You are what you Read (and Write): Inferring People's Personalities from Produced or Consumed Text to Combat Targeted Advertising

Matthew Edwards, University of Bristol
Adam Sutton, University of Bristol
Almog Simchon, University of Bristol
Stephan Lewandowsky, University of Bristol

Motivation: Targeted advertising has become more prevalent with improved access to social media data combined with big data learning algorithms. This is seen as particularly harmful in political advertising. This work is part of an attempt to mitigate the harm via first reverse-engineering the process. We survey users from ``Reddit'' for their Big Five personality, and then train neural network models to see if we can predict personalities based on produced and consumed text.

Produced Text

The Data

- **1,229 Reddit users** provided personality inventories and consented for their Reddit activity to be scraped.
- **904k comments** have been scraped in total from their posts on Reddit.
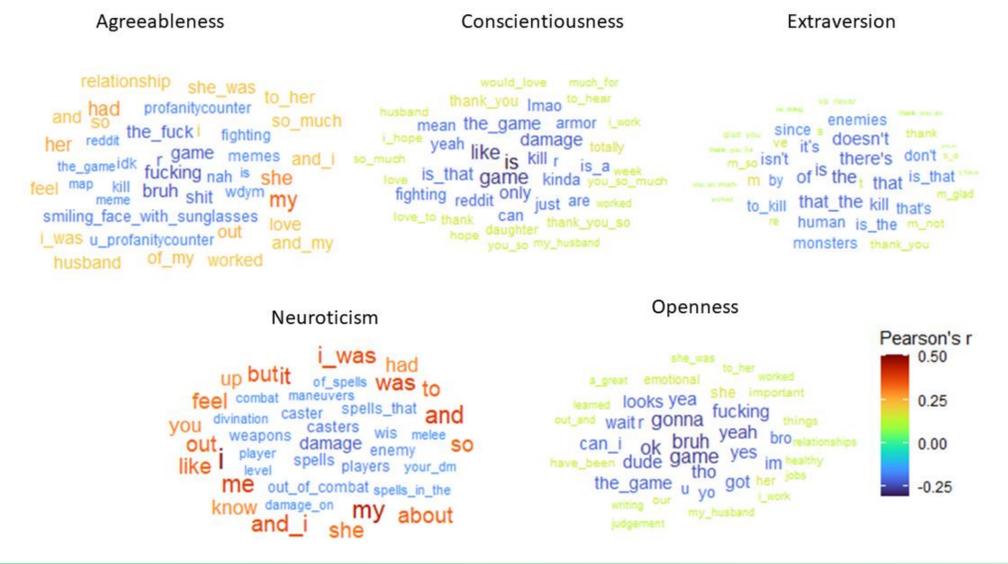
Model

- **Objective:** Predict the personality score of a user given all of their comments
- **Input:** All of a users collected comments from all subreddits.
- **Output:** A user's Personality score for a single personality dimension.

Performance

Personality Dimension	Pearson's r [95% CI]
Extraversion (EXT)	0.32 [0.22, 0.43]
Agreeableness (AGR)	0.31 [0.20, 0.42]
Conscientiousness (CON)	0.33 [0.27, 0.38]
Neuroticism (NEU)	0.33 [0.21, 0.45]
Openness to Experience (OPN)	0.32 [0.26, 0.39]

Consumed Text

The Data

- **15,041 NEWS articles** are sampled which have been consumed by at least **1** of:
- **226 Unique Users**, who also provided their personality inventory

Model

- **Objective:** Predict the average personality of all users who've consumed a given NEWS article or Title.
- **Input:** Textual content that users has been read by multiple users.
- **Output:** A news articles' average personality score for a single personality dimension.

Performance

Personality Dimension	Pearson's r [95% CI]
Extraversion (EXT)	0.3 [0.28, 0.32]
Agreeableness (AGR)	0.21 [0.19, 0.23]
Conscientiousness (CON)	0.25 [0.21, 0.29]
Neuroticism (NEU)	0.33 [0.3, 0.36]
Openness to Experience (OPN)	0.17 [0.14, 0.21]

Personality Predictive Words from Produced Text

Personality Predictive Words from News Titles
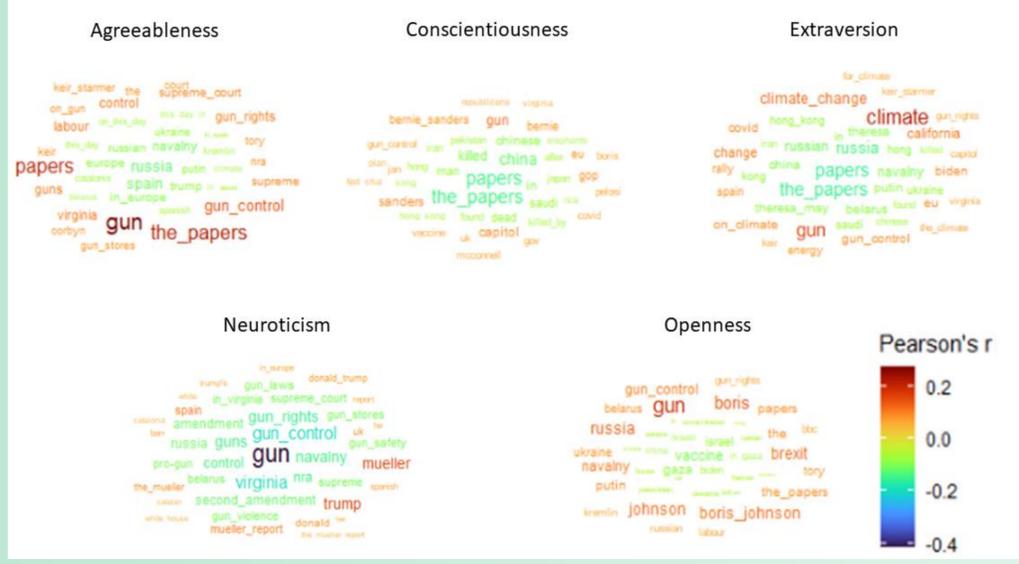
Conclusions: This work shows that personality can be learned from text that a person has produced and consumed. Additional participants for training these models may result in improved performances. This work will be used to create a tool where text can be submitted and personality scores for will be generated to inform users about potential harm.