Real-time Social Media Analytics through Semantic Annotation and Linked Open Data

The University Of Sheffield.

Diana Maynard, Mark Greenwood, Ian Roberts, George Windsor and Kalina Bontcheva

Nesta...

Analysing Social Media

- Real-time social media analytics needed by journalists, companies, governments etc.

- Processing and understanding text from social media is hard: messages are often brief, ungrammatical, and noisy

- Interpretation is heavily dependent on temporal, geographical, and social network context

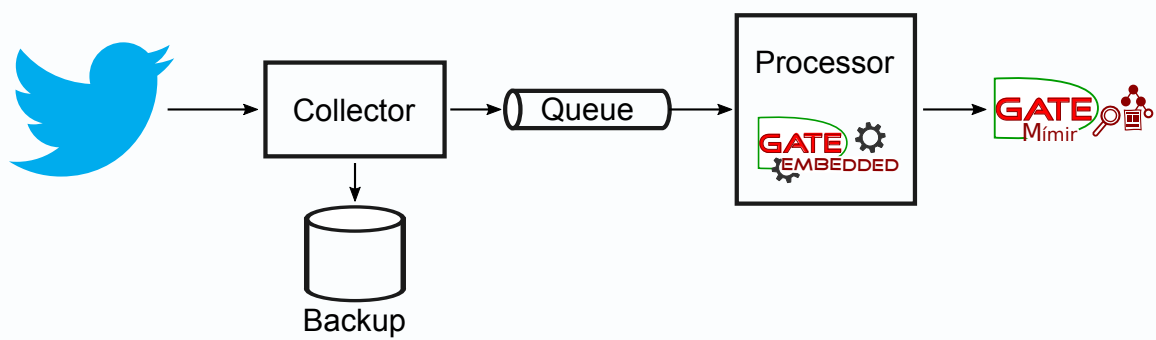## What kind of questions can we answer?

- Find all positive sentiment expressions about the "UK economy" theme in tweets written by Labour candidates for constituencies in Greater London".

- Which political issues do people care about the most? Which ones do they engage with and why?

- Which party's manifesto is the most positive, and how does this influence public voting?

Semantics is Crucial!

- Information isn't always explicit in the text

- We need semantic information in order to aggregate knowledge, e.g. all tweets from MPs in a certain geographical region

Searching Index "2015-03-09"

(DocumentAuthor author_party="Conservative Party") OVER
(Topic theme="uk_economy")

Search

Richard Short
@TorySharty

With so many protected from Labour's pension raid are they sure it will even generate £2.7bn #bbcsp
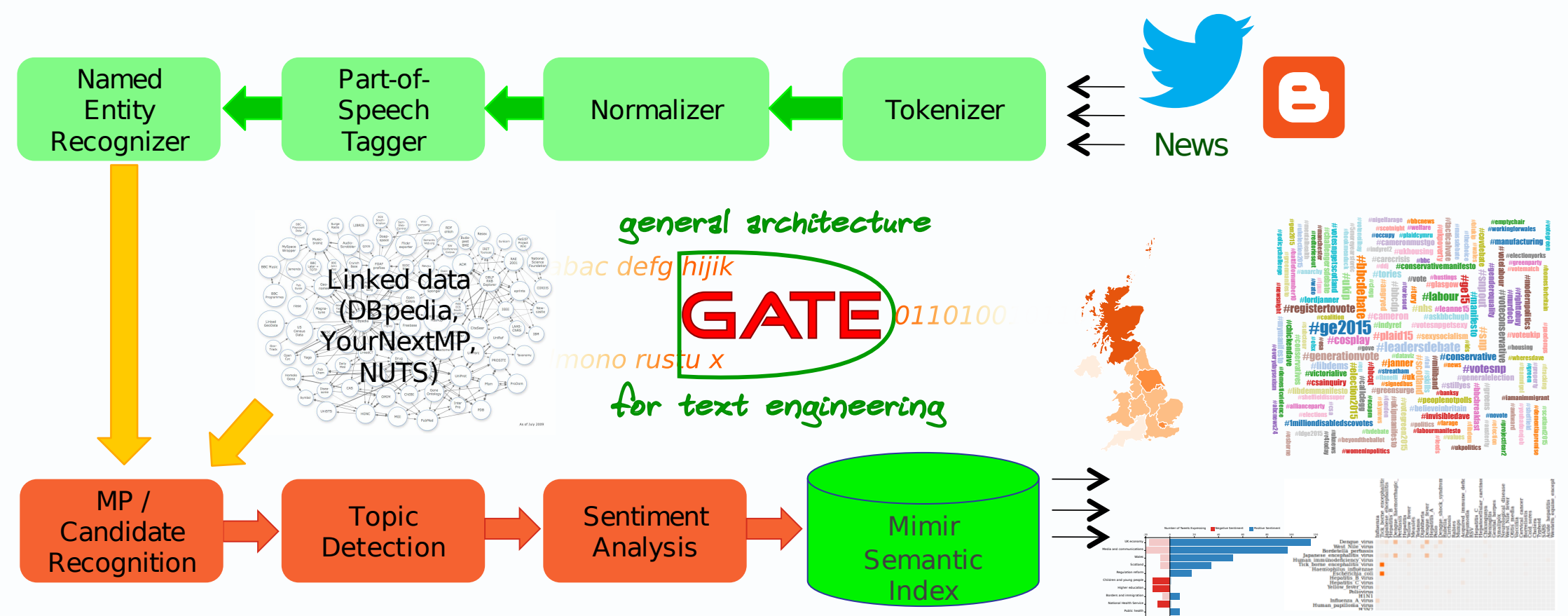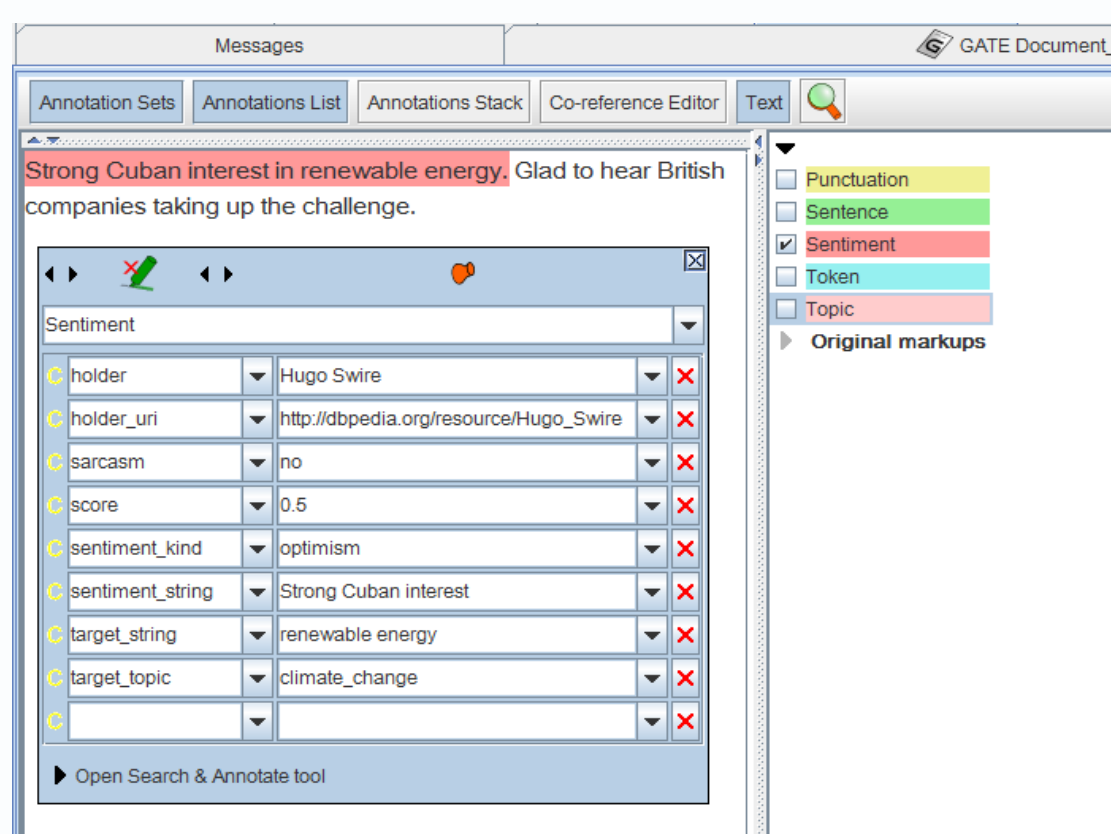
Real-Time Big Data Analysis

- Data collection via Twitter streaming API
- Documents stored as JSON and annotated via GCP
- Documents indexed via MIMIR
- Search and visualisation via MIMIR/Prospector

Political Futures Tracker


What sentiments are expressed?

Climate change is less boring than other political topics

- Climate change isn't mentioned frequently by politicians

- But it shows high levels of engagement:
  - frequent retweets
  - contains sentiment
  - optimistic
  - contains URLs and @mentions

- People engage more with climate change than other topics because they feel they can actively contribute
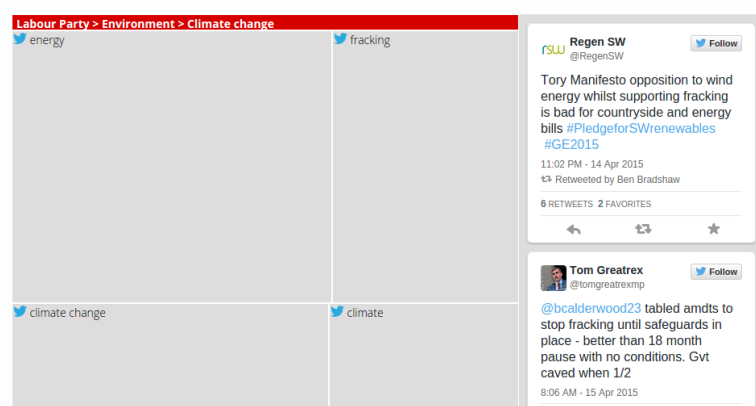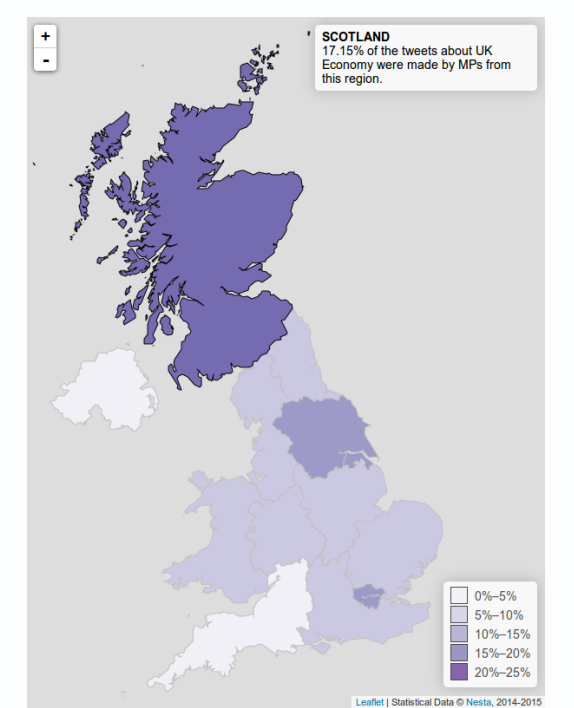
Sentiment visualisation: where did Miliband go wrong?
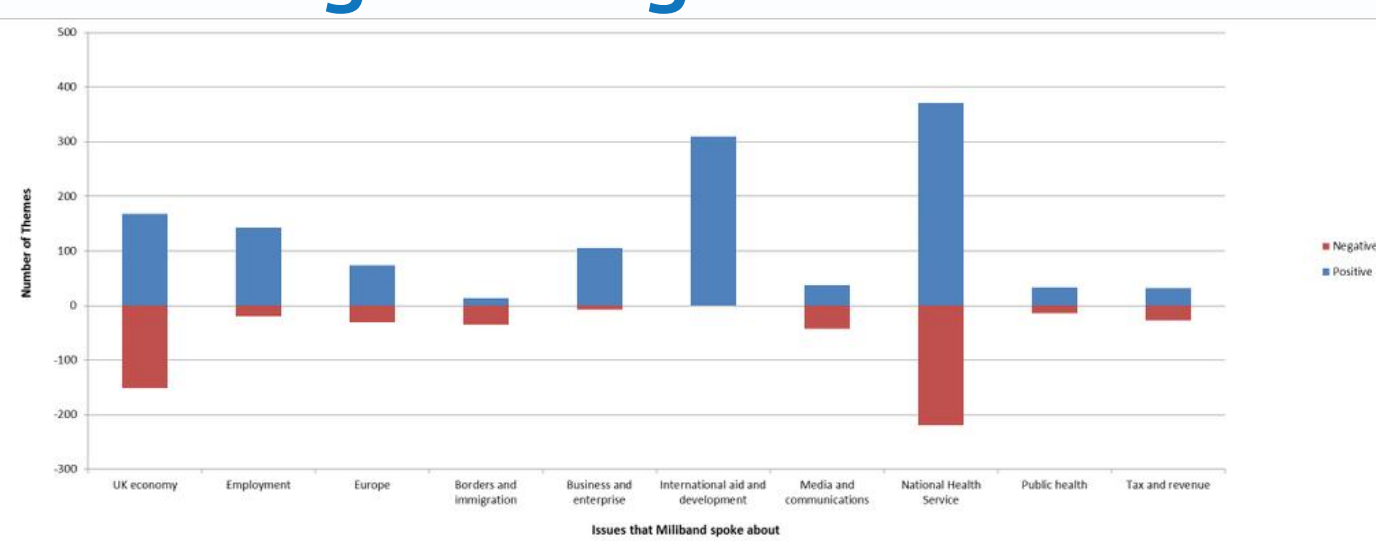
- Main topics and related sentiment of discussion on Twitter generated by the #Leadersdebate, between the leaders of the 7 main parties.

- Milliband was negative about the economy and the NHS

Topic visualisation

- Where did people talk about the economy?

SEVENTH FRAMEWORK PROGRAMME

Nesta...