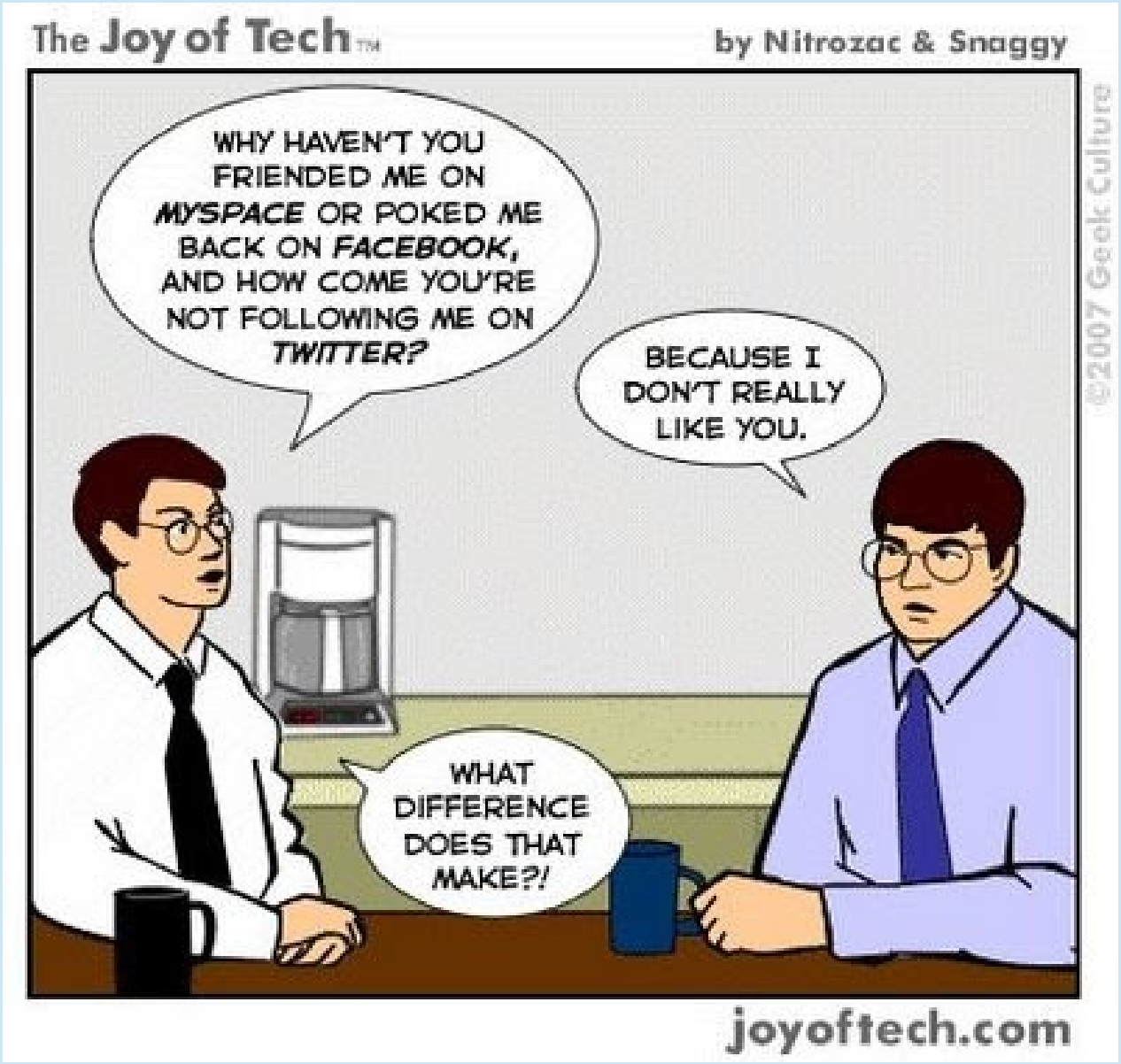
The University Of Sheffield.

general architecture
GATE
for text engineering

Do we really know what people mean when they tweet?

Dr. Diana Maynard
University of Sheffield, UK

We are all connected to each other...



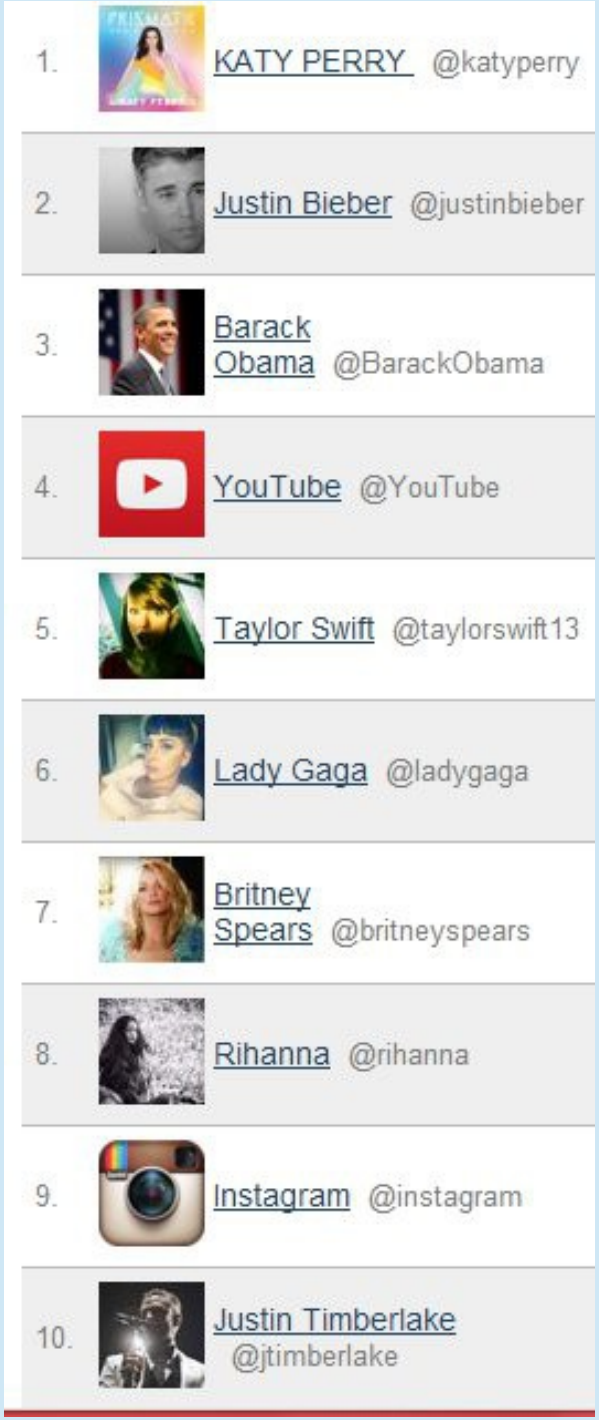
- Information, thoughts and opinions are shared prolifically on the social web these days
- 72% of online adults use social networking sites

Your grandmother is three times as likely to use a social networking site now as in 2009

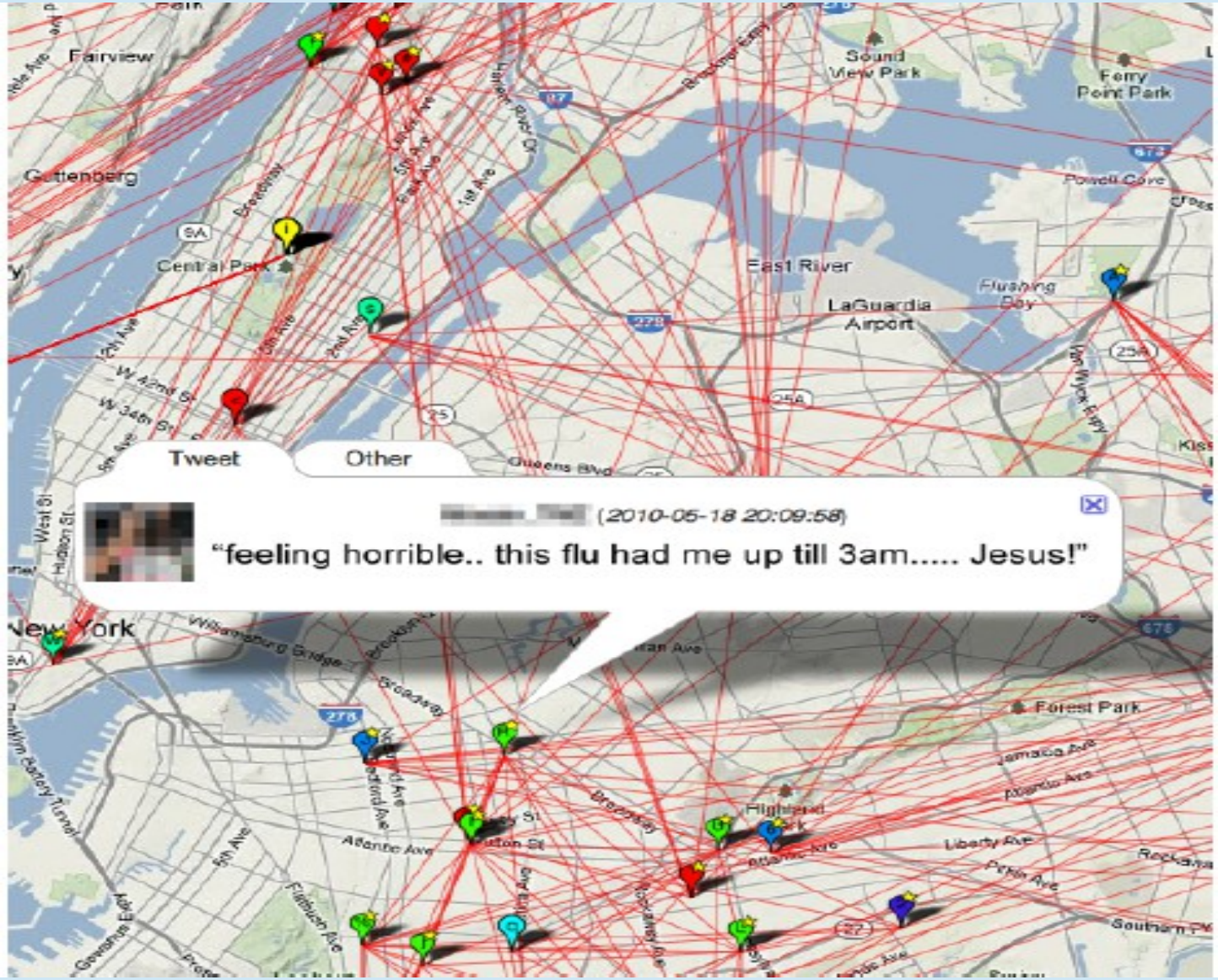
- In Britain and the US, approx 1 hour a day on social media
- 30 % of Americans get all their news from Facebook
- Facebook has more users than the whole of the Internet in 2005
- 40% of Twitter users don't actually tweet anything

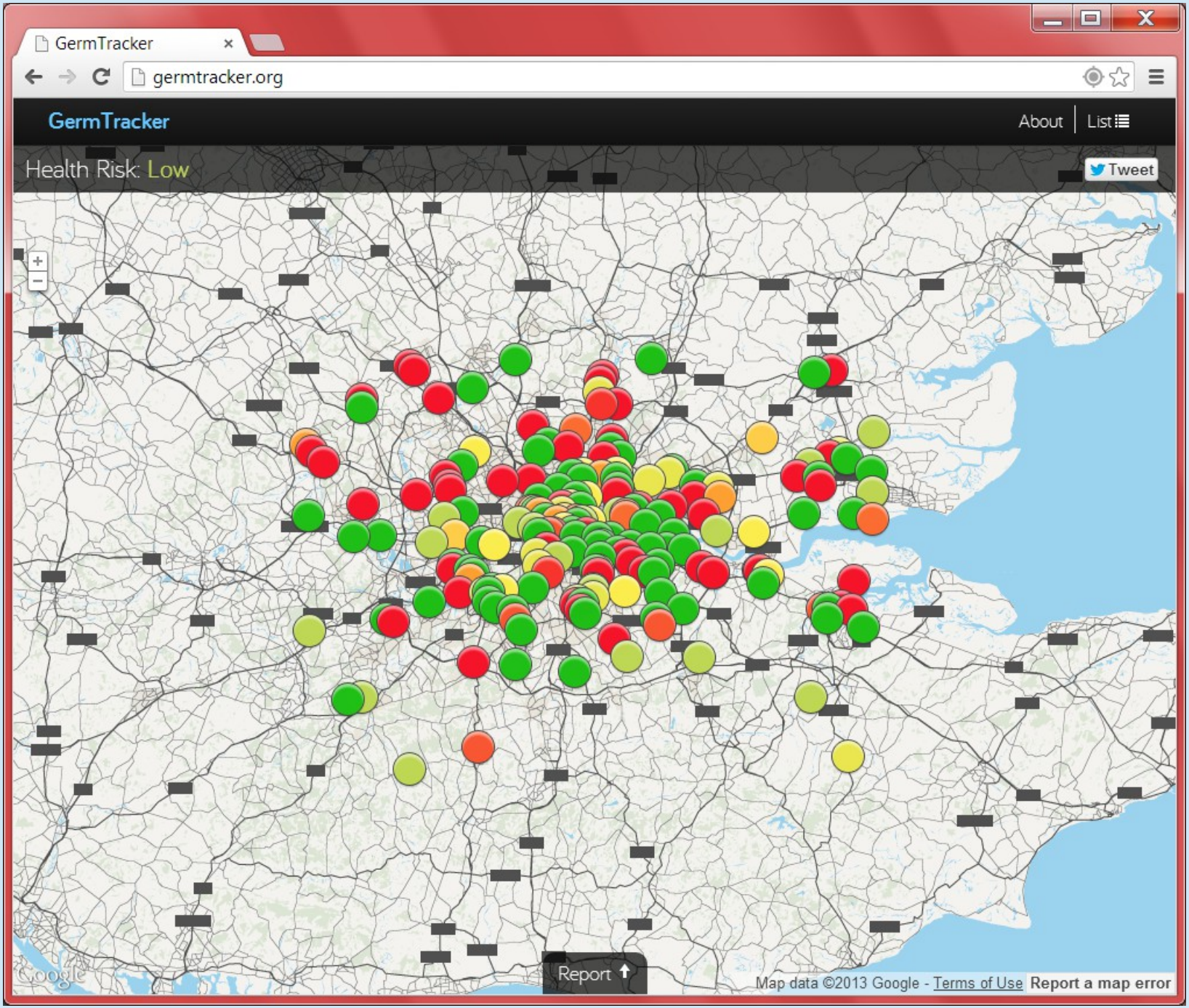
What are people reading about?

- Of the top 10 Twitter accounts with the highest number of followers:
  - 7 pop stars
  - 2 social media sites
  - and Barack Obama
- As you can imagine, there's a lot of mindless drivel on social media sites



There can be surprising value in mindless drivel




GermTracker
germtracker.org
GermTracker
About
List
Health Risk: Low
Tweet
Report
Map data ©2013 Google - Terms of Use
Report a map error

Opinion mining is about finding out what people think...

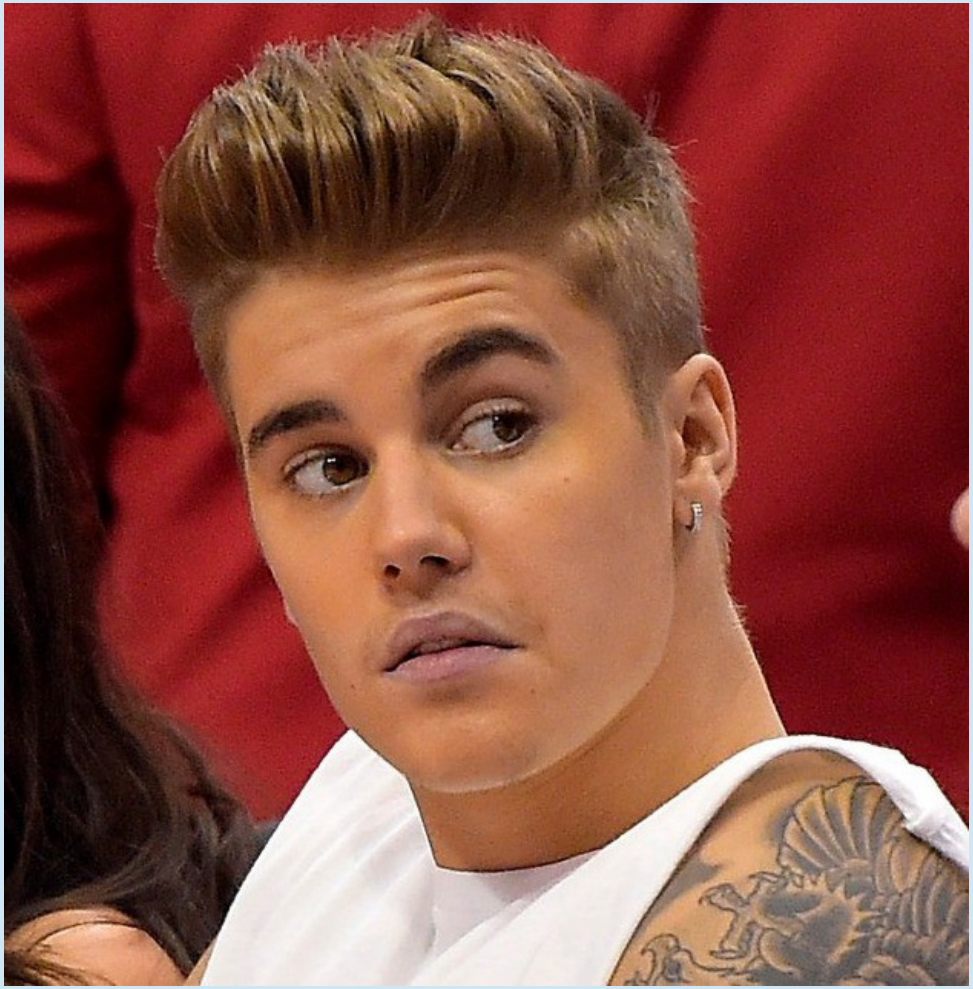
Voice of the Customer

- Someone who wants to buy a camera
  - Looks for comments and reviews
- Someone who just bought a camera
  - Comments on it
  - Writes about their experience
- Camera Manufacturer
  - Gets feedback from customer
  - Improves their products
  - Adjusts marketing strategies

More than just analysing product reviews

- Understanding customer reviews and so on is a huge business
- But also:
    - Tracking political opinions: what events make people change their minds?
    - How does public mood influence the stock market, consumer choices etc?
    - How are opinions distributed in relation to demographics?
- NLP tools are crucial in order to make sense of all the information

“Climate change? Oh great!” says Justin Bieber



What else do we want to investigate?

- Decarbonet project: investigating public perception of climate change
- How do opinions change over time and what causes these changes?
- How can we influence opinion?
- Sarcasm detection
- What social media content should be preserved/forgotten? Interestingness of opinions (ARCOMEM and ForgetIt projects)

But there are lots of tools that “analyse” social media already….

- Streamcrab http://www.streamcrab.com/
- Semantria http://semantria.com
- Social Mention http://socialmention.com/
- Sentiment140: http://www.sentiment140.com/
- TipTop: http://feeltiptop.com/

Why are these sites unsuccessful?

- They don't work well at more than a very basic level
- They mainly use dictionary lookup for positive and negative words
- Or they use ML, which only works for text that's similar in style to the training data
- They classify the tweets as positive or negative, but not with respect to the keyword you're searching for
    - keyword search just retrieves any tweet mentioning it, but not necessarily about it as a topic
    - no correlation between the keyword and the sentiment

“Positive” tweets about fracking

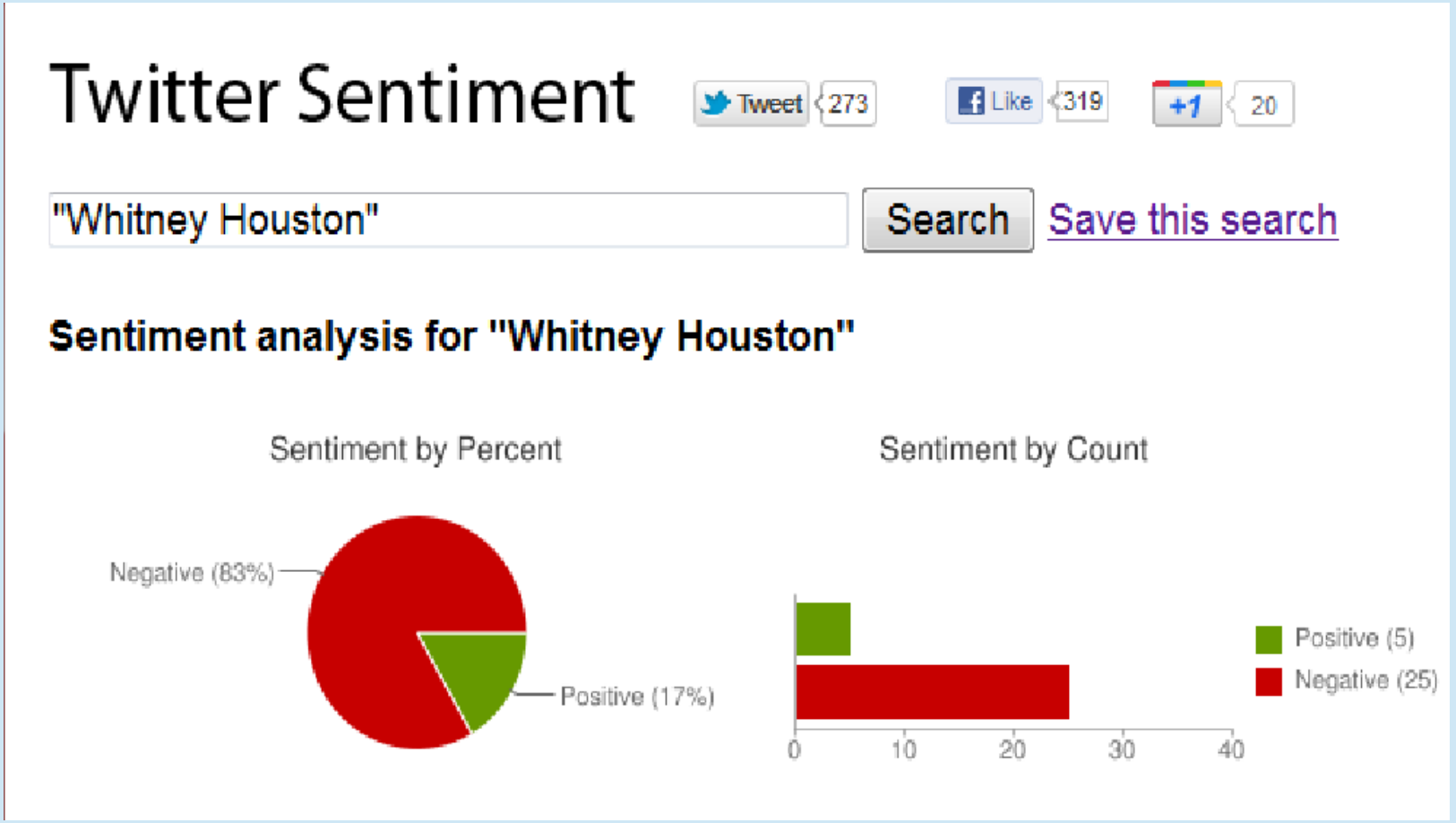
- *Help me stop fracking. Sign the petition to David Cameron for a #frack-free UK now!*
- *I'll take it as a sign that the gods applaud my new anti-fracking country love song.*
- *#Cameron wants to change the law to allow #fracking under homes without permission. Tell him NO!!!!!*

It's not just about looking at sentiment words

“It's a great movie if you have the taste and sensibilities of a 5-year-old boy.”
“I'm not happy that John did so well in the debate last night.”
“I'd have liked the film if it had been shorter.”
“You're an idiot.”

- Conditional statements, content clauses, situational context can all play a role

Whitney Houston wasn't very popular...



Death confuses opinion mining tools

TeghanSimone: Radio playing **Whitney Houston**.. I swear I'm about to cry... So sad
Posted 5 minutes ago

- Opinion mining tools are good for a general overview, but not for some situations

Opinion spamming

Suppose we run a contest where people retweet our ad repeatedly, and the winner's whoever loses the most followers.

Spam opinion detection (fake reviews)

- Sometimes people get paid to post “spam” opinions supporting a product, organisation or even government
- An article in the New York Times discussed one such company who gave big discounts to post a 5-star review about the product on Amazon
- http://www.nytimes.com/2012/01/27/technology/for-2-a-star-a-retailer-gets-5-star-reviews.html?_r=3&ref=business
- Could be either positive or negative opinions
- Generally, negative opinions are more damaging than positive ones

How to detect fake opinions?

- Review content: lexical features, content and style inconsistencies from the same user, or similarities between different users
- Complex relationships between reviews, reviewers and products
- Publicly available information about posters (time posted, posting frequency etc)

Irony and sarcasm

- *I had never seen snow in Holland before but thanks to twitter and facebook I now know what it looks like. Thanks guys, awesome!*
- *Life's too short, so be sure to read as many articles about celebrity breakups as possible.*
- *I feel like there aren't enough singing competitions on TV . #sarcasmexplosion*
- *I wish I was cool enough to stalk my ex-boyfriend ! #sarcasm #bitchtweet*
- *On a bright note if downing gets injured we have Henderson to come in*

How do you know when someone is being sarcastic?

- Use of hashtags in tweets such as #sarcasm, emoticons etc.
- Large collections of tweets based on hashtags can be used to make a training set for machine learning
- But you still have to know which bit of the tweet is the sarcastic bit

Man , I hate when I get those chain letters & I don't resend them , then I die the next day .. #Sarcasm

I am not happy that I woke up at 5:15 this morning. #greatstart #sarcasm

You are really mature. #lying #sarcasm

Case study:
Rule-based Opinion Mining on Tweets

Why Rule-based?

- Although ML applications are typically used for Opinion Mining, there are advantages to using a rule-based approach when training data isn't easily available
- For example, working with multiple languages and/or domains
- Rule-based system is more easily adaptable
- Novel use of language and grammar makes ML hard

GATE Components

- TwitIE
    - structural and linguistic pre-processing, specific to Twitter
    - includes language detection, hashtag retokenisation, POS tagging, NER
- (Optional) term recognition using TermRaider
- Sentiment gazetteer lookup
- JAPE opinion detection grammars
- (Optional) aggregation of opinions
    - includes opinion interestingness component

Basic approach for opinion finding

- Find sentiment-containing words in a linguistic relation with terms/entities (opinion-target matching)
- Dictionaries give a starting score for sentiment words
- Use a number of linguistic sub-components to deal with issues such as negatives, adverbial modification, swear words, conditionals, sarcasm etc.
- Starting from basic sentiment lookup, we then adjust the scores and polarity of the opinions via these components

A positive sentiment list

- awesome category=adjective score=0.5
- beaming category=adjective score=0.5
- becharm category=verb score=0.5
- belonging category=noun score=0.5
- benefic category=adjective score=0.5
- benevolently category=adverb score=0.5
- caring category=noun score=0.5
- charitable category=adjective score=0.5
- charm category=verb score=0.5

A negative sentiment list

Examples of phrases following the word “go”:

- down the pan
- down the drain
- to the dogs
- downhill
- pear-shaped

Opinion scoring

- Sentiment gazetteers (developed from sentiment words in WordNet) have a starting "strength" score
- These get modified by context words, e.g. adverbs, swear words, negatives and so on

  The film was awesome --> *The film was* ***really*** *amazing*.

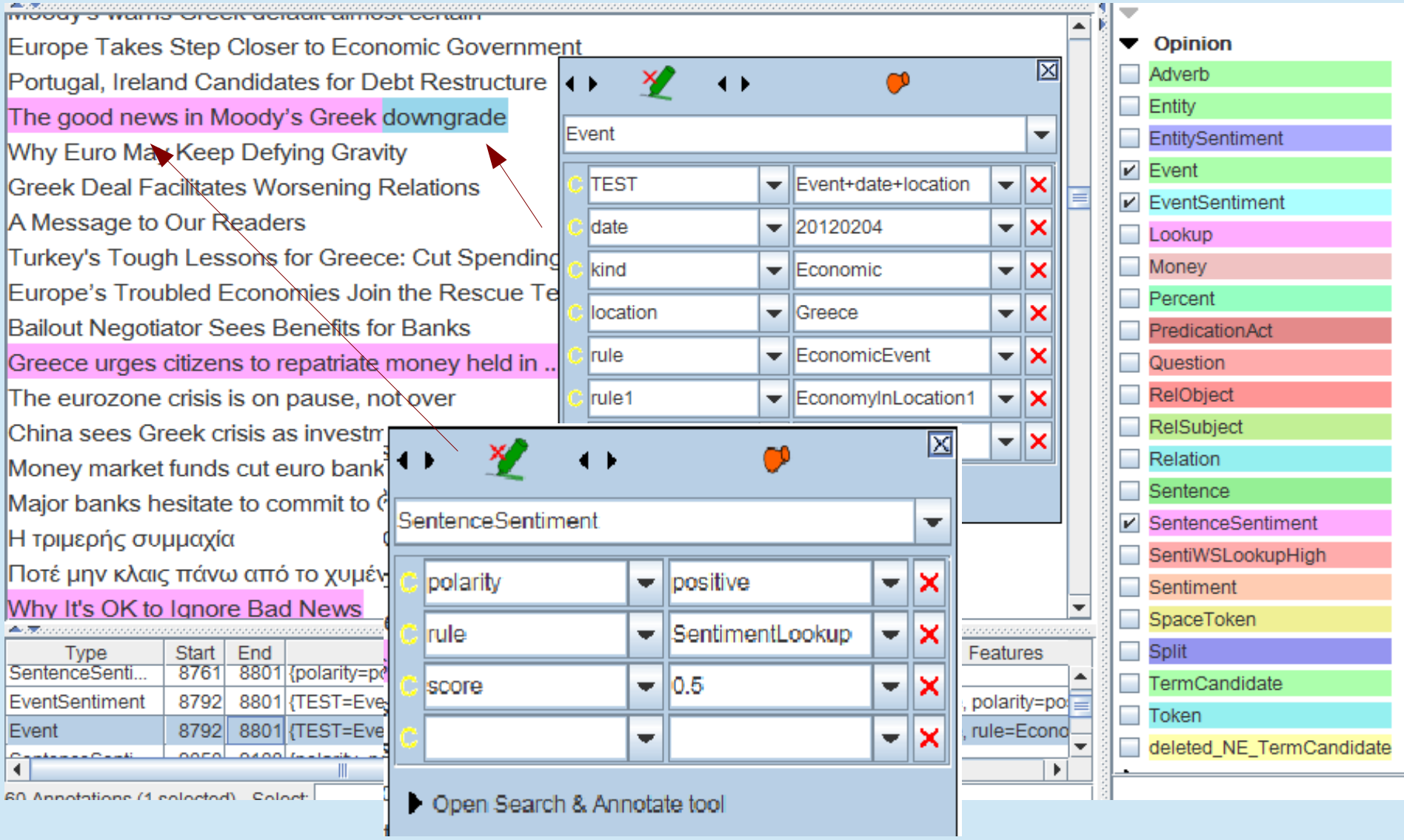
  The film was awful --> *The film was* ***absolutely*** *awful.*.

  The film was good --> The film was ***not so*** *good.*
- Swear words modifying adjectives count as intensifiers

  The film was good --> The film was ***damned*** *good.*
- Swear words on their own are classified as negative

  Damned *politicians.*

Example: Opinions on Greek Crisis



Using Machine Learning for the task

- If we can collect enough manually annotated training data, we can also use an ML approach for this task
- Product reviews: use star-based rating (but these have flaws)
- Other domains, e.g. politics: classify sentences or tweets (the ML *instances)*, many of which do not contain opinions.
- So the ML *classes* will be *positive*, *neutral and negative*
- (Some people classify *neutral* and *no opinion* as distinct classes, but we find the distinction too difficult to make reliably)

Training on tweets

- You can use hashtags as a source of classes
    - Example: collect a set of tweets with the **#angry** tag, and a set without it, and delete from the second set any tweets that look angry
    - Remove the **#angry** tag from the text in the first set (so you're not just training the ML to spot the tag)
    - You now have a corpus of manually annotated angry/non-angry data
- This approach can work well, but if you have huge datasets, you may not be able to do the manual deletions
- You can also train on things like **#sarcasm** and #irony

Summary

- Opinion mining is **hard** and therefore **error-prone** (despite what vendors will tell you about how great their product is)
- There are many types of sentiment analysis, and many different uses, each requiring a different solution
- It's very unlikely that an off-the-shelf tool will do exactly what you want, and even if it does, performance may be low
- Opinion mining tools need to be **customised** to the task and domain
- Anything that involves processing **social media** (especially messy stuff like Facebook posts and Twitter) is even harder, and likely to have lower performance
- For tasks that mainly look at aggregated data, this isn't such an issue, but for getting specific sentiment on individual posts/reviews etc, this is more problematic

So where does this leave us?

- Opinion mining is ubiquitous, but it's still far from perfect, especially on social media
- There are lots of linguistic and social quirks that fool sentiment analysis tools.
- The good news is that this means there are lots of interesting problems for us to research
- And it doesn’t mean we shouldn’t use existing opinion mining tools
- The benefits of a modular approach mean that we can pick the bits that are most useful
- Take-away message: use the right tool for the right job

More information

- GATE http://gate.ac.uk (general info, download, tutorials, demos, references etc)
- The EU-funded Decarbonet and TrendMiner projects are dealing with lots of issues about opinion and trend mining from social media
    - http://www.decarbonet.eu
    - http://www.trendminer-project.eu/
- Tutorials
    - Module 12 of the annual GATE training course: “Opinion Mining” (2013 version) http://gate.ac.uk/wiki/TrainingCourseJune2013/
    - Module 14 of the annual GATE training course: “GATE for social media mining”

Some GATE-related opinion mining papers

(available from http://gate.ac.uk/gate/doc/papers.html)

- Diana Maynard, Gerhard Gossen, Marco Fisichella, Adam Funk. Should I care about your opinion? Detection of opinion interestingness and dynamics in social media. To appear in Journal of Future Internet, Special Issue on Archiving Community Memories, 2014.
- Diana Maynard and Mark A. Greenwood. Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis. Proc. of LREC 2014, Reykjavik, Iceland, May 2014.
- Diana Maynard, David Dupplaw, Jonathon Hare. Multimodal Sentiment Analysis of Social Media. Proc. of BCS SGAI Workshop on Social Media Analysis, Dec 2013
- D. Maynard and K. Bontcheva and D. Rout. Challenges in developing opinion mining tools for social media. In Proceedings of @NLP can u tag #usergeneratedcontent?! Workshop at LREC 2012, May 2012, Istanbul, Turkey.
- D. Maynard and A. Funk. Automatic detection of political opinions in tweets. In Raúl García-Castro, Dieter Fensel and Grigoris Antoniou (eds.) The Semantic Web: ESWC 2011 Selected Workshop Papers, Lecture Notes in Computer Science, Springer, 2011.
- H.Saggion, A.Funk: Extracting Opinions and Facts for Business Intelligence. Revue des Nouvelles Technologies de l'Information (RNTI), no. E-17 pp119-146; 2009.

Some demos to try

- http://sentiment.christopherpotts.net/lexicon/

  Get sentiment scores for single words from a variety of sentiment lexicons
- http://sentiment.christopherpotts.net/textscores/

  Show how a variety of lexicons score novel texts
- http://sentiment.christopherpotts.net/classify/

  Classify tweets according to various probabilistic classifier models
- http://demos.gate.ac.uk/arcomem/opinions/

  Find and classify opinionated text, using GATE-based ARCOMEM system

Questions?