Investigating radical content on TikTok
A computational approach

Nader Hotait

Humboldt University of Berlin
Berlin Institute for Empirical Integration and Migration Research (BIM)

University of Mannheim
Graduate School of Economic and Social Sciences

Content

(1) Introducing the RaPoTik project and its objectives
(2) Current Research Design
(3) Collecting and analyzing TikTok user data
(4) Open Challenges

Radicalization Potentials on TikTok

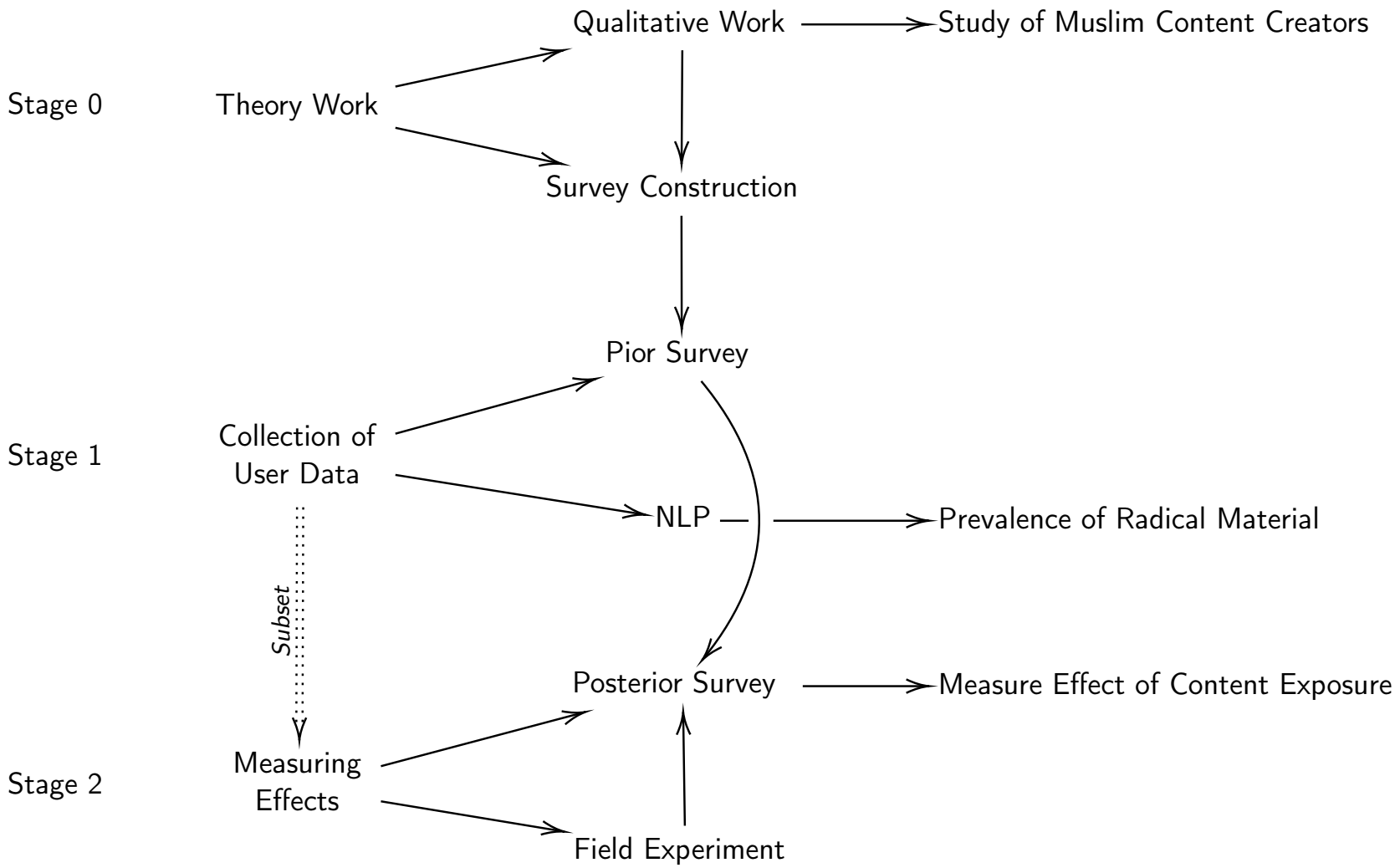
- Funded by the Berlin State Commission against Violence
- Goals:
  (1) Reconstruct the prevalence of radical material in user feeds ("For You Page")
  (2) Measure the effect of content exposure on radicalization
- Focus on religious radicalization of German-Muslim TikTok users

Background

- The Internet as the main domain for radical actors and material[1,2,3,9,12,13,14]
- Contested relationship between social media consumption and radicalization[7,17]
- Increasing scrutiny of recommender systems and their alleged role in creating echo chambers/filter bubbles[4,10,16,18]
- Remains unclear for TikTok

Why TikTok?

- Extensive radical material available[6,15]
- Algorithms possibly opening extremist rabbit holes[8]
- Youngest and most active social media user base[11]
- Possible effect of material exposure to radicalization
- Ideological reinforcements through filter bubbles
- Higher vulnerability in younger people[5]


Qualitative Work
Study of Muslim Content Creators
Stage 0
Theory Work
Survey Construction
Pior Survey
Stage 1
Collection of
User Data
NLP
Prevalence of Radical Material
Subset
Posterior Survey
Measure Effect of Content Exposure
Stage 2
Measuring
Effects
Field Experiment

Collecting User Data

- Incentivizing German Muslims ($\approx 65 - 110$) and Non-Muslims (≈ 50) to participate in our study
- Muslims are incentivized to participate in the field experiment and Non-Muslim only in the prior survey
- However, both groups are instructed to download their TikTok user data and hand it over to us, accompanied by a survey
- The survey collects data on socio-demographic and ideological indicators
- Both survey and data submission are labeled with an ID, matched, and anonymized
- User data comes in a .json files with various user information

Data Wrangling

- From .json files information on watched videos (links and date) are extracted
- From those links videos are then downloaded, transcribed, and metadata collected
  - The data (>30TB expected) is hosted on designated Humboldt University servers
  - Computing power is borrowed from cloud computing implications from Humboldt University as well
  - Audio-to-text and OCR are applied from established NLP libraries are applied to generate usable text data

Data Analysis (Suggestions)

- Extracting important features from the qualitative/mixed-methods stage
- Use prior knowledge to apply Topic Modelling/Detection for topic distribution over time
- Answer to established indicators of radicalism through text features
  - Answer how certain videos relate to indicators of radicalism (e.g. sentiment, syntax)
  - Sentiments and affect intensity analysis towards markers of radicalism
- Standard supervised learning-training/test workflow with either quantified text variables in something like Glmnet or using text elements as classifiers similar to Naive Bayes/SVM/CNN

What issues do we face?

- No API access
  - Legality
  - Efficiency
- What library and models to choose from?
- Lacking fidelity for OCR
- Model validity for radical content
- Presence of A/B testing
- ...

[1] Anne Aly et al. “Introduction”. In: *Violent Extremism Online: New Perspectives on Terrorism and the Internet*. Ed. by Anne Aly et al. London: Routledge, 2016, pp. 1–7.

[2] Mia Bloom, Hicham Tiflati, and John Horgan. “Navigating ISIS’s Preferred Platform: Telegram”. In: *Terrorism and Political Violence* 31.6 (2019), pp. 1242–1254. DOI: 10.1080/09546553.2017.1339695.

[3] Manuela Caiani and Linda Parenti. *European and American extreme right groups and the internet*. London, 2013, pp. 1–235.

[4] Elizabeth Dubois and Grant Blank. “The echo chamber is overstated: the moderating effect of political interest and diverse media”. In: *Information, Communication & Society* 21.5 (2018), pp. 729–745. DOI: 10.1080/1369118X.2018.1428656.

[5] Julie Emmelkamp et al. "Risk factors for (violent) radicalization in juveniles: A multilevel meta-analysis". In: *Aggression and Violent Behavior* 55 (2020), p. 101489. ISSN: 1359-1789. DOI: https://doi.org/10.1016/j.avb.2020.101489. URL: https://www.sciencedirect.com/science/article/pii/S1359178920301932.

[6] Friedhelm Hartwig and Albrecht Hänig. *Monitoring der Peripherie des religiös begründeten Extremismus (PrE)*. Berlin: Bundeszentrale für politische Bildung, 2021.

[7] Ghayda Hassan et al. "Exposure to Extremist Online Content Could Lead to Violent Radicalization:A Systematic Review of Empirical Evidence". In: *International Journal of Developmental Science* 12 (2018), pp. 71–88. ISSN: 2191-7485. DOI: 10.3233/DEV-170233.

[8] Olivia Little and Abbie Richards. "TikTok's algorithm leads users from transphobic videos to far-right rabbit holes". In: *Media Matters for America* (Oct. 2021). [Online; Accessed 20 Jan 2022]. URL: https://www.mediamatters.org/tiktok/tiktoks-algorithm-leads-users-transphobic-videos-far-right-rabbit-holes.

[9] Jialun Qin et al. "Analyzing terror campaigns on the internet: Technical sophistication, content richness, and Web interactivity". In: *International Journal of Human Computer Studies* 65.1 (2007), pp. 71–84. DOI: 10.1016/j.ijhcs.2006.08.012.

[10] Alastair Reed et al. "Radical Filter Bubbles: Social Media Personalisation Algorithms and Extremist Content". In: *Global Research Network on Terrorism and Technology* 8 (2019). https://static.rusi.org/20190726_grntt_paper_08_0.pdf, (Accessed: 24/10/2022).

[11] Daniel Ruby. "TikTok User Statistics (2022): How many TikTok Users Are There?" In: *Demand Sage* (2022). [Online; Accessed 09/25/2022]. URL: https://www.demandsage.com/tiktok-user-statistics/.

[12] Martin Rudner. ""Electronic jihad": The internet as Al Qaeda's catalyst for global terror". In: *Studies in Conflict and Terrorism* 40.1 (2017), pp. 10–23. DOI: 10.1080/1057610X.2016.1157403.

[13] Marc Sageman. *Leaderless Jihad: Terror networks in the twenty-first century*. Philadelphia: University of Pennsylvania Press, 2008.

[14] Gabriel Weimann. "Why do terrorists migrate to social media?" In: *Violent Extremism Online: New Perspectives on Terrorism and the Internet*. Ed. by Anne Aly et al. London: Routledge, 2016, pp. 45–64.

[15] Gabriel Weimann and Natalie Masri. "Research Note: Spreading Hate on TikTok". In: *Studies in Conflict & Terrorism* 0.0 (2020), pp. 1–14. DOI: 10.1080/1057610X.2020.1780027. eprint: https://doi.org/10.1080/1057610X.2020.1780027. URL: https://doi.org/10.1080/1057610X.2020.1780027.

[16] Joe Whittaker et al. "Recommender systems and the amplification of extremist content". In: *Internet Policy Review* 10.2 (2021). DOI: 10.14763/2021.2.1565.

[17] Michael Wolfowicz, Badi Hasisi, and David Weisburd. "What are the effects of different elements of media on radicalization outcomes? A systematic review". In: *Campbell Systematic Reviews* 18.2 (2022), e1244. DOI: 10.1002/cl2.1244.

[18] Michael Wolfowicz, David Weisburd, and Badi Hasisi. "Examining the interactive effects of the filter bubble and the echo chamber on radicalization". In: *Journal of Experimental Criminology* (2021). ISSN: 1572-8315. DOI: 10.1007/s11292-021-09471-0. URL: https://doi.org/10.1007/s11292-021-09471-0.