

Edge Analytics—Real-time Analysis of Social Media

Objective

First responders at the edge want to monitor open source social media feeds in real-time to enhance public safety.

Operational Challenges

Increased cognitive load

Volume and velocity of raw data increases the cognitive load on edge users.

Dynamic environment

Relevance is dynamic

Discovery

What needs to be monitored is unknown and changing

Local Context

Global trends may not be relevant to local events.

Hypothesis and Technical Challenges

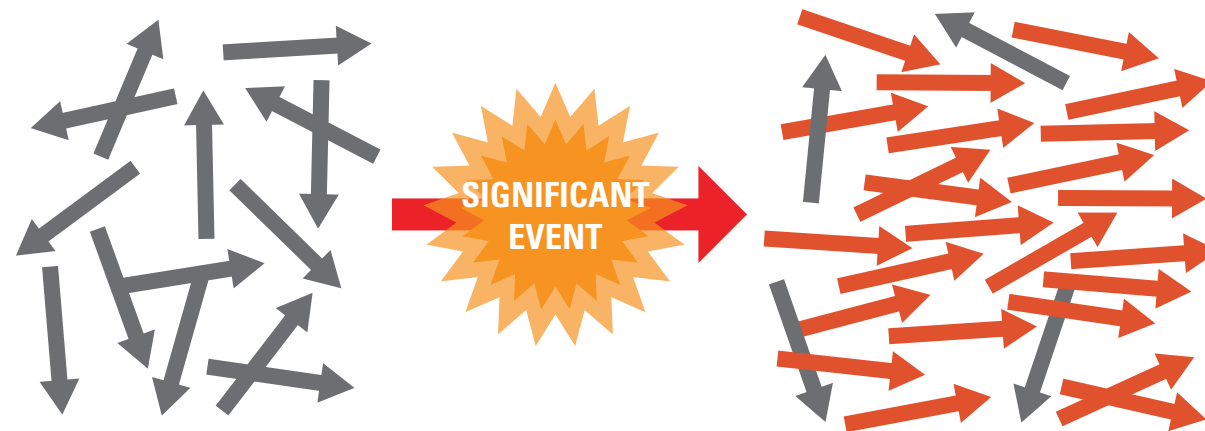
Hypothesis:

Given a local context where patterns of interest are scattered without any dominant trend or pattern, events and emerging trends can be identified in sufficient time to inform operations.

Technical Challenges:

How can we summarize streaming data?

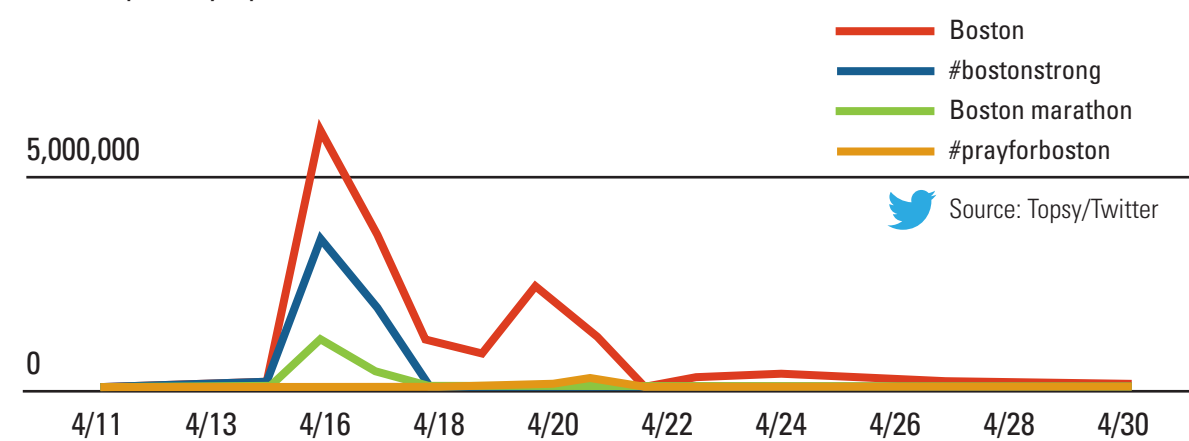
How fast can we determine changes?



Given a local context, Significant Events are detectable as dominant trends or patterns

Boston dominated the news agenda

Tweets per day April 2013, ET

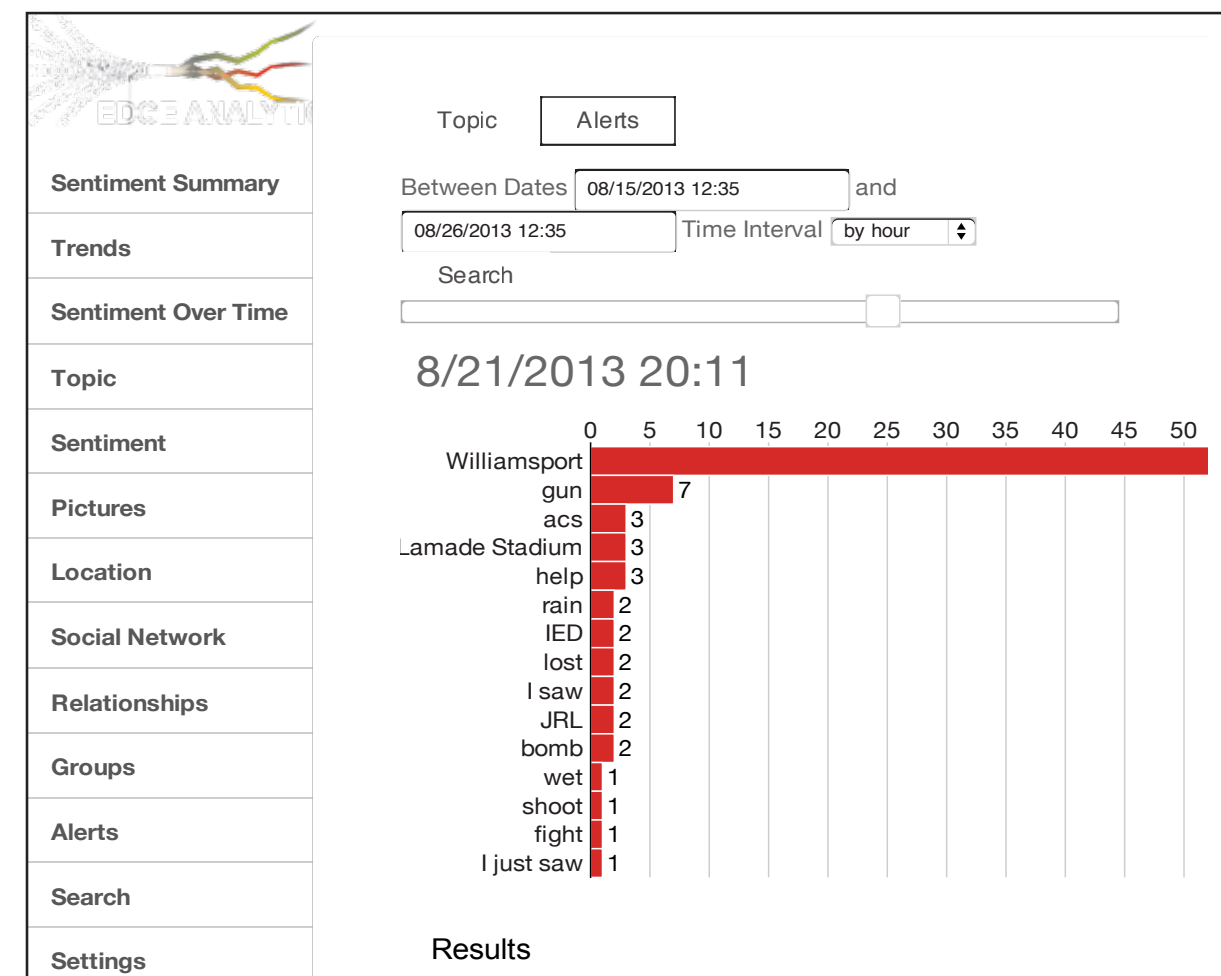


Can the peak be detected in time to influence reaction?

Solution

Monitor social media streams (e.g., Twitter) in real time

- Sentiment analysis
- Topic modeling
- Social network analysis

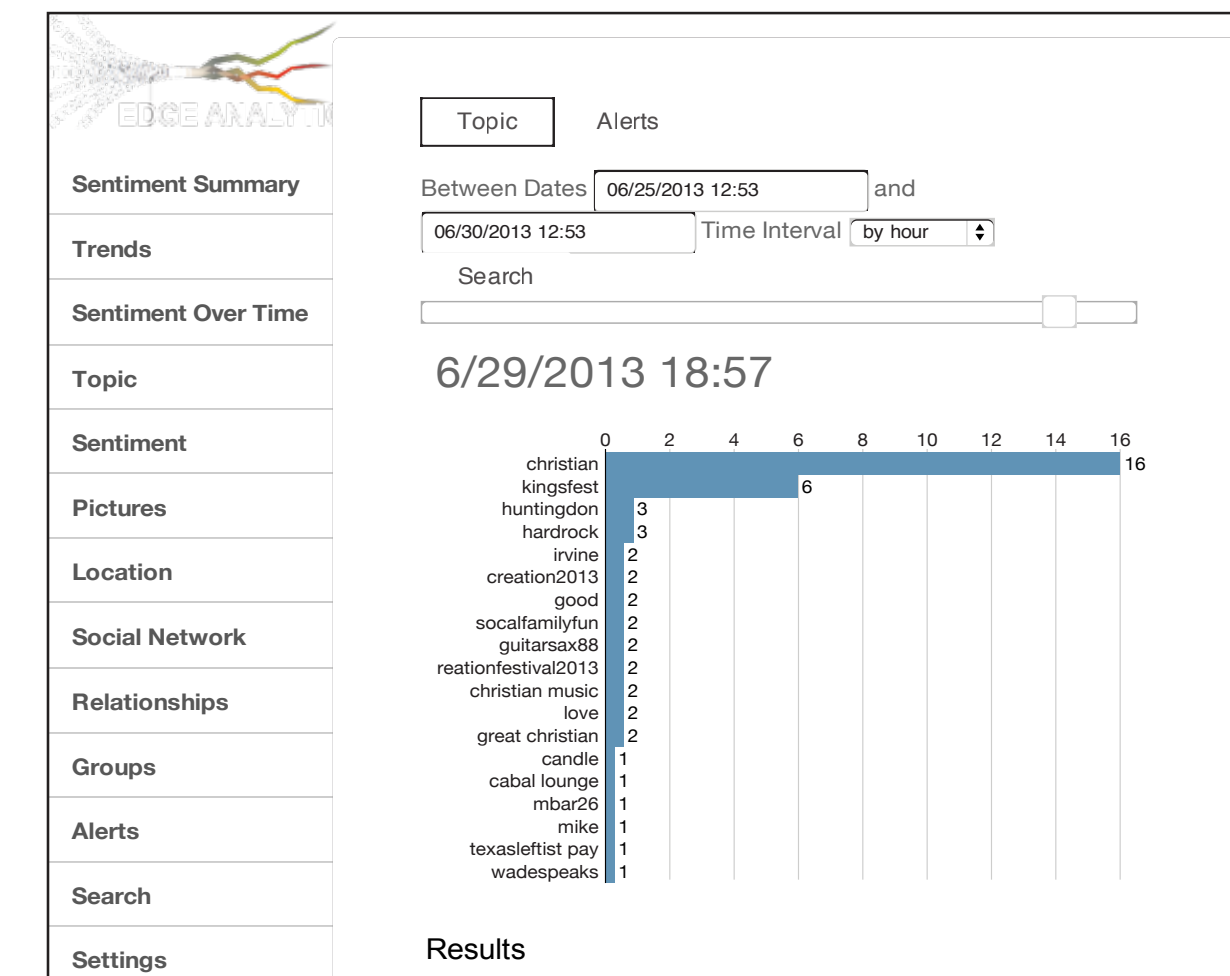


Flexible infrastructure

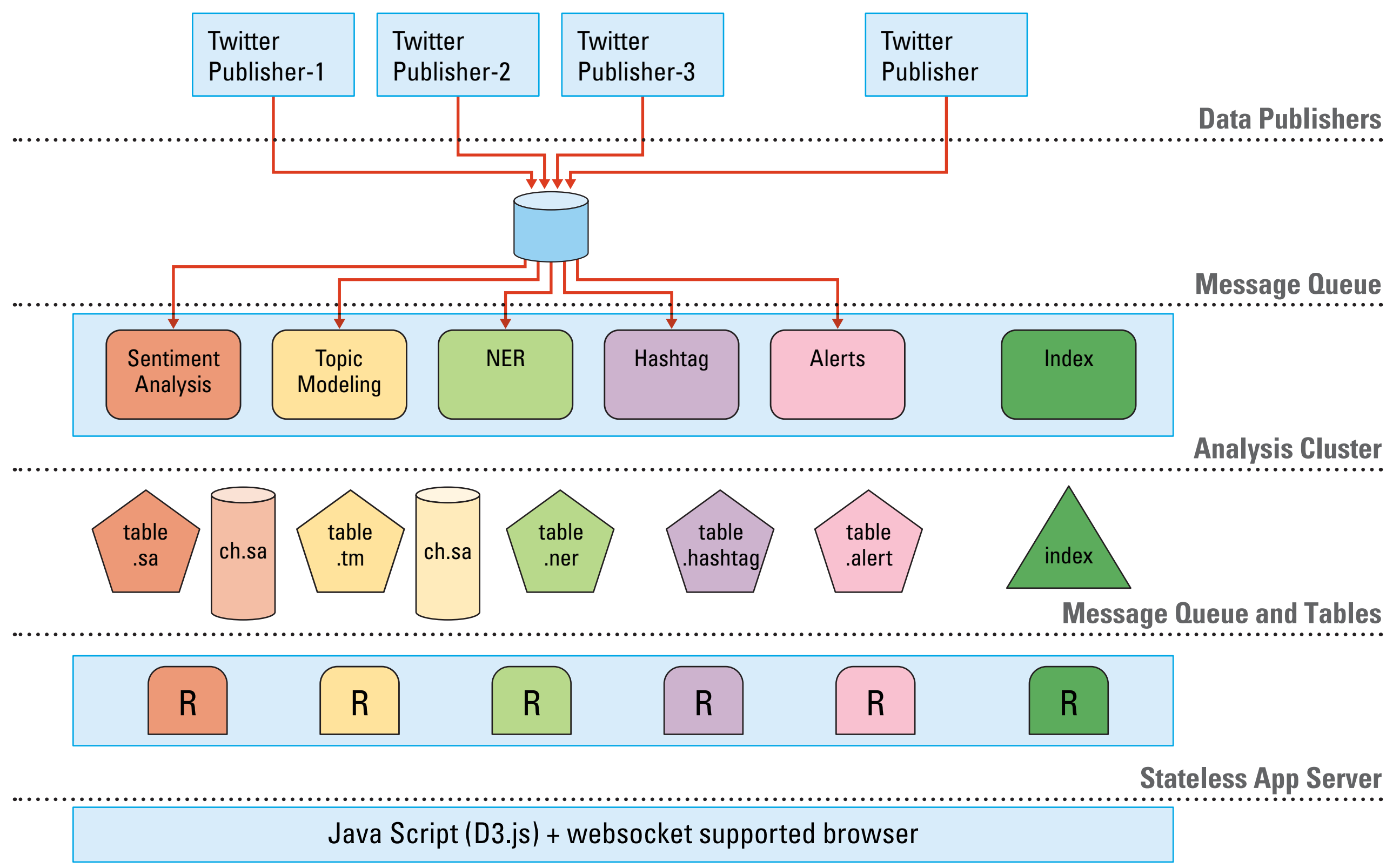
- Performance: Process real-time streams
- Scalability: Quickly from a laptop to enterprise cloud
- Platform: Support multiple programming languages
- Development Time: Rapid and parallel development

Reduce Cognitive Load

- Flexible searching
- Custom alerting
- Transform data into patterns
- Visualize patterns



- Reuse: Existing & open source components
- Fault tolerant: Ability to support failures
- Modifiability: Quickly add new data/analysis

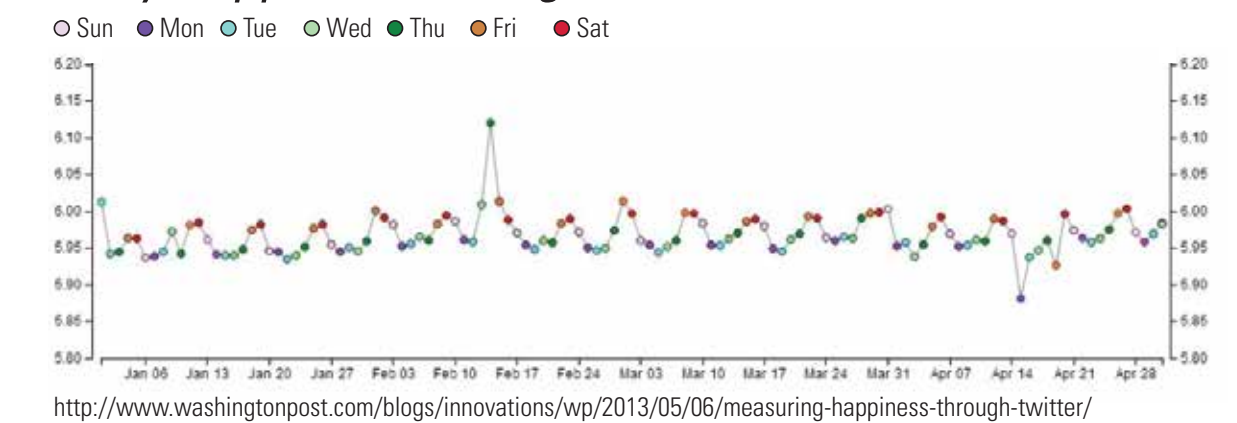


Types of Trend Analysis

Macro

- Collective trends
- Noise-tolerant
- Privacy protecting
- Large data volume

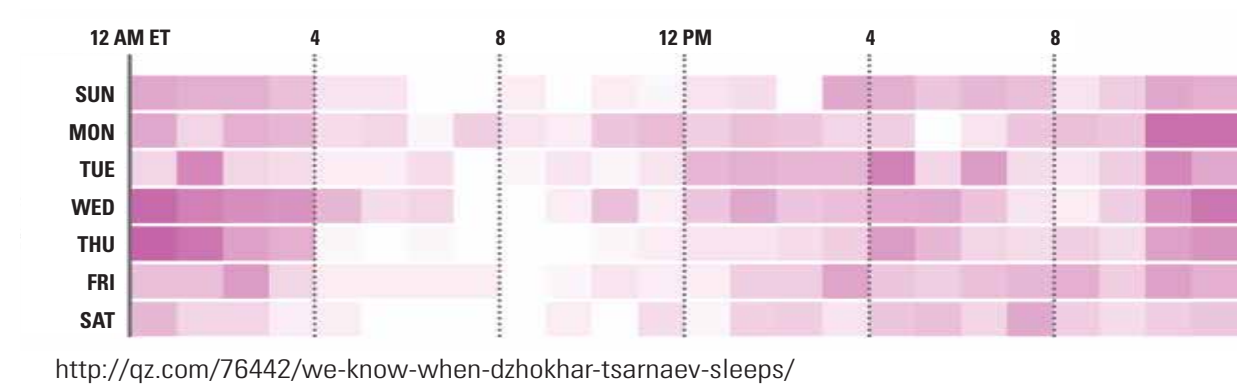
Daily Happiness Averages for Twitter, 2013



Micro

- Targeted and context specific
- Requires higher accuracy
- Privacy issues

Sleep patterns of the Boston bomber based on Twitter Metadata



Field Experiments and Results

Pilots

1. Huntingdon County EMA (Creation Festival 2013)
2. Weapons of Mass Destruction Civil Support Team (Little League World Series)

Observations & Feedback

HCEMA and WMD-CST showed the potential

- HCEMA - "Harlem Shake": Event took 7 minutes to verify after Edge Analytics picked it up
- WMD-CST - "Bomb threat": Detected and determined to be a hoax

Challenges and Future Work

- False positives from keywords
- Dynamic and adaptive discovery of relevant artifacts
- Non-textual inputs
 - Voice
 - Multi-lingual challenges and Leetspeak
 - Images and Video
- Feedback to support preventative/influence operations