

#filterbubble

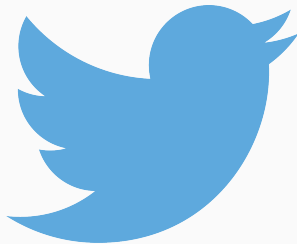
Františka Sandroni

Jakub Dostál

Filter bubble

Izolace od dostatečně širokého spektra informací.

- social networks
- preferential algorithms
- Eli Pariser (2011)



Filter bubble

Issues:

- content **homogeneity**
- objectivness loss
- **radicalization**

Goals:

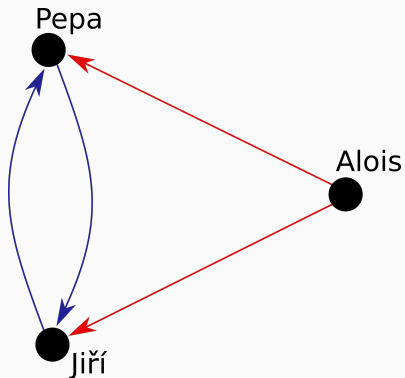
- Method for filter bubble analysis
- quantitative approach
- filter bubble **detection**

Method

1. data collection
2. studied groups selection
3. content affecting studied people
4. posts on given topic
5. sentimental analysis
6. filter bubble measures

Twitter

- social network
- news channel
- following, followers
- Twitter API

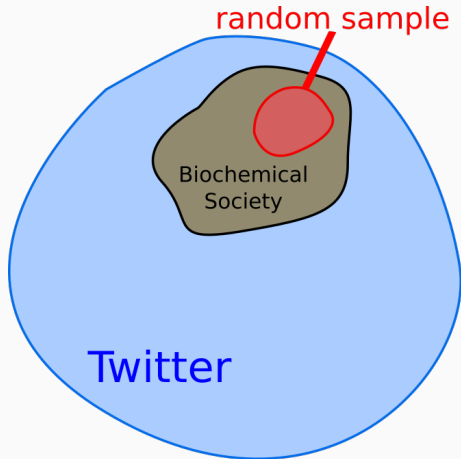


Method

1. data collection
2. **studied groups selection**
3. content affecting studied people
4. posts on given topic
5. sentimental analysis
6. filter bubble measures

Studied groups selection

- studied users
- **corresponding community**
- random sample from followers



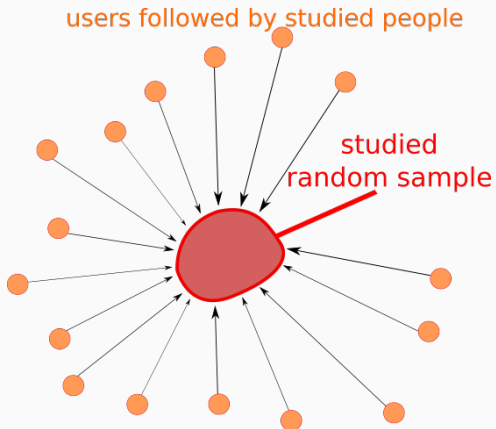
Twitter → Biochemical Society → **studied people**

Method

1. data collection
2. studied groups selection
3. **content affecting studied people**
4. posts on given topic
5. sentimental analysis
6. filter bubble measures

Tweets collection

- content affecting studied people
- content from **followed people**
- mutual tweets



Method

1. data collection
2. studied groups selection
3. content affecting studied people
4. **posts on given topic**
5. sentimental analysis
6. filter bubble measures

Tweets filtering

- posts with given keywords
- keyword: *Trump*

I had fish and chips for lunch. ✗

I'm glad Donald **Trump** is the president of the USA. ✓

The president of the USA is a gentleman. ✗

Method

1. data collection
2. studied groups selection
3. content affecting studied people
4. posts on given topic
5. **sentimental analysis**
6. filter bubble measures

Sentimental analysis

- positive vs. negative tweets
- classification
- machine learning, big data

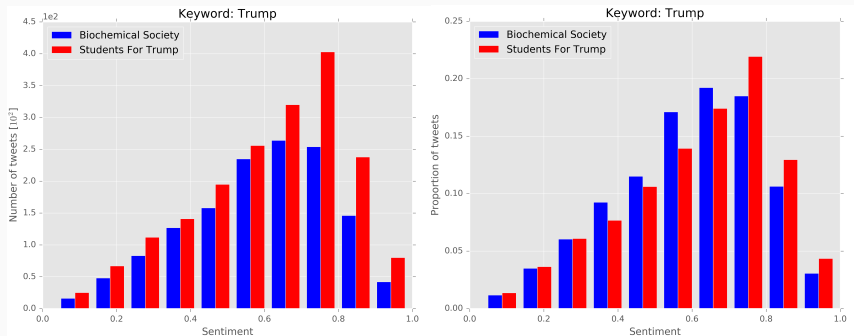
Donald Trump is a terrible person.
(0.14)

Donald Trump is a great person.
(0.95)

Method

1. data collection
2. studied groups selection
3. content affecting studied people
4. posts on given topic
5. sentimental analysis
6. **filter bubble measures**

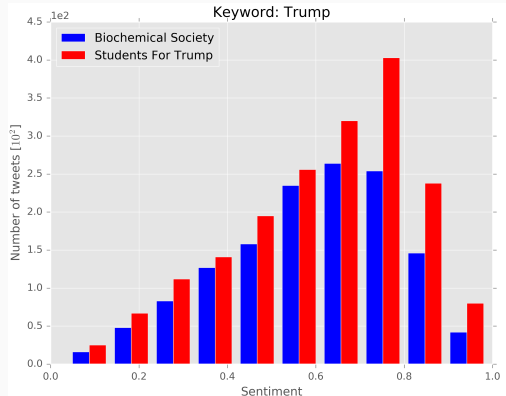
Keyword: Trump



Number of tweets histogram, keyword: "Trump".

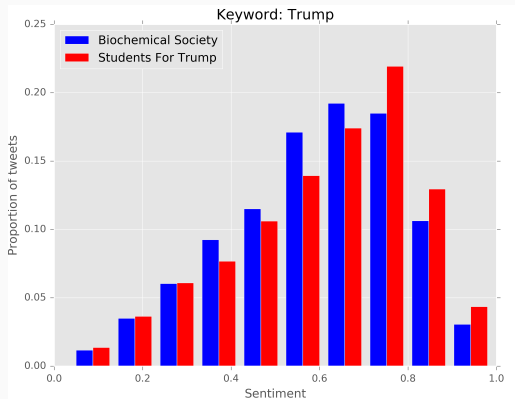
- *Biochemical Society*
- *Students for Trump*

Keyword: Trump



- slight difference in number of tweets
- *Students for Trump* slightly more positive

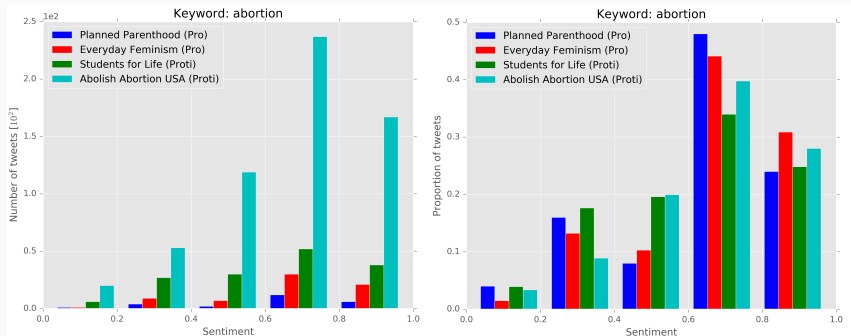
Keyword: Trump



Normalized histogram - sentiment proportion.

- *Biochemical Society* \rightarrow sentiment < 0.5
- *Students for Trump* \rightarrow sentiment > 0.5

Keyword: abortion



Number of tweets histogram, keyword: "abortion".

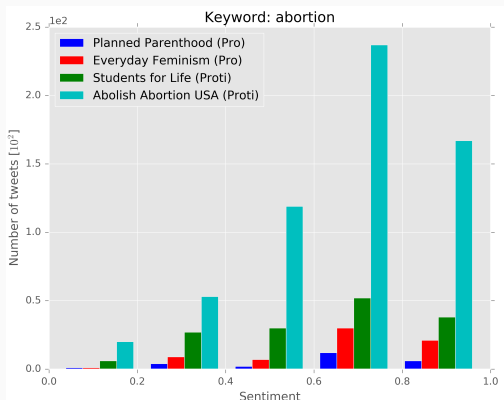
■ *Planned Parenthood*

■ *Everyday Feminism*

■ *Student for Life*

■ *Abolish Abortion USA*

Keyword: abortion

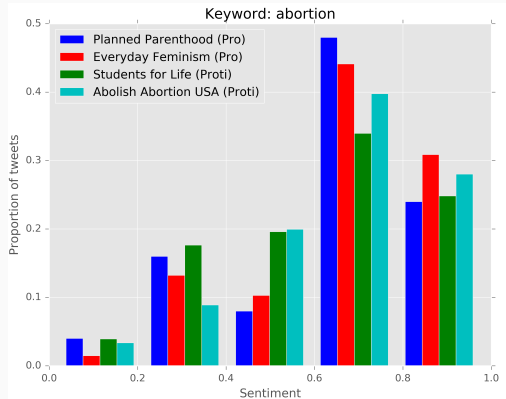


- huge difference in number of tweets

- threat for objectivity

- more activity in groups against abortion

Keyword: abortion



Normalized histogram - sentiment proportion.

- against abortion \rightarrow sentiment < 0.5
- for abortion \rightarrow sentiment > 0.5

Conclusion

New method:

- large scale - **noise reduction**
- more **straightforward** than traditional research
- quantitative research

Measurements:

- Trump → low content homogeneity
- abortion → **threat** for objectivity

Tweet examples

Everyday Feminism:

- *"We're proud of all of the abortion providers in this room - thank you for your brave; compassionate care. #Proud2Provide #LifesWork"* (0.9776)

Abolish Abortion USA:

- *"And look at the Planned parenthood abortion rooms.....887 babies killed a day and 300,000 dead babies a year....!!!"* (0.2255)