

# What do natural language platforms have to offer health research: leveraging soft-intelligence during the COVID-19 pandemic



Dr Christopher Marshall



## Natural language processing (NLP)

- Artificial intelligence
- Sentiment analysis
- Topic analysis/modelling
- NLP platforms
  - Machine learning
    - Rapidly improving all the time

## Soft intelligence

- Difficult to define
  - Soft vs Hard intelligence?
  - *“informal, localised, often narrative data of a kind not easily captured...”*
- ‘General chatter’
  - social media, news stories, interviews, online feedback/comments etc.
- Often overlooked

## Health research and NLP

- A LOT of health-related information out there
  - Usual suspects (journal articles, technical reports, trial registries)
  - Soft-intelligence (e.g. social media – Covid-19)

## Covid-19 and NLP

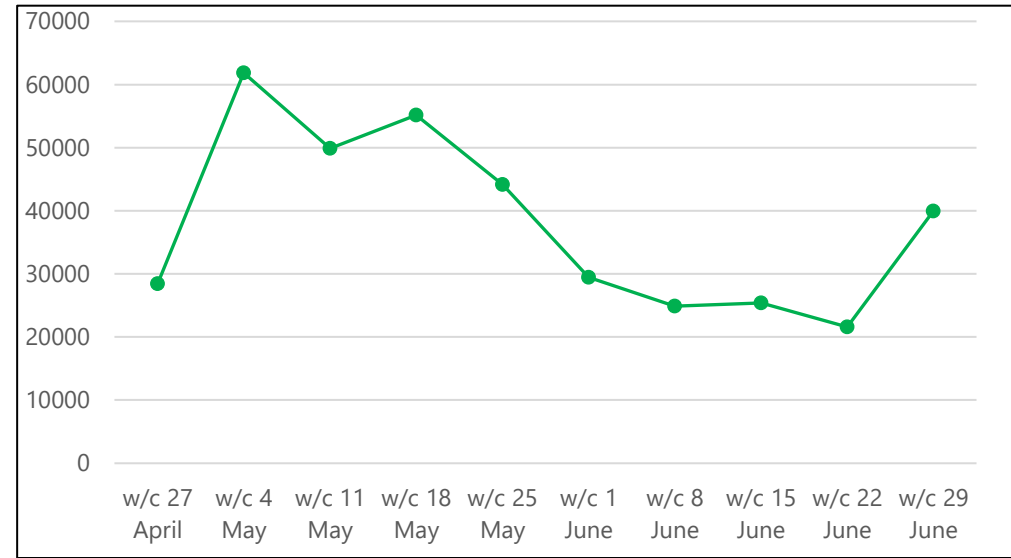
- Prominent role in outbreak response efforts
- NLP tools to tame the Covid-19 literature
  - 45,000 articles since January 1
  - NLP tools to sift through the evidence
- 16 March, White House, Call to action
  - 13,000 tagged full-texts to assist development
- Dozens of new NLP tools
  - *Covid-19 Research Explorer*
  - *COVIDScholar*
  - *Covid-19 Primer* – also pulls in some soft-intelligence
- Is there anything more we can do with soft-intelligence?

# Using Twitter to track UK public opinion in the UK about mental health during the COVID-19 pandemic

- Mental health one of the most discussed health topics during the pandemic
- Developed a search strategy of relevant terms
- Collected tweets from April 30 to 4 July 2020
- Used a specialist NLP text analytics platform, which supports topic and sentiment analysis
- **Objectives:**
  - Track tweet frequency across the UK
  - Identify key emerging topics being discussed on Twitter
  - Explore the overall trend in sentiment over time
  - Regional-level analysis

# Initial results: volume

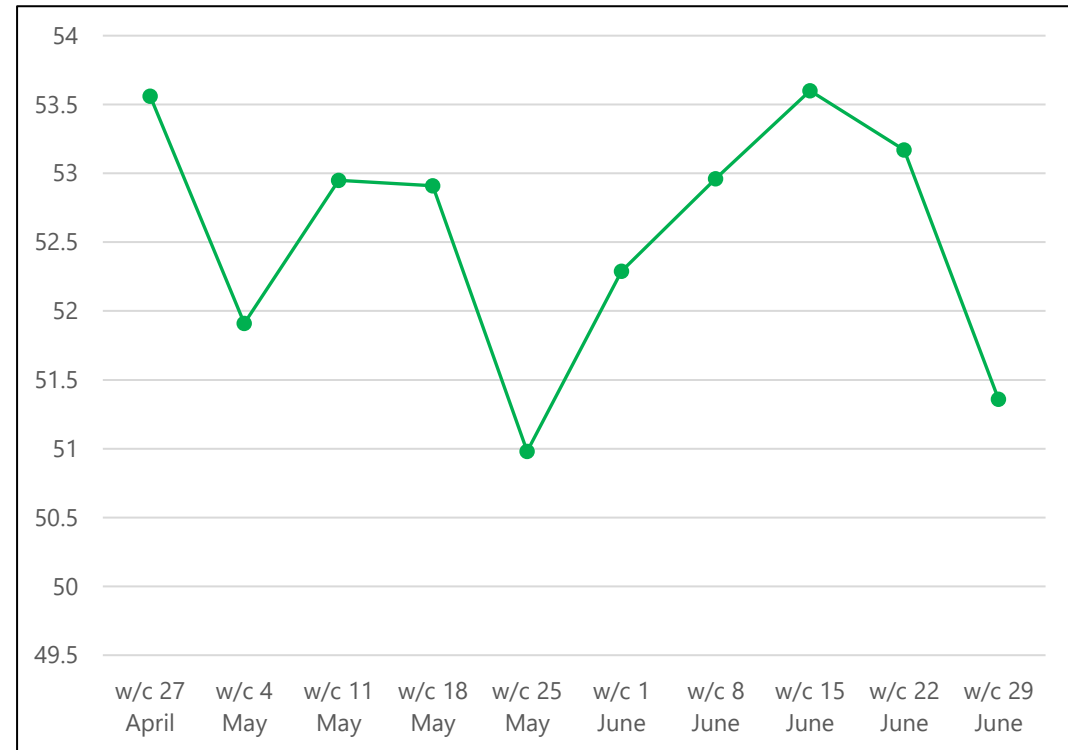
- 380,728 tweets analysed
  - Highest volume w/c 4 May
  - Lowest volume w/c 22 June
  - Overall downward trend
- Sharp increase in the final week
  - Further lockdown relaxations?
- Regional breakdown of volume



UK region	n Tweets
London	36,156
North West England	21,082
Scotland	17,395
South East England	17,368
Yorkshire and the Humber	13,822
West Midlands	11,108
South West England	10,216
East of England	9,399
East Midlands	8,419
Wales	6,523
North East England	6,053
Northern Ireland	6,030

# Initial results: sentiment

- Of the 380,728 tweets identified and analysed for sentiment:
  - **70,262 (18%) were identified as having positive sentiment**
  - 272,906 (72%) were identified as having neutral sentiment
  - **37,560 (10%) were identified as having negative sentiment**
- Overall sentiment score: 52% (suggests slightly positive sentiment)



# Useful references

- Gilbert, J. P., Ng, V., Niu, J., & Rees, E. E. (2020). Artificial intelligence in public health: A call for an ethical framework when using social media data for artificial intelligence applications in public health research. *Canada Communicable Disease Report*, 46(6), 169.
- Wang, J., Deng, H., Liu, B., Hu, A., Liang, J., Fan, L., ... & Lei, J. (2020). Systematic Evaluation of Research Progress on Natural Language Processing in Medicine Over the Past 20 Years: Bibliometric Study on PubMed. *Journal of medical Internet research*, 22(1), e16816.
- Baclic, O., Tunis, M., Young, K., Doan, C., Swerdfeger, H., & Schonfeld, J. (2020). Challenges and opportunities for public health made possible by advances in natural language processing. *Canada Communicable Disease Report*, 46(6), 161-168.
- Artificial-intelligence tools aim to tame the coronavirus literature, Technology Feature, Nature (June, 2020): <https://www.nature.com/articles/d41586-020-01733-7>
- Conway, M., Hu, M., & Chapman, W. W. (2019). Recent advances in using natural language processing to address public health research questions using social media and consumergenerated data. *Yearbook of medical informatics*, 28(1), 208.

# Acknowledgements and disclosures

- The NIHR Innovation Observatory is funded by the National Institute for Health Research (NIHR)
- The views expressed are those of the author(s) and not necessarily those of the NIHR or the UK Department of Health and Social Care