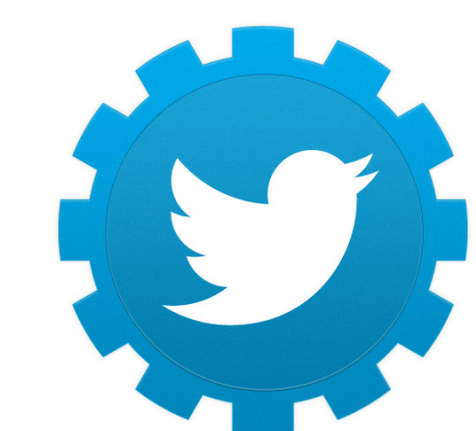
By DENISE GRADY SEPT. 30, 2014

Newsflash leverages Twitter to detect these events in real time, and weights them using Newsflash's Acceleration and Clustering algorithms. Journalists use Newsflash to discover stories as they happen.



Architecture @pipe_line

A Python application interfacing with the Twitter API and web clients over sockets.



NEWSFLASH

Parse: Stem, Tokenize

Extract: URLs, Handles

Update: Bounding, Ranking, Acceleration



"In breaking news you're either first or last"

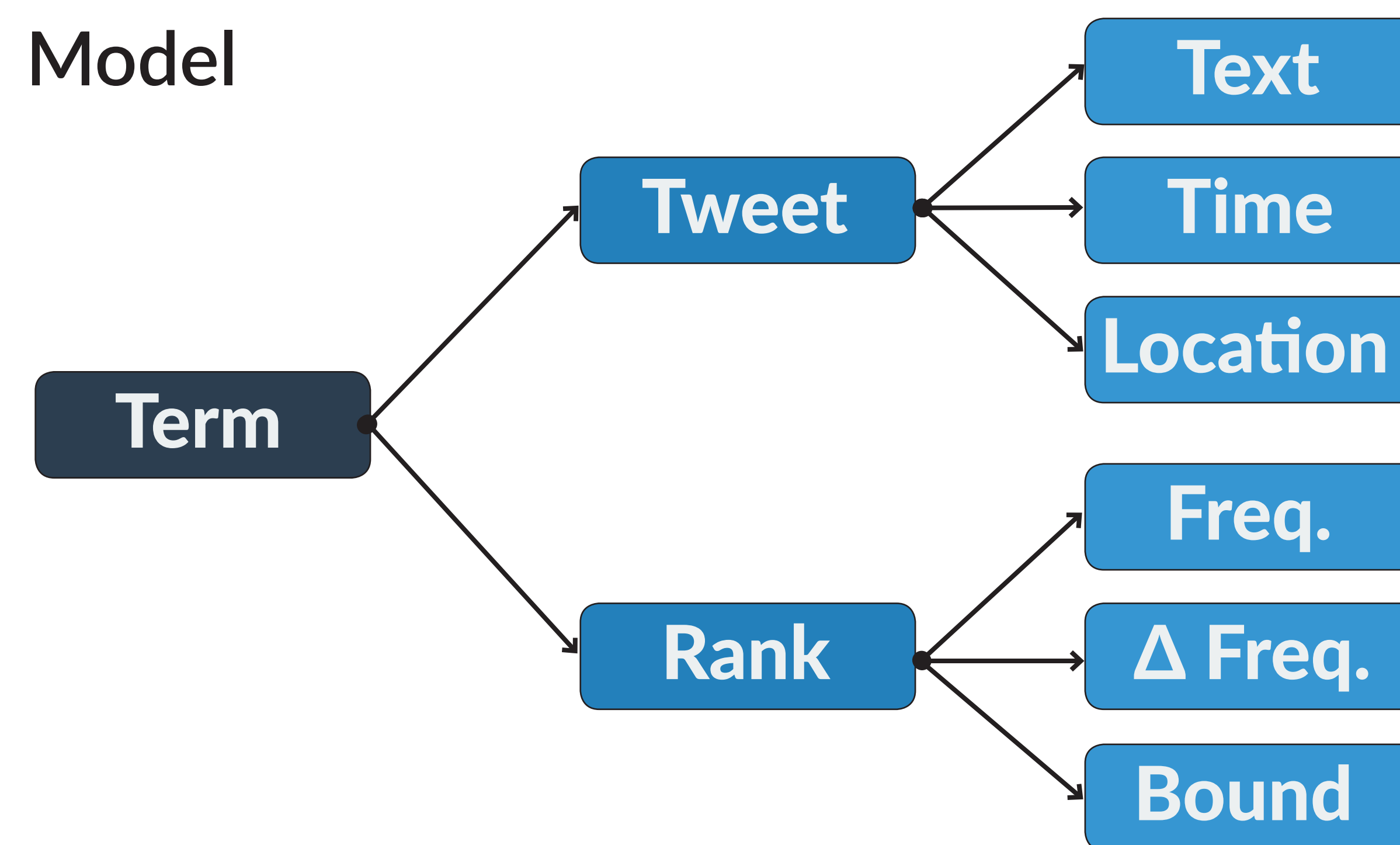
Realtime Discovery of Latent Trends in Tweets



Analysis @data_models

Newsflash's core focus is on extraction of topics of velocity within small geographic areas.

Model



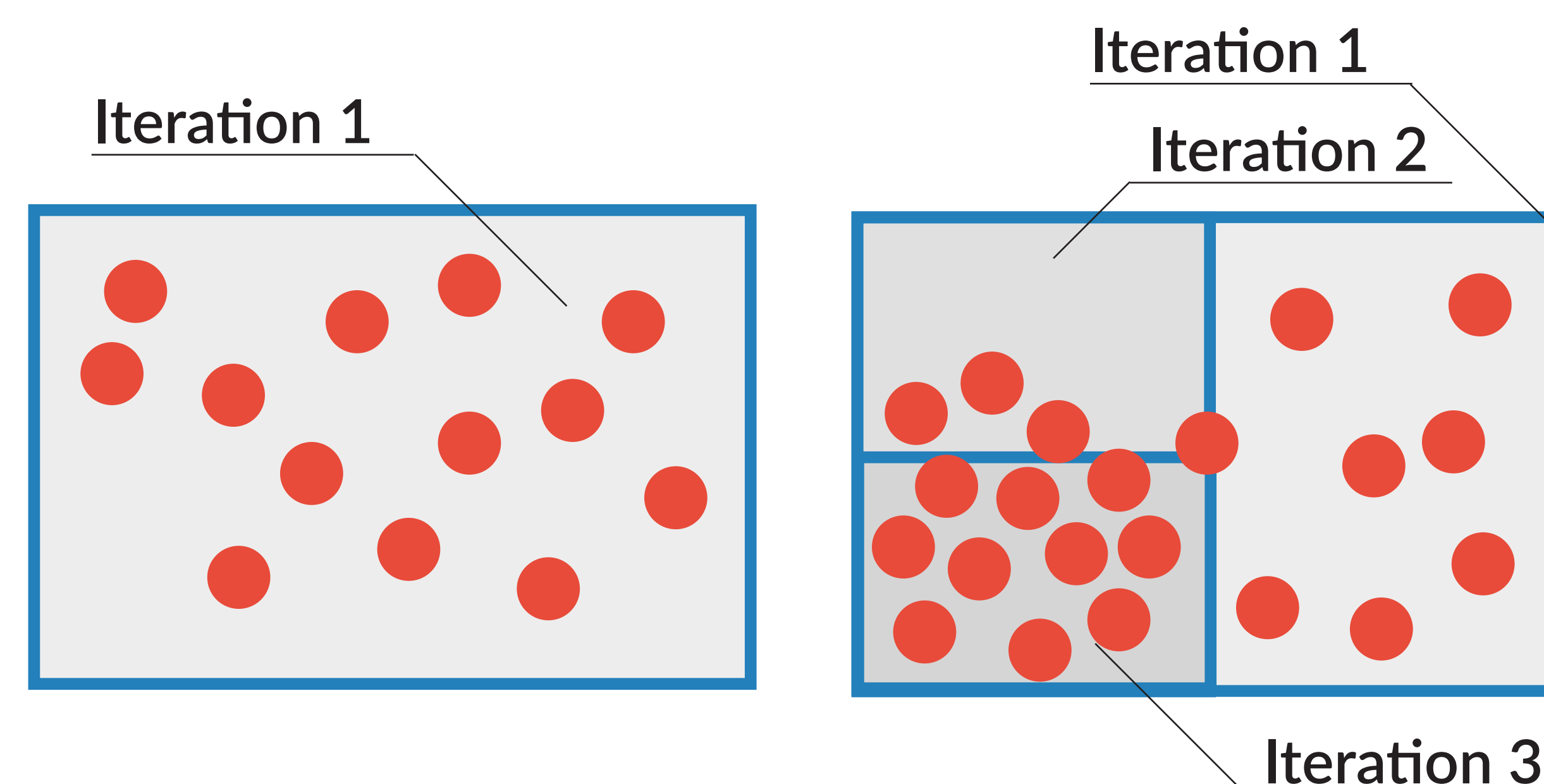
Training

Newsflash trains an internal model on a corpus of collected Tweets at launch. After training, the model continues to learn in realtime from a live stream.

Terms are weighted within the model by a multifeatured vector, with the strongest influence being *acceleration*.

Geographic Clustering

Terms in the trained model are clustered using an iterative greedy bounding box approach derived from entropic decision trees.



Aaron Behr and Alex Hadik

@abehr, @ahadik



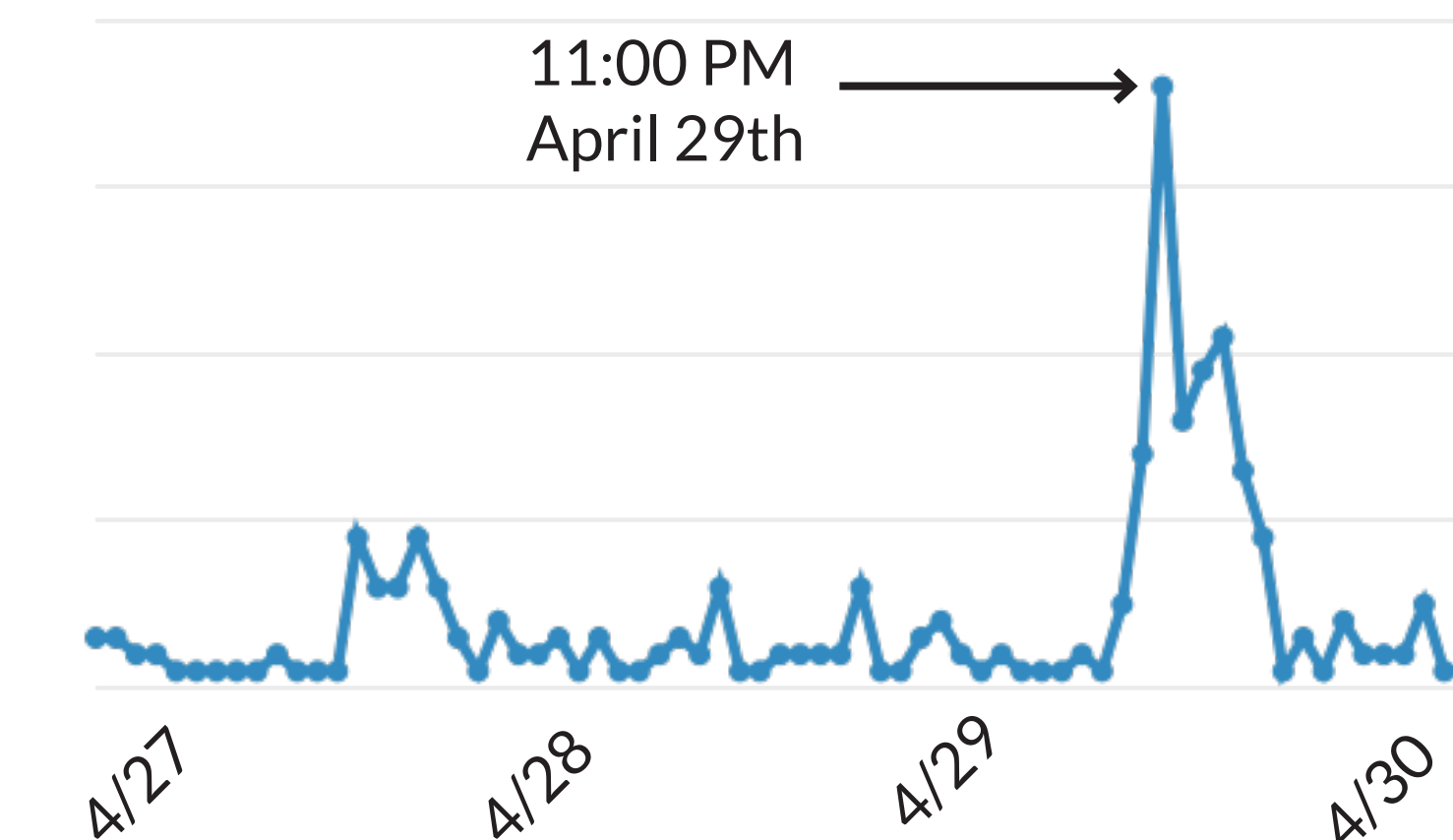
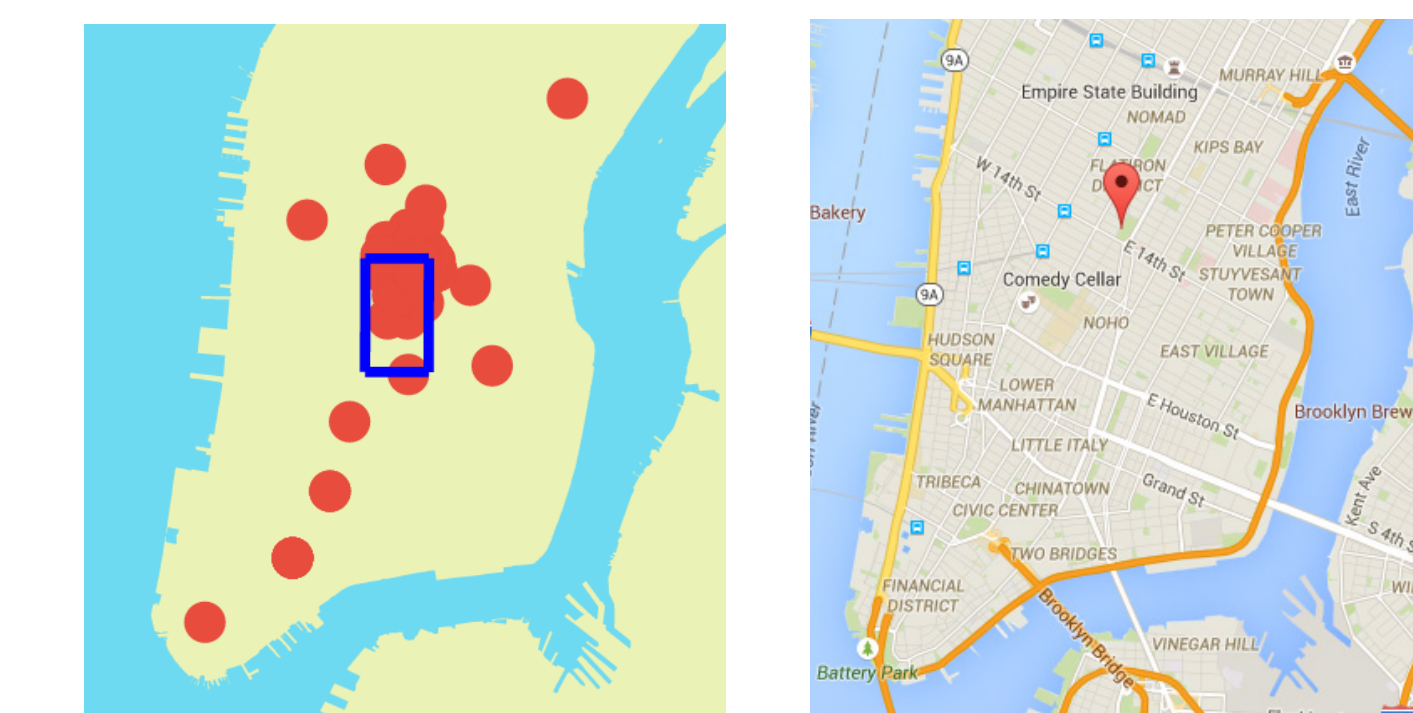
Results @trendspotter

Newsflash successfully extracts topics exhibiting abnormal frequencies and geographic clustering. [#winning](#)

Newsflash presents D3.js visualizations of current activity on Twitter as well as the rankings of the currently trained Newsflash model. This interface is used to explore the topics the short term trending topics.

Over 140 Arrested as New Yorkers Protest the Death of Freddie Gray

By BENJAMIN MUELLER and JOHN SURICO APRIL 30, 2015



Using Newsflash to analyze the term “protest” from April 26th to 30th identifies the Freddie Gray protests in Union Square. Newsflash not only successfully determines geographic bounds based on tweet density, but also correctly identifies increased velocity of the topic.



Future Work @to_do

Newsflash stands out from existing analysis platforms by finding instantaneous trends but can go further.

- User defined parameter controls and sorting.
- Improved accuracy of cluster detection by bounding box algorithm.
- Further exploration of ranking methods.
- Geographic generalization enabling Newsflash to work in multiple cities.
- Further product development towards full platform.