UNIVERSITY OF TWENTE.

#Datagrant project: presentation UT text analysis café

Dong Nguyen & Tijs van den Broek

2-6-2016

●●●○○ T-Mobile NL 14:21 84%

The Washington Post (WP Company Llc)

Sections The Washington Post

On I.T.

Twitter grants select researchers access to its public database

A 0

A sign displays the Twitter logo on the front of the New York Stock Exchange ahead of the company's IPO in New York in

February 4, 2014

Twitter announces to allow a handful of researchers access to its historical data

Cancer Early Detection Campaigns on Twitter

Annual Cancer Deaths	
Per Year	6.202.080
Per Month	516.840
Per Day	16.992
Per Hour	708
Per Minute	12

Cancer Early Detection Campaigns on Twitter

Annual Cancer Deaths	
Per Year	6.202.080
Per Month	516.840
Per Day	16.992
Per Hour	708
Per Minute	12

One third of these deaths could have been **prevented**

How Cancer Deaths can be Prevented

1. Doing research
2. Regular screening and early diagnosis
3. Vaccination
4. Lifestyle

How Cancer Deaths can be Preven

1. Doing research

2. Regular screening and ea
3. Vaccination
4. Lifestyle

How Cancer Deaths can be Preven

1. Doing research
2. Regular screening and early diagnosis
3. Vaccination
4. Lifestyle

How Cancer Deaths can be Preven

1. Doing research
2. Regular screening and early diagnosis
3. Vaccination
4. Lifestyle

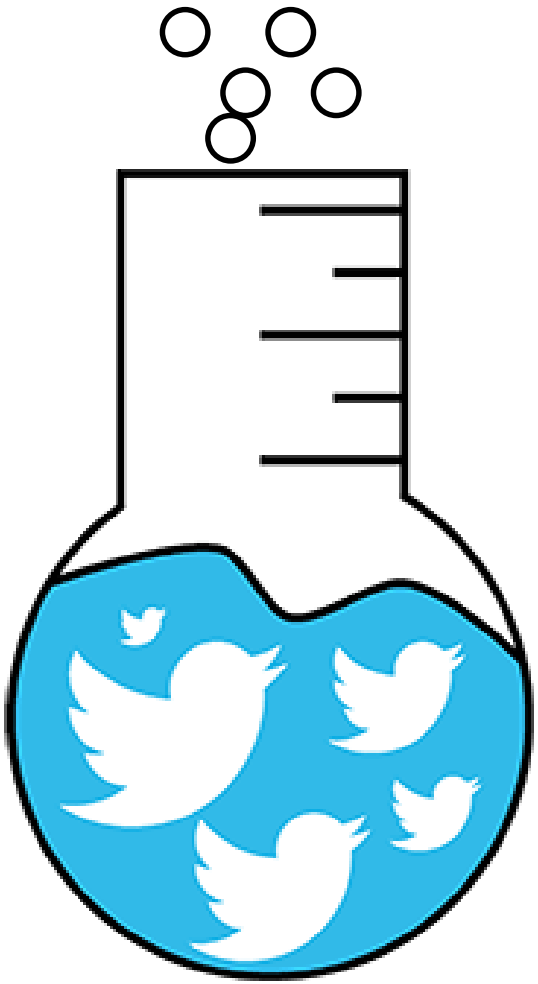
How Cancer Deaths can be Preven

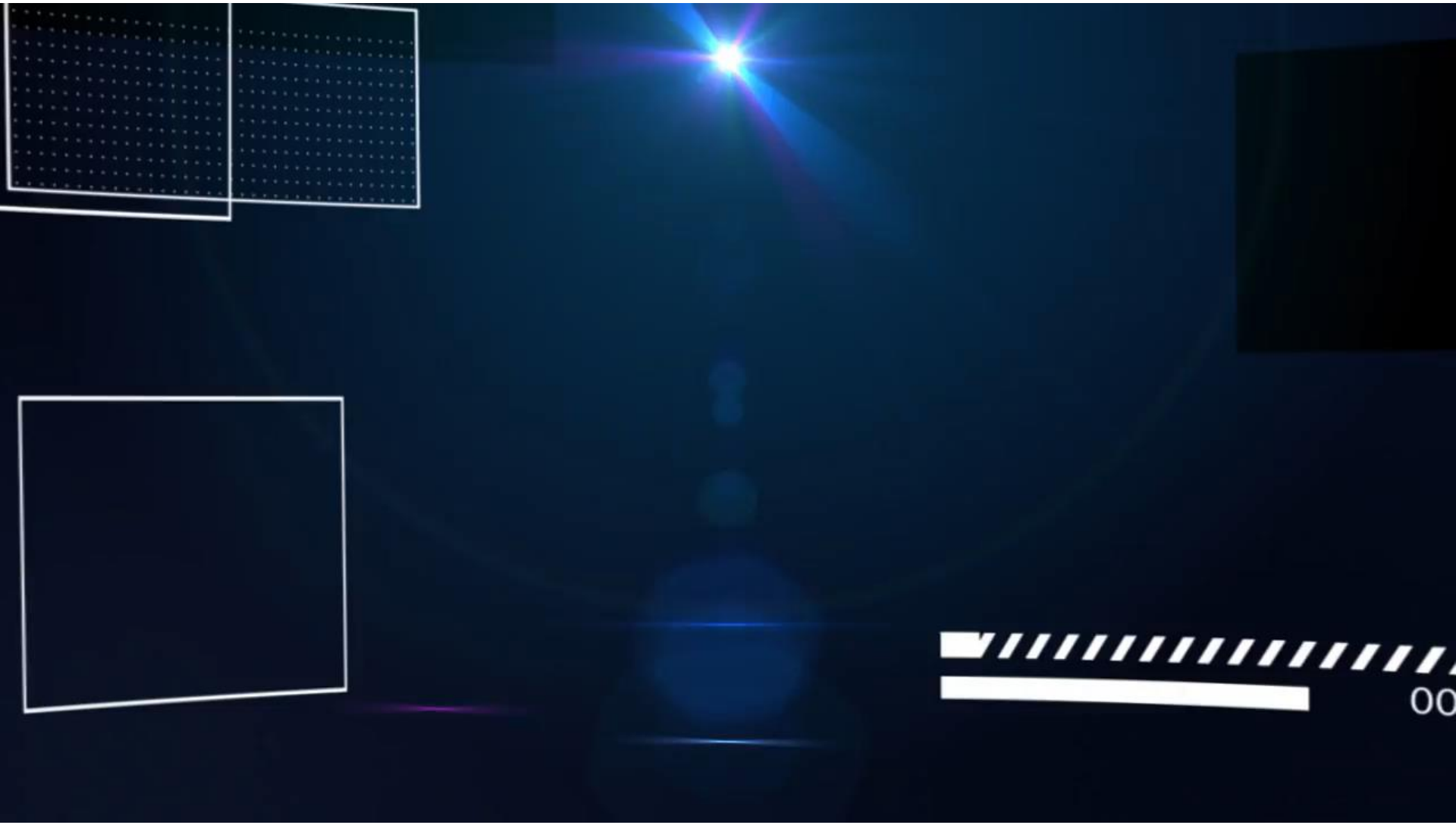
1. Doing research
2. Regular screening and early diagnosis
3. Vaccination
4. Lifestyle

Twitter#DataGrant from Twent University

The data grant proposal aims to study the diffusion process and effectiveness of cancer early detection campaigns on Twitter

Twitter #DataGrant



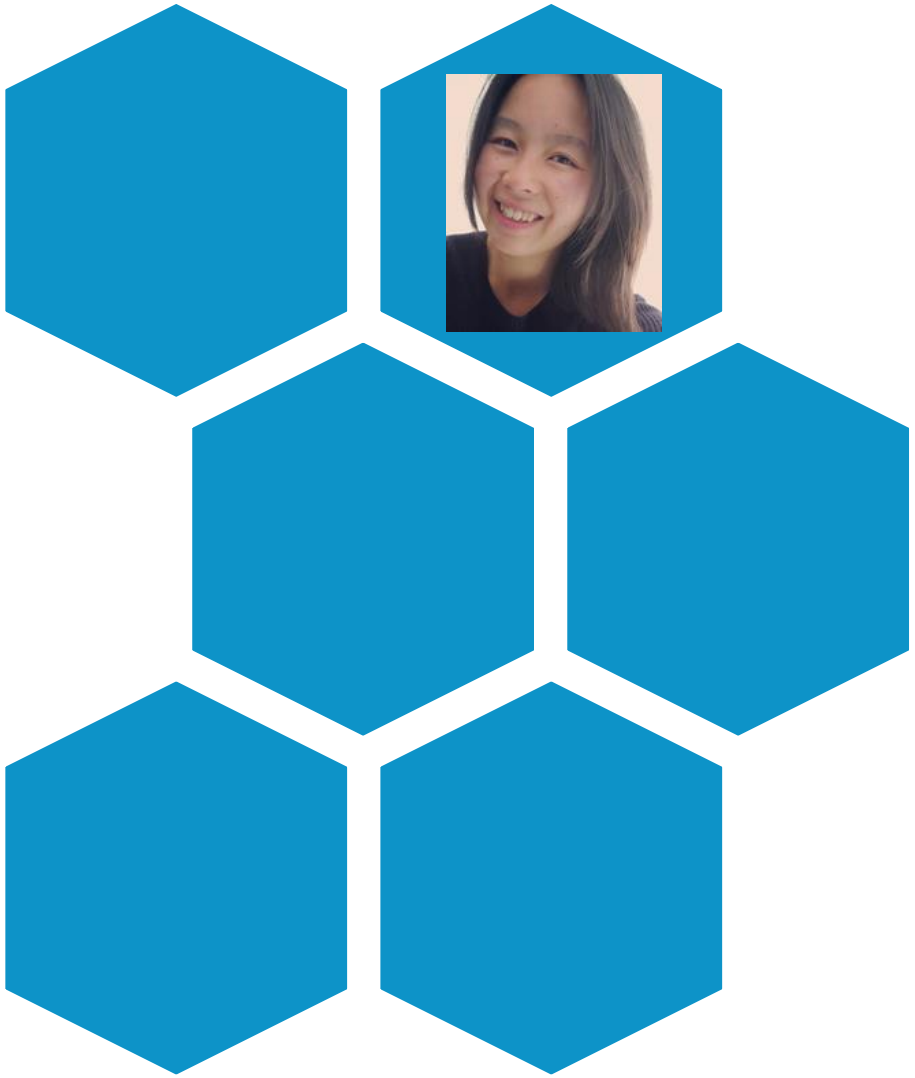


Results

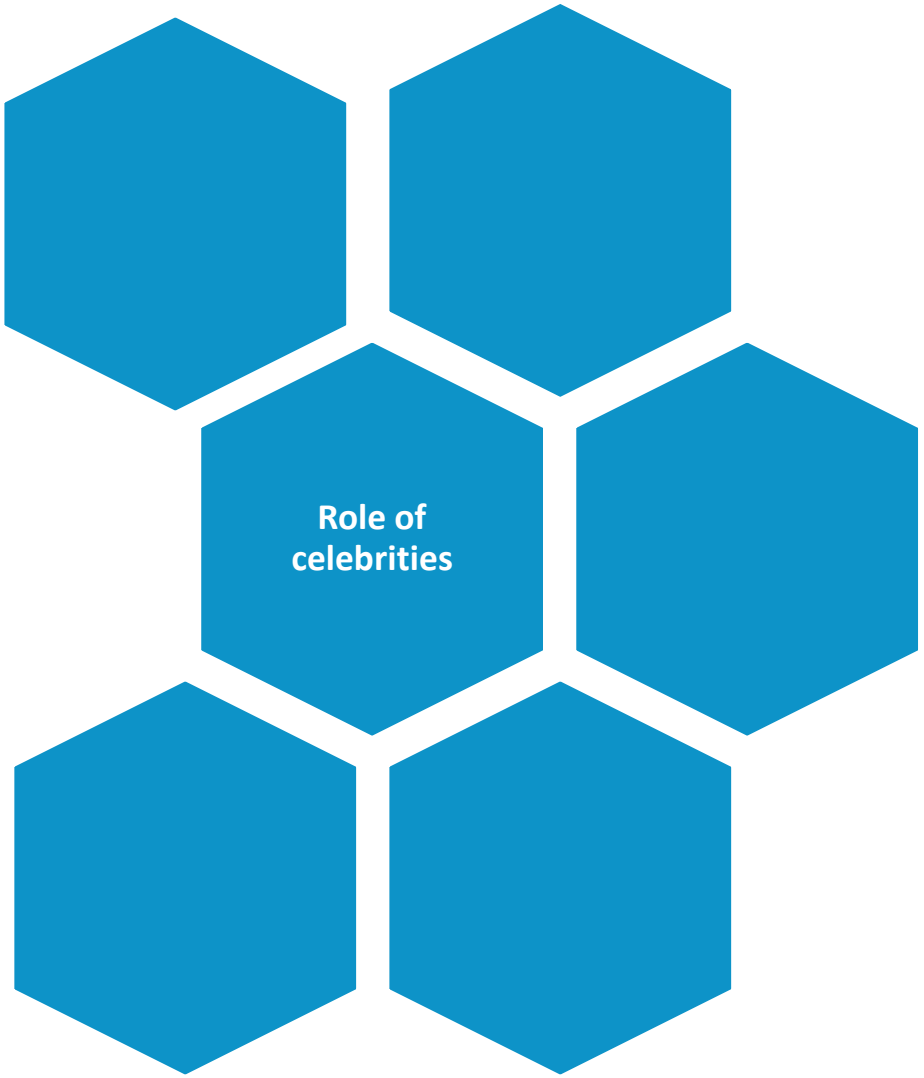
Classifying Countries

Results

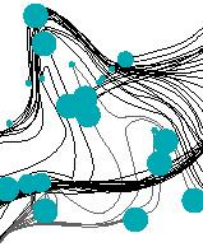
Classifying Motivations

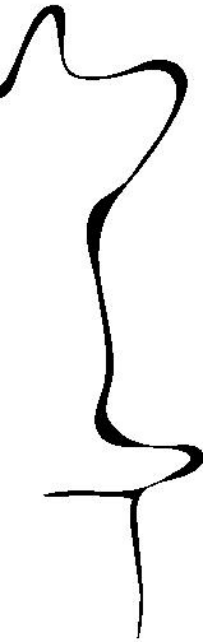


Results

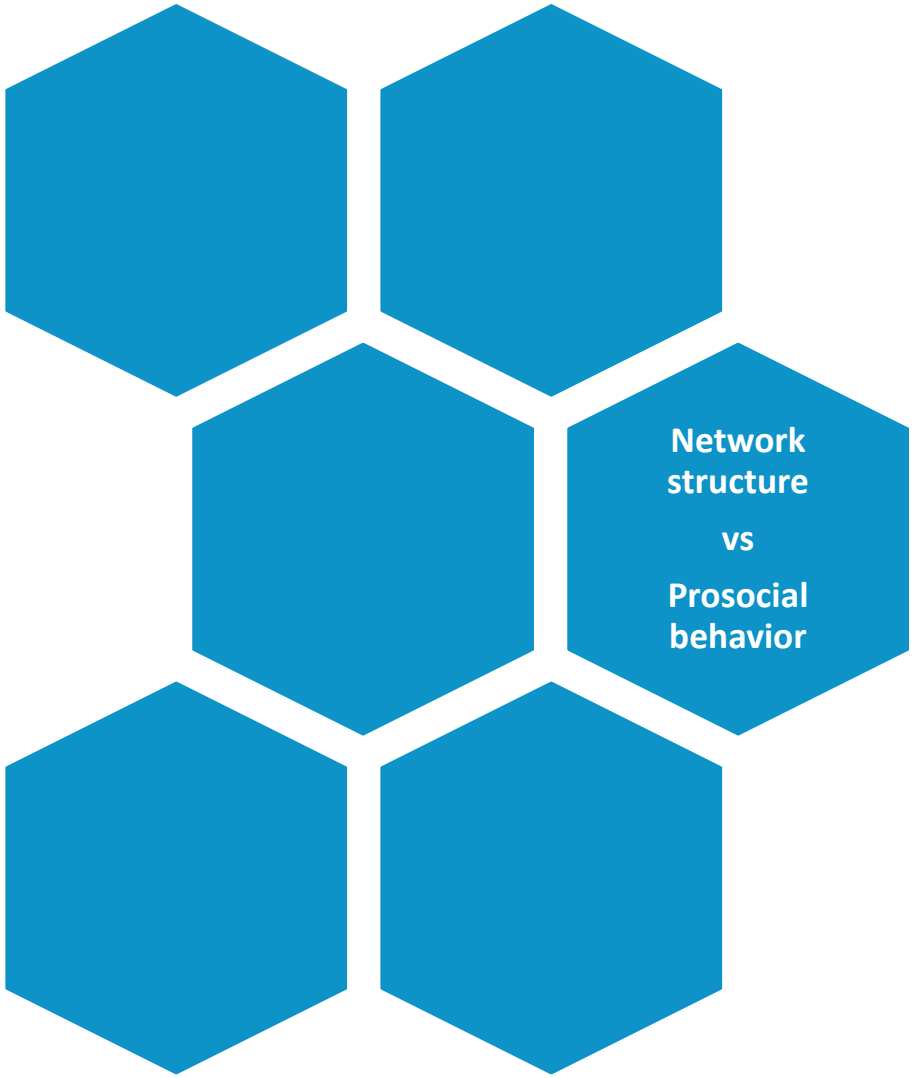


Role of Celebrities





Results



Network Structure vs. Altruism

Results

Role of Identity

Work in progress

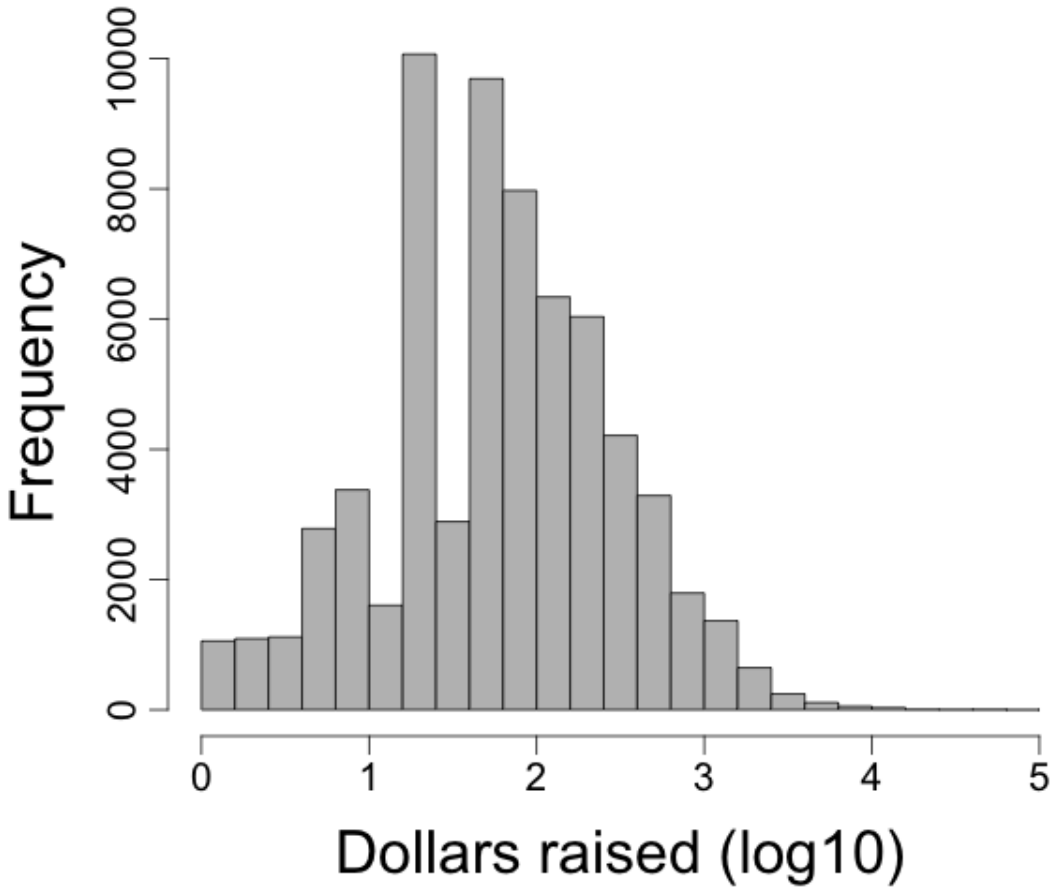
- Impact of tweets and offline fundraising events on fundraising performance
- Multilevel study in US counties (linking data about ratio prostate and testicular cancer)
- Motivation of campaign members over time
- The effectiveness of Movember teams and their leaders

Identifying Motivations to Participate in Online Health Campaigns

D. Nguyen, T. van den Broek, C. Hauff, D. Hiemstra and M. Ehrenhard: #SupportTheCause: Identifying Motivations to Participate in Online Health Campaigns at EMNLP 2015.

166,422 US participants in

100,615 raised **no** money



Knowing individual motivations helps to explain the amount of campaign donations raised by participants

Social Identity Model of Collective Action (van Zomeren et al., 2008)

- Injustice: A shared emotion that includes both affective (e.g., anger) and cognitive perceptions (ideology) of an unfair situation

Social Identity Model of Collective Action (van Zomeren et al., 2008)

- Injustice: A shared emotion that includes both affective (e.g., anger) and cognitive perceptions (ideology) of an unfair situation

"I had testicular cancer"

"my dad"

"because men's health is important to me"

Social Identity Model of Collective Action (van Zomeren et al., 2008)

- Injustice: A shared emotion that includes both affective (e.g., anger) and cognitive perceptions (ideology) of an unfair situation
- Social identity: A sense of belonging together that emerges out of common attributes, experiences and external labels

Social Identity Model of Collective Action (van Zomeren et al., 2008)

- Injustice: A shared emotion that includes both affective (e.g., anger) and cognitive perceptions (ideology) of an unfair situation
- Social identity: A sense of belonging together that emerges out of common attributes, experiences and external labels

"my friends asked me again to join them"

"a great excuse to grow a stache"

Social Identity Model of Collective Action (van Zomeren et al., 2008)

- Injustice: A shared emotion that includes both affective (e.g., anger) and cognitive perceptions (ideology) of an unfair situation
- Social identity: A sense of belonging together that emerges out of common attributes, experiences and external labels
- Collective efficacy: The shared belief that ones group is capable of resolving its grievances through a campaign

Social Identity Model of Collective Action (van Zomeren et al., 2008)

- Injustice: A shared emotion that includes both affective (e.g., anger) and cognitive perceptions (ideology) of an unfair situation
- Social identity: A sense of belonging together that emerges out of common attributes, experiences and external labels
- Collective efficacy: The shared belief that ones group is capable of resolving its grievances through a campaign

"this campaign can make a difference!"

Nick Wheeler

My Motivation

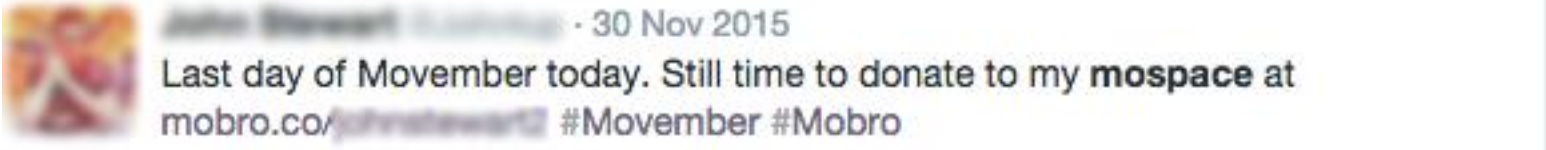
Because I have just turned 50 and have become aware how many people this disease affects. Selling shirts is not enough. It is more important to beat this terrible cancer.

- Nick Wheeler

1

FIRST YEAR MO!

Linking profiles



In total 5,519 users linked. 2,108 were manually annotated for their motivation.

Dataset statistics

	Train	Test
# Participants	1,494	614
% US / UK	54.8/45.2	53.3/46.7
% Injustice	37.6	40.2
% Social identity	48.7	46.9
% Collective efficacy	36.1	35.0

Annotation agreement

	Cohen's Kappa
Injustice	0.71
Social identity	0.67
Collective efficacy	0.47

Based on 200 double annotations

Automatic classification of Movember profiles

- Logistic Regression
- Unigrams, bigrams, topics, text length, country

	F1
Injustice	0.816
Social Identity	0.788
Collective efficacy	0.627

Final system

Feature analysis

Injustice	Social Identity	Collective Efficacy
LDA topic[a]	fun	LDA topic[b]
cancer	team	beat
friend	moustache	and family
lost	mo	change
father	grow	yourself
had	mustache	all of
survivor	LDA topic[c]	awareness
prostrate	fuzz	for movember
for my	movement	awareness of
my	look	last

Table 4: Top-weighted features for free-text motivation experiments.

[a] topic about family/friends who had cancer
[b] topic about raising funds for research
[c] topic about the Movember campaign

Automatic classification of Twitter profiles

- Logistic regression
- Unigrams, bigrams, topics, text length, country, behavior, urls, user mentions,..

Low performance, why?

- Few tweets per user
- Nature of Twitter influences content

	F1
Injustice	0.458
Social Identity	0.531
Collective efficacy	0.399

Final system

Findings

- Campaign participants with an **injustice** motivation raise significantly ($p < 0.001$) more money
- Participants that are part of a **team** raise significantly more money ($p < 0.001$)
- Participants with a **social identity** motivation are more often part of a team

	Injustice	Identity	Efficacy
UK ($)	203.74	128.36	123.39
US ($)	234.47	156.07	169.03

n=90,484

Summary

- Explored machine learning methods to automatically identify the motivations of Movember participants

- We found a strong link between motivations and donations, and motivations and team membership

Thank you!

D. Nguyen, T. van den Broek, C. Hauff, D. Hiemstra and M. Ehrenhard: #SupportTheCause: Identifying Motivations to Participate in Online Health Campaigns at EMNLP 2015.