

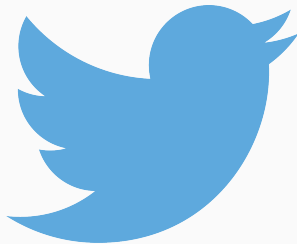
#filterbubble

Františka Sandroni

Jakub Dostál

Izolace od dostatečně širokého spektra informací.

- sociální sítě
- preferenční algoritmy
- Eli Periser (2011)



Filter bubble

Problémy:

- homogenita obsahu
- ztráta objektivit
- radikalizace

Naše práce:

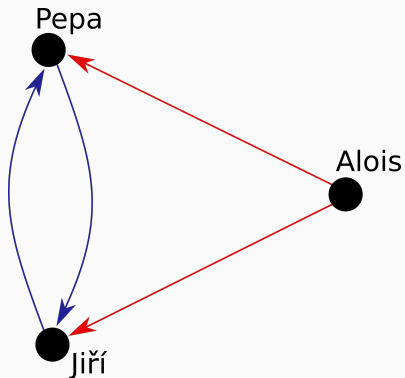
- metoda pro **studium** informační bubliny
- kvantitativní výzkum
- **detekce** filter bubble

Metoda

1. sběr dat
2. výběr studované skupiny
3. příspěvky sledované studovanými lidmi
4. filtrace příspěvků na dané téma
5. analýza vybraných příspěvků
6. srovnání dat různých skupin

Twitter

- sociální síť
- informační kanál
- **following**, **followers**
- Twitter API

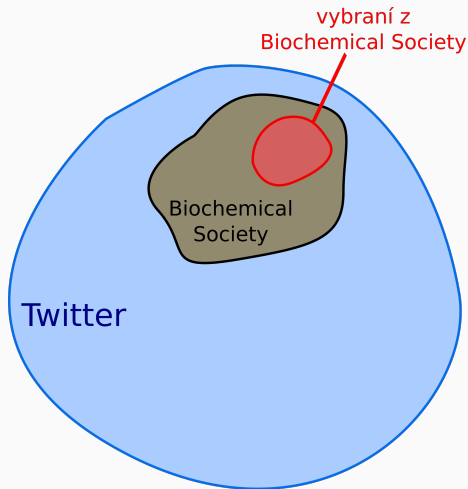


Metoda

1. sběr dat
2. výběr studované skupiny
3. příspěvky sledované studovanými lidmi
4. filtrace příspěvků na dané téma
5. analýza vybraných příspěvků
6. srovnání dat různých skupin

Výběr studované skupiny

- pozorovaná skupina
- příslušnost ke komunitě
- náhodný výběr ze sledujících



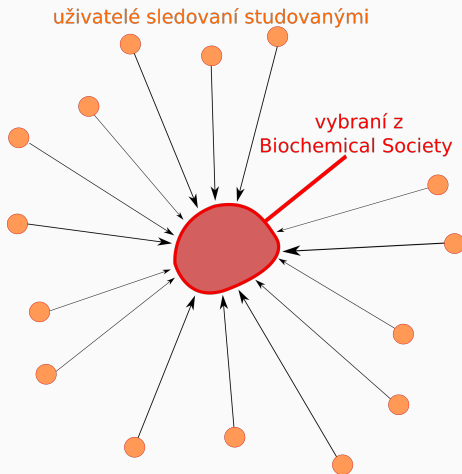
Twitter → Biochemical Society → **studování lidé**

Metoda

1. sběr dat
2. výběr studované skupiny
3. **příspěvky sledované studovanými lidmi**
4. filtrace příspěvků na dané téma
5. analýza vybraných příspěvků
6. srovnání dat různých skupin

Sběr příspěvků

- příspěvky viditelné studovanými
- okolí uživatele
- vzájemné příspěvky



Metoda

1. sběr dat
2. výběr studované skupiny
3. příspěvky sledované studovanými lidmi
4. **filtrace příspěvků na dané téma**
5. analýza vybraných příspěvků
6. srovnání dat různých skupin

Filtrace příspěvků

- přítomnost klíčového slova
- klíčové slovo = *Trump*

Na snídani jsem měl vločky s mlékem. ✗

Jsem velmi rád, že Donald **Trump** je prezidentem USA. ✓

V USA mají nejlepšího prezidenta. ✗

Metoda

1. sběr dat
2. výběr studované skupiny
3. příspěvky sledované studovanými lidmi
4. filtrace příspěvků na dané téma
5. **analýza vybraných příspěvků**
6. srovnání dat různých skupin

Sentimentální analýza

- pozitivní vs. negativní text
- klasifikace
- machine learning, big data

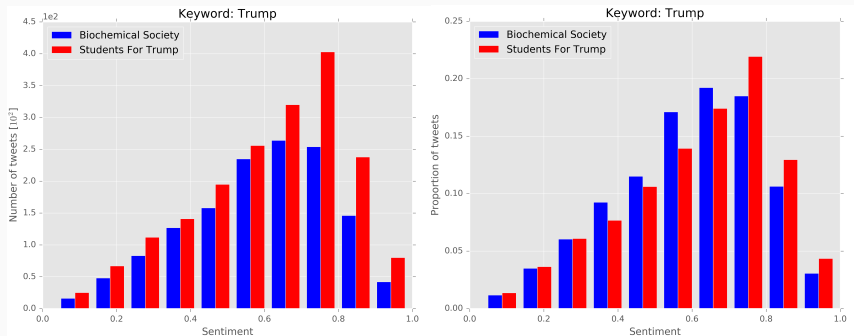
Donald Trump je hrozný člověk.
(0.14)

Donald Trump je skvělý člověk.
(0.95)

Metoda

1. sběr dat
2. výběr studované skupiny
3. příspěvky sledované studovanými lidmi
4. filtrace příspěvků na dané téma
5. analýza vybraných příspěvků
6. **srovnání dat různých skupin**

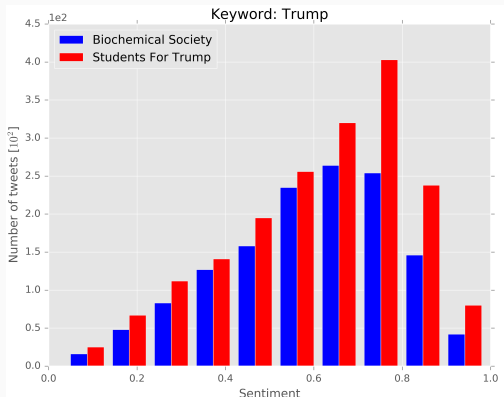
Klíčové slovo: Trump



Histogramy počtu tweetů s klíčovým slovem „Trump“.

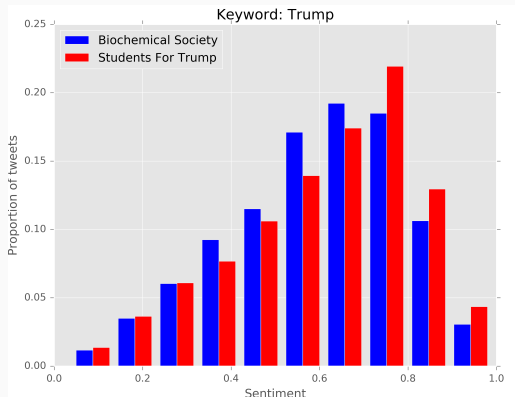
- *Biochemical Society* - financování vědy
- *Students for Trump* - motivace

Klíčové slovo: Trump



- mírně rozdílný počet příspěvků
- *Students for Trump* spíše pozitivní

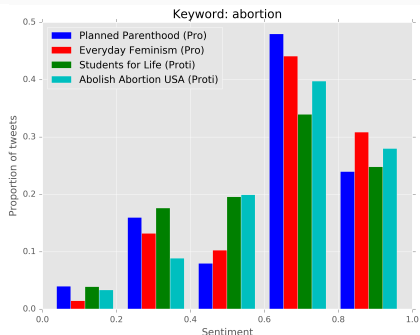
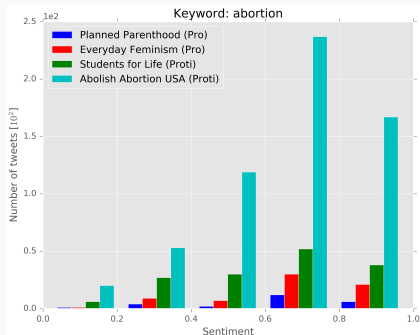
Klíčové slovo: Trump



Normalizovaný histogram - proporce sentimentu.

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

Klíčové slovo: potrat



Histogramy počtu tweetů s klíčovým slovem „abortion“.

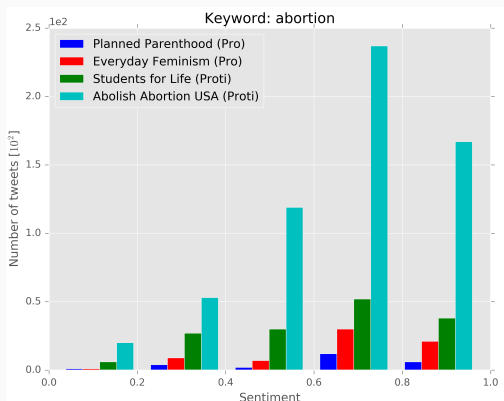
■ *Planned Parenthood*

■ *Everyday Feminism*

■ *Student for Life*

■ *Abolish Abortion USA*

Klíčové slovo: potrat

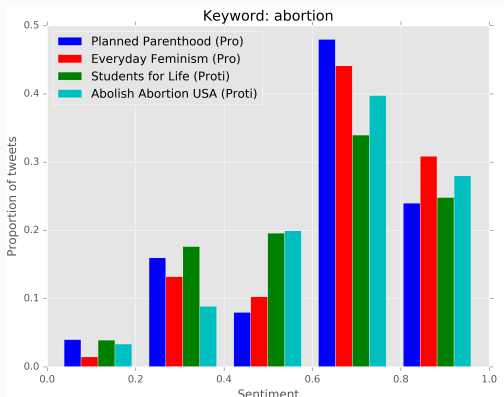


■ velký rozdíl v počtu tweetů

■ objektivita je ohrožena

■ větší aktivita skupin proti potratu

Klíčové slovo: potrat



Normalizovaný histogram - proporce sentimentu.

- proti potratu \rightarrow sentiment < 0.5
- pro potrat \rightarrow sentiment > 0.5

Nová metoda:

- velké množství dat - **potlačení šumu**
- **přímější** než tradiční výzkum
- první krok ke kvantitativnímu výzkumu

Měření:

- Trump → slabá homogenita obsahu
- potrat → **ohrožení** objektivity

Příklady tweetů

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)*

- detekce informační bubliny
- ochrana před homogenitou
- komplexnější informovanost
- zmírnění preferenčních algoritmů