

Social Media Data

Akitaka Matsuo
Essex IADS

Social media analysis

- It's blooming. Why?
 - Unobtrusiveness
 - (Can be) inexpensive
 - Access to the wide range of people
 - Elite
 - Mass (possibly representative)

Concerns

- Bias in the data (e.g. predicting election outcomes with Tweets)
 - Over/underrepresentation of specific population
 - Some groups are more active
- Bots
- Ethical concerns (especially experimental research with real-world interventions)
 - Informed consents
 - Boundary of influence
- Data access in future

Twitter API in the Past

- Two types of API
 1. REST API
 - Queries for specific information about users and tweets
 - Search recent tweets
 - Examples: user profile, list of followers and friends, tweets
 2. Streaming API
 - Connect to the stream of tweets as they are being published

Twitter after Elon

- The free access is gone. Starting at \$100, realistically minimum \$5,000
 - Mostly out of reach of academic researchers
- What now?
 - Non API access
 - <https://www.nature.com/articles/d41586-023-03483-8>
 - Or other platforms, but which?



Facebook API

- Facebook API was also a great source of academic research for the research of public
- Not really the case anymore, but still you can get the large number of data from public Facebook pages (e.g. government organizations, political figures, and political parties)
 - It is difficult to know the profile of people who engage
 - New attempts: <https://fort.fb.com/researcher-apis>