Encountering #feminism on Twitter

Viviene Cree & Steve Kirkwood, The University of Edinburgh, and Daniel Winterstein, SoDash

@VivCree @KirkwoodSteve @Winterstein
Structure of this presentation:
Focus on Methodology

♦ Aims of the project (Viv)
♦ Methodology:
  ▪ Background (Viv)
  ▪ What we did, how we did it, limitations (Daniel)
♦ Findings (Daniel)
♦ Future work (Viv)
♦ Questions
Aims

♦ Began with a realisation and a claim...
♦ But what is feminism? Does the idea of a '4th wave' (or any wave) make sense? How is feminism being expressed, and by whom? And what is the reaction to this? Led to another question -
♦ What can an analysis of social media (in this case, use of Twitter) tell us about feminism today?
♦ A series of sociological questions but using AI techniques to explore them
Methodology: Background

♦ Two cultures: social work academics with an interest in feminism & discourse analysis, and researchers from a social media analysis team.
♦ Funding from SSPS research grant
♦ A challenging enterprise – How to analyse conversations, at scale?
♦ Building cultural bridges – How do sociologists & data scientists communicate with one another when we have little shared language? We’ve learnt a lot about what we can and cannot answer...
Methodology 1: The Data

- Tweets between Dec 2012 and May 2014, using 1% Twitter 'sprinkler' data: 1 million+ messages a day
- Filtered the data against a list of topic keywords - led to a dataset of 500,000 messages

“I love a feminist bitch... oh it make my dick hard - @XDannyXBrownX”

“Warning: do NOT EVER take a women's studies course. Feminist suck and there is a high risk of suicide.”
Methodology 2: Analysis

♦ How can we code a data-set of ½ million messages?
♦ Tweets were coded ('tagged'):
  1. We designed tagsets around the project's research questions: topic, position, message-type, tone
  2. Each of our 5 research team tagged 100 tweets.
  3. Discussed and refined the tagsets
  4. Final tagged data: 800 position tags, 1,000 topic tags. Just 0.02% of the data!
  5. AI classifier trained on the tagged tweets.
  6. Pattern of AI errors measured on the tagged data
  7. AI classifier auto-tagged all the remaining tweets
  8. Error correction applied
Tagging Process

◊ Our initial tagsets included:
  • Feminist: 1\textsuperscript{st} wave / 2\textsuperscript{nd} wave / 3\textsuperscript{rd} wave / 4\textsuperscript{th} wave
  • Tone: angry / calm / joking
◊ We iterated – adjusting where things weren't working
◊ Unsuccessful:
  • “Rare” topics, e.g. children
  • Tone was too subjective to judge.
  • 1\textsuperscript{st} wave, 2\textsuperscript{nd} wave: could not assess from a tweet.
  • Message type: mostly inform-explain-proclaim.
◊ Successful:
  • The main topics: rape, abortion, violence
  • Pro / anti feminism
Limitations/issues

♦ How representative is the data?
♦ Coding - tagging is difficult & time-consuming
♦ AI for individual messages is not very accurate – but we can use error correction.
Findings

- Pro and anti feminism camps can be distinguished with reasonable accuracy – confirming that debate on Twitter is often polarised.
- Feminism on Twitter does not easily fit with the waves of feminism model.
- Rape is the dominant topic, followed by abortion.
  - Geo-political events mostly did not affect the topics (notable exception: Wendy Davis).
- High-level of casual mysogyny: e.g. tweets about rape made up about 70% of our data, mostly casual use.
  - The proportion is largely constant over time
  - What does this tell us?
Future work: This Project

- Trace specific topics/campaigns & their critiques over time: e.g. ‘#endfathersday’ and ‘#yourslipisshowing’
- How could we improve the method used?
  - More tagging needed for really robust results:
    - More training data improves AI accuracy
  - Provide a larger sample size for error correction
- More work to ensure consistency of tagging:
  - Were we each tagging in the same way?
  - Writing fuller descriptions of each tag might help
- Interestingly, these issues of validity & reliability in coding are shared with other research projects.
Future work: Methodology

♦ Streamline the development of good tagging
  • A 'best practice' guide for big-data tagsets on the basis of our experience?
♦ Improve software support for academic research
  • Report confidence intervals
  • Measure team tagging consistency
♦ Shorten the research loop: high-level question → precise question → data analysis → interpretation → further questions → …
  • The “research loop” with social media data is already fast, compared with interviews or surveys.
  • But it could be faster still.
QUESTIONS?
Viv Cree, @VivCree, v.cree@ed.ac.uk
Daniel Winterstein, @Winterstein, daniel@sodash.com
Steve Kirkwood, @KirkwoodSteve