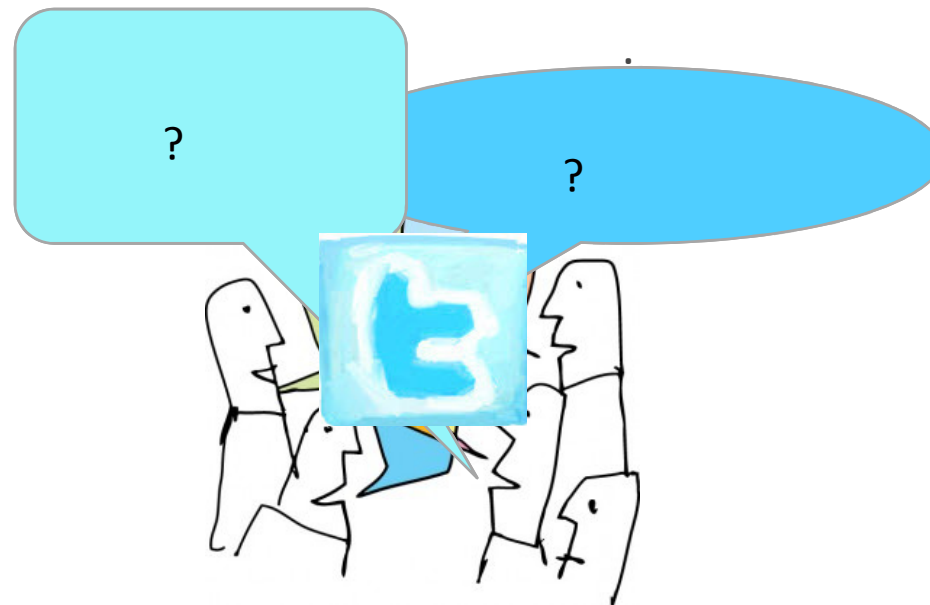


Echolytics

Social Media Analytics



Overview

- Project
- Results
- Future Work

Project

Goal

- Using social computing and sentiment analysis techniques to retrieve, organize, and analyze people's judgments about social and environmental activities of companies according to their tweets.
- Using the outcome of such effort to rate businesses according to people's perception rather than objective criteria.

A horizontal header bar with a dark grey, textured background. The letters 'CSR' are written in white, sans-serif font on the right side. Below the header bar, there are three colored segments: a blue segment on the left, a cyan segment in the middle, and a green segment on the right.

CSR

➤ What is CSR and why is it important?

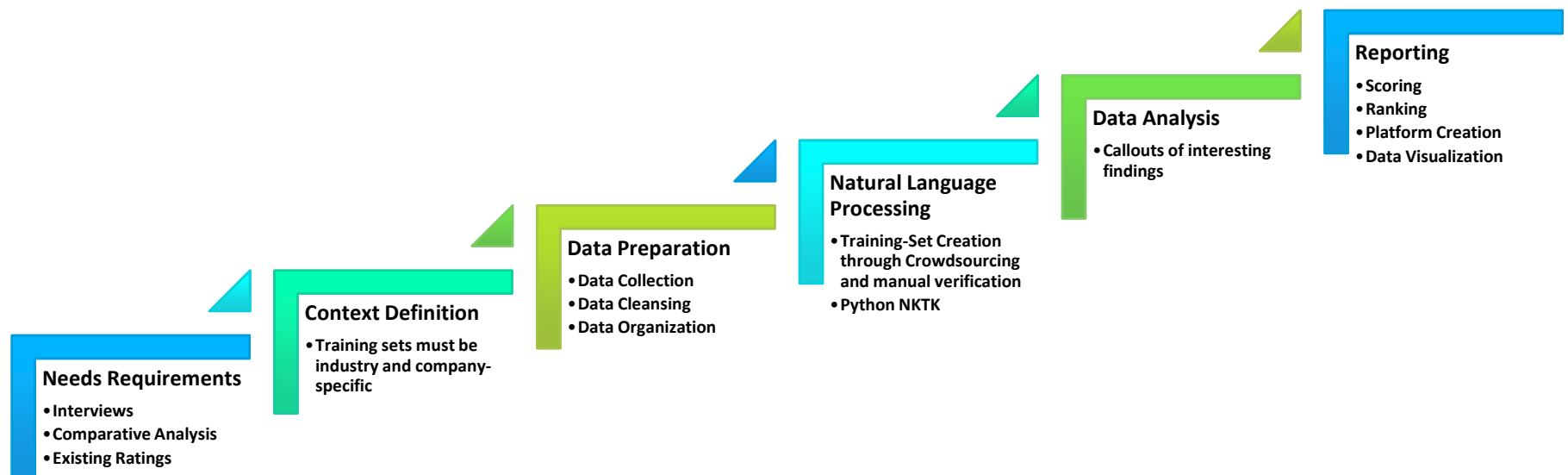
Data Source

- Twitter (Phase 1)
 - Jan 2012 – Feb 2012 *
 - 6 companies in 2 industries
- Facebook
- Blogs (news, experts, discussions, etc.)



Methodology

Computing “Goodness” from Social Media Data



Steps Taken

- Data collection and meta-data appending
- Data cleansing and organization
- Define CSR (by crowdsourcing)
- Define contexts (industry and company-specific, sentiment – by crowdsourcing)
- Training Set Creation for NLP Classifier

Steps Taken

- NLP classification on tweets
 - Refers to company or not (yes/no/undetermined)
 - Sentiment (+/-/0/undetermined/mixed)
 - Relates to CSR (yes/no/undetermined)
- NLP to identify Top Terms (words and bigrams, POS/chunking) and Top Hashtags
- Manual review of data
- Visualizations and website creation

Classifying Tweets

Tweets:

Bout to be so fat #panera

@laurroggg You're on drugs if you think #panera makes you feel fat!

I love Chipotle more than McDonald's

Rate that tweet!

- Refers to company?
- CSR related?
- Sentiment?

Rate That Tweet!

Tweet:

I have no more chipotle chips :(



Rate that tweet!

- Refers to company?
- CSR related?
- Sentiment?



Rate That Tweet!

Tweet:

Clorox makes me ITCH

Rate that tweet!

- Refers to company?
- CSR related?
- Sentiment?

Rate That Tweet!

Tweet:

I would eat elsewhere but #Panera lists
calorie counts #healthierating

Rate that tweet!

- Refers to company?
- Is CSR related?
- Sentiment?

Rate That Tweet!

Tweet:

RT @HyperBrendan: RT @abcnews: Jobs go as iconic brands Mortein & Dettol stop being made in Australia from July
<http://t.co/y1Xbj8v5> || Libs #blamecarbontax

Rate that tweet!

- Refers to company?
- CSR related?
- Sentiment?

Rate That Tweet!

Tweet:

Maybe if Panera wasn't so damn delicious, I would 1) have more money and 2) consume half as many carbs.
#paneraisafoodgroup

Rate that tweet!

- Refers to company?
- Is CSR related?
- Sentiment?

Rate That Tweet

Tweet:

So tired just got done colorix ing the bathroom and I put Lysol together well that made a bomb and I I can't breath my lungs r clogged lol

Rate that tweet!

- Refers to company?
- Is CSR related?
- Sentiment?

Rate That Tweet!

Tweet:

McDonald's to Stop Using 'Pink Slime' in
Hamburgers in US.

Rate that tweet!

- Refers to company?
- Is CSR related?
- Sentiment?

Results

Reporting Platform Prototype

Echolytics analyzes tweets, eliminates noise, and ranks companies according to the public's perception of them.

Top Do-Good Companies

Rank	Company Name	Goodness	CSR Buzz	Details	Echo Score
1	Panera			Preview Full Report	90
2	McDonalds			Preview Full Report	87
3	Chipotle			Preview Full Report	76

Imperfect Science

➤ Results: Feasible but not an exact science

Operationalization Choices

- Using Twitter and only Twitter
- Which companies are chosen and which industries affect everyone's scores (future all Top 100)
- Date range
- Company aliases

Procedures Used

- Data collection – English
- Data cleansing - unicode characters and emoticons
- CSR definition – depends on audience
- Training set must be larger because CSR tweets are only a small percent of all tweets
- Manual oversight of NLP program to identify Top Terms and Top Hashtags

Human Intelligence Needs

- We gathered all tweets contained these companies names (and a few **aliases**) **between** 1/1/12 and 2/29/12
- **Define CSR** (by crowdsourcing)
- NLP **classification** on tweets
 - Refers to company or not (yes/no/undetermined)
 - Sentiment (+/-/0/undetermined/mixed)
 - Relates to CSR (yes/no/undetermined)

Future Work

Post-Experiment Steps

- Needs Assessment
- Usability Testing
- Further automation of automatable pieces
- Guidelines/tips for human intel procedures
- Training Set supervision

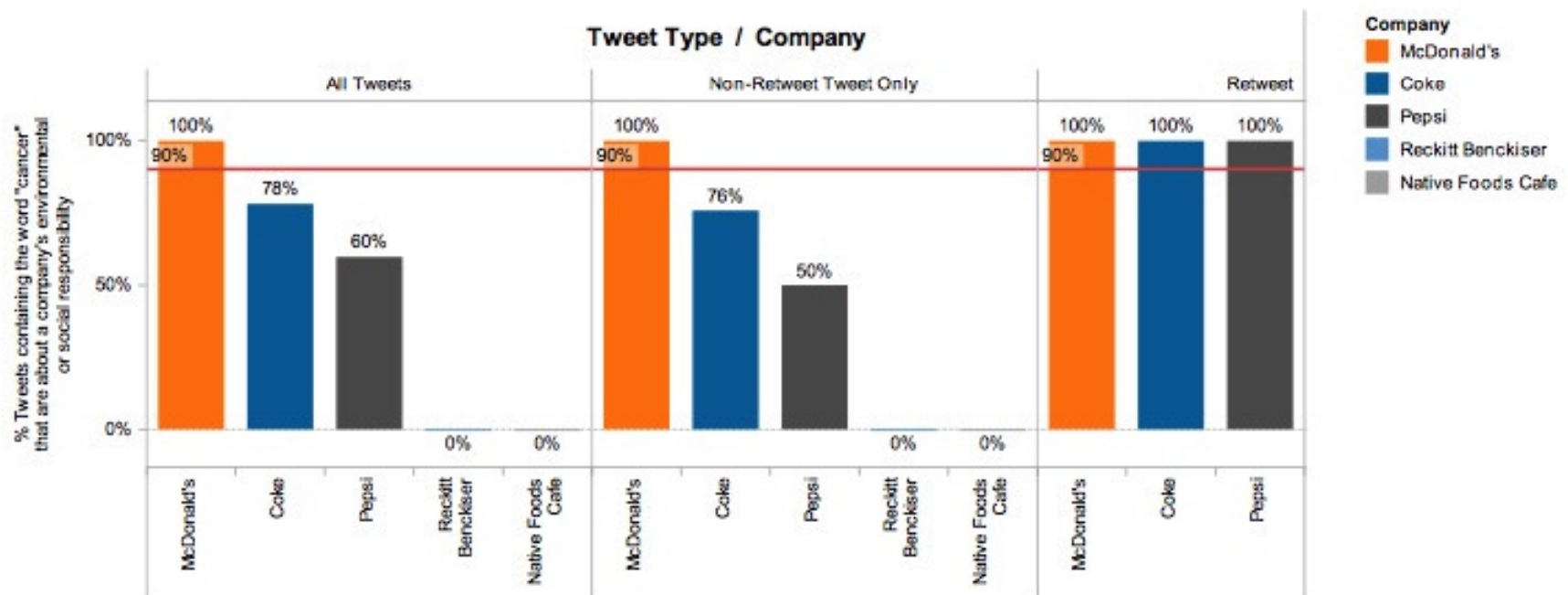
Additional Features

- Many other NLP features would be possible
 - more sophisticated topics identification, topic comparisons and tracking, trends, “viral”ity,/amount and speed of spread
 - user analyses
 - Geography
 - indiv vs company
 - Influencers
 - speech patterns-user characteristics mapping
 - separate retweets from non-retweets for all reports

Context Matters

Identifying Tweets about environmental or social responsibility:

Among Tweets that contain certain company names, does also containing word "cancer" predict (at least 90% of the time) that the Tweet is about the company's environmental or social responsibility?



Audience

- Investors (selected for Phase 1)
- Consumers
- Consumer Influencers
- Watchdog Entities
- Corporations

Questions?