TWITTER CANCER AWARENESS CAMPAIGNS AS #IDENTITYPROJECTS FOR PREVENTION AND EARLY DETECTION BEHAVIOR

Lunch Lecture Inter-Actief – April, 4th 2017

@AnnaPriante – a.priante@utwente.nl
Agenda

- Something about me
- PhD project
- Social Identity Classifier and applications
- Q&A
#WholAm in 160 characters and some more

Anna Priante
@AnnaPriante
PhD researcher on online social movements @UTwente | Topic: Twitter cancer awareness campaigns, online identities and offline behaviors | Musician | Foodblogger
TWITTER CANCER AWARENESS CAMPAIGNS AS #IDENTITYPROJECTS
A multi-level study of the influence of online identities on offline behaviors
Anna Priante

Mijn motivatie:
To raise awareness about the importance of healthy behaviors and early detection to prevent cancer

MOVE-update
8 WEKEN GELEDEN
Ik heb gefietst 10 kilometer

DELEN VIA FACEBOOK  VERWIJDEREN  AANPASSEN  VERSTOPEN

MOVE-update
8 WEKEN GELEDEN
Ik ben actief geweest gedurende 1 uur

DELEN VIA FACEBOOK  VERWIJDEREN  AANPASSEN  VERSTOPEN

Teamcaptain
TIJJS VAN DEN BROEK

Post een teamupdate.
#MOVE DAY 2 Every day I go to work by bike. Just 10 km per day but it is good to move movember modatagranttwente

Anna Priante
@AnnaPriante

29 November 2015 ·

SupportTheCause: I'm helping Movember change the face of men's health with my donation. Join me in supporting this important cause! http://mosista.co/annapriante

Anna Priante

28 November 2015 ·

Ultimo giorno, last run of the day! Join me and support the cause!
ARE THESE ONLINE CAMPAIGNS EFFECTIVE?
UNDER WHICH CONDITIONS ARE THEY EFFECTIVE?
TWITTER CANCER AWARENESS CAMPAIGNS AS #IDENTITYPROJECTS
A multi-level study of the influence of online identities on offline behaviors
Effectiveness of online campaigns

Online (low-threshold) behavior

Meaningful offline behavior
How can we study campaigns effectiveness?

Conceptualize, explore and test how online identities explain the effectiveness of online campaigns.

UNIVERSITY OF TWENTE.
Identities on Twitter

Profile & Descriptions

Anna Priante
@AnnaPriante
PhD researcher on online social movements @TUtwente | Research & Twitter cancer awareness campaigns and offline behaviors | Foodblogger | Musician.

Tjjs van den Broek
@tyjsvandbroek ti segue
Social scientist & engineer at TNO | University of Twente | Twitter Da awarder - Home brewer

Michel Ehrenhard
@mythiodae ti segue
Cinderzoeker @UTwiene | disruptieve verandering, nieuwe organisatiemakings, ondernemen leidinggeven, decision making | ICT | Twitter @datagram | voetbal | hiking

Hashhtags

#movember #modatagranttwente #moustache #menshealth

Tweets & retweets

#MOVE DAY 2 Every day I go to work by bike. Just 10 km per day but it is good to move #movember #modatagranttwente

Don't leave a MO Behind
SOCIAL SCIENCES
MEET
BIG DATA

UNIVERSITY OF TWENTE.
‘Big social data’: all tweets from 9 campaigns and 6 cancer types since 2008

> 300 million Tweets
#WhoAmI IN 160 CHARACTERS?
CLASSIFYING SOCIAL IDENTITIES BASED ON
TWITTER PROFILE DESCRIPTIONS

A. Priante, D. Hiemstra, T. van den Broek, A. Saeed, M. Ehrenhard, A. Need


Main contributions:
1. Social theory can be used to guide NLP methods, and NLP methods provide input to revisit social theory
2. Three social identity classifiers of Twitter profile descriptions (English) for scaling-up online identity research to massive datasets
From language to identity and behavior

LANGUAGE

Social Identity

Who am I?

Behaviors

Anna Priante
@AnnaPriante
PhD researcher on online social movements @UTwente | Topic: Twitter cancer awareness campaigns, online identities and offline behaviors | Musician | Foodblogger
Social Theory

- **SOCIAL IDENTITY**: individual’s self-concept derived from social roles or memberships to social groups

**5-Category Online Social Identity Classification**

- **Relational**
  - All I Can Say Is I Love Justin bieber, if Ur A Belieber follow me an i'll follow u. justin bieber if u followed me i would scream my head off and faint love u
  - Husband to a great Wife, Father to a great Daughter

- **Occupational**
  - Community Manager and vidjagame enthusiast. I also write jokes.
  - IT | Gamer | Fitness & Foodie | GT86/FR-S Owner

- **Political**
  - #feminist #activist
  - Cronulla Woman of the Year 2015. It’s time to make #pancreaticcancer MATTER
  - Volunteer nicoleltzsimons.com - Friend of @thedesmondtutu

- **Ethnic/Religious**
  - Loves being an Englishman
  - Follower of Jesus Christ

- **Stigmatized**
  - gay white gamer geek male.
  - geek • insomniac • workaholic
Data & Methods

- Movember campaign
- Random tweets
- From UK, US & AUS (country classifier)

Profile description

N = 2,633

- Human Annotation (2 coders)
- Codebook
- Inter-rater reliability (KALPHA) > 0.850

Qualitative Content Analysis

Automatic Text Classification

- Binary text classification
- TF-IDF weighting & ChiSquare for FS
- 4 classification algorithms (SVM, BNB, LR, RF)
- Stratified 10-Fold Cross Validation
- Evaluation measures: precision (P), recall (R) and F-score.
Experiment 1: PARTIALLY WORKS!

- 5-category online social identity classification
- Stratified 10F Cross Validation on 3 training sets
- Optimal classification algorithm: **Logistic Regression** (evaluation on 3 test sets)
Experiment 1: PARTIALLY WORKS!

- 5-category online social identity classification
- Stratified 10F Cross Validation on 3 training sets
- Optimal classification algorithm: **Logistic Regression** (evaluation on 3 test sets)

<table>
<thead>
<tr>
<th>RELATIONAL</th>
<th>OCCUPATIONAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>R</td>
</tr>
<tr>
<td>0.757</td>
<td>0.648</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>POLITICAL</th>
<th>ETHNIC/RELIGIOUS</th>
<th>STIGMATIZED</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>R</td>
<td>F</td>
</tr>
<tr>
<td>0.600</td>
<td>0.200</td>
<td>0.300</td>
</tr>
</tbody>
</table>
REVISITING SOCIAL THEORY

- Political, Ethnic/Religious and Stigmatized identities are more engaged in collective action
- They have a collective, action-oriented nature and may often overlap
- We merge Political, Ethnic/Religious and Stigmatized identities (PES identity)

Distributions of social identities on Twitter (based on the annotated sample).

- Relational: 28%
- Occupational: 36%
- Political: 4%
- Ethnic/Religious: 13%
- Stigmatized: 7%
- N/a: 45%
Experiment 2: WORKS!

- 3-identity category merger (PES identity)
- Stratified 10F Cross Validation on Combined(1) training set
- Optimal classification algorithm: Logistic Regression (evaluation on Combined(2) test set)

<table>
<thead>
<tr>
<th>PES</th>
<th>P</th>
<th>R</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.857</td>
<td>0.466</td>
<td>0.604</td>
</tr>
</tbody>
</table>

We succeeded in developing 3 social identity classifiers for relational, occupational and PES identities!

Social theory can be used to guide NLP methods, and NLP methods provide input to revisit social theory
WHAT’S NEXT?

THE ROLE OF OCCUPATIONAL IDENTITY AND COMMUNICATION NETWORKS IN CHARITABLE GIVING DURING ONLINE CAMPAIGNS
Occupational Identity (Classifier) + Charitable Giving

Communication Network

Frequency of Tweets

Structural position in the network
Practical contributions for Movember

- Campaign organizers should consider to target people with occupational identities because they are likely to engage in charitable giving.
- Limit behavioral choices to those relevant to charitable giving by lightening the intensity of Twitter activity in favor of the creation of communicative networks, social identification, interconnectedness and interdependence.
THANK YOU FOR YOUR ATTENTION!

QUESTIONS?

a.priante@utwente.nl
@AnnaPriante