

1. Title of the Project

What online risks do internet users fear? A longitudinal study using machine learning and network analysis

2. Project Coordinators

Dr. Saif Shahin (TSHD, DCU)

Dr. Sara Pabian (TSHD, DCC)

3. Project Description

The internet, and digital technology more broadly, has brought the world together, created new forms of livelihood for millions of people, and allowed us to educate or entertain ourselves in multifarious ways. At the same time, digitalisation has enabled or exacerbated many social issues. To illustrate, more than 400,000 children and adolescents in the Netherlands annually report they have experienced cyberbullying.¹ Unmitigated hate speech on Facebook has been alleged to have contributed to the genocide of Myanmar's Rohingya minority.² Disinformation on social media has kept millions of people from getting vaccinated against Covid19.³ Meanwhile, online echo chambers have given rise to political polarisation and ethnocentric radicalisation that are tearing societies apart.⁴

Digitalisation has also produced particular challenges for vulnerable populations. Every second person of colour in Canada, for instance, reports racial discrimination on the internet.⁵ Anti-immigration groups are known to network online to create 'moral panic' against Latino immigrants⁶ in the US and Muslim immigrants in Europe.⁷ Women and sexual minorities, meanwhile, are much more likely to experience harassment and abuse online than heterosexual men.⁸

Scholars of digital society have produced a vast body of research on each of these 'online risks', typically dealing with their antecedents, effects, and possible solutions. However, little is known about how common users of the internet view or weigh these issues, who shapes their perception of these issues, and how they believe the internet can be made safer.

This study aims to shed light on internet users' understanding of the different issues related to internet safety and how it has evolved over time. To do so, it will focus on the Safer Internet Day (SID), an annual campaign to 'make the internet a safer and better place for all'.⁹ Initiated by the EU SafeBorders project in 2004, SID is now a global event marked by a series of actions by affiliated organisation in more than 200 countries and territories. Their aim is 'to raise awareness of emerging online issues and current concerns'.¹⁰ Specifically, the study will analyse tweets related to SID over 12 years [2012-2023] to understand:

RQ1: What are the different types of online risks related to internet safety that internet users tweet about?

RQ2: How have these issues changed over the years?

RQ3: What types of Twitter users are likely to influence online conversations about different issues related to internet safety?

Tweets that use hashtags related to SID, such as #SaferInternetDay and #SID22 [for the year 2022], will be the primary data for this study. For each year, data will be collected for a week starting from the SID. For example, as the SID in 2022 was observed on 8 February, all tweets using SID-related hashtags posted from 8-14 February will be collected. The study will rely on Twitter's latest API, which enables access to the platform's historical data for academic research.¹¹ In addition to the content of the tweet, the API provides a wealth of metadata, including the tweet's author, bio, location, number of followers, retweets, and so on. Only publicly available data will be collected, in line with the ethical guidelines of the Association of Internet Researchers.¹²

The dataset is expected to comprise of more than half a million tweets — an estimate based on the 2022 sample, already collected by the project coordinators, which includes nearly 66,000 tweets. The study will draw on a computational mixed-methods approach to analyse the data. Two particular methods will be employed.

Study 1: Machine Learning

In the first stage, the study will use machine learning to distinguish the different risks related to internet safety raised by internet users each year. Machine learning will involve hand-coding a small sample of tweets for various categories [risks] that will be used to train an algorithm to 'recognise' different risks. The algorithm's efficiency will be tested on another sample of tweets. Once the algorithm reaches acceptable levels of reliability, it will be used to analyse the rest of the corpus. The analysis will disclose the different online risks that internet users have been tweeting about (RQ1) and how the composition of risks — in terms of the proportion of tweets about each risk — has changed over the years (RQ2).

Study 2: Network Analysis

Tweets also enable the formation of 'networks' through retweets and replies — representing how different Twitter users connect with each other in conversations about different online risks. In the second stage, the study will employ network analysis to identify Twitter users that are likely to dominate these networks and influence conversations about internet safety (RQ3). The study will also draw upon their user bio, location, follower count and other metadata to develop a typology of users that are likely to influence online conversations on these risks.

This project carries theoretical as well as societal significance. First, it can guide both scholarly research and public outreach agendas related to internet safety. Greater effort may be warranted for understanding and dealing with those issues that internet users are themselves more concerned about. At the same time, the research will distinguish risks that, while important, might not be high on the radar of internet

users. These findings can improve digital education in schools and universities, so that younger people in particular, who form the majority of the internet users, are better aware of the challenges of life in contemporary digital society.

This research will bring together different theoretical perspectives and empirical evidence on online risks, as well as methodological expertise, from the fields of culture studies and communication sciences, but also disciplines such as computational and educational sciences, and social psychology. Dr. Saif Shahin, one of the project coordinators, is well-versed in the use of computational research and has published multiple articles using these methods. The other coordinator, dr. Sara Pabian, is an expert in the field of online risks. The objective of the project is a joint academic publication, which will be unique and of added value due to its multidisciplinary approach. The results may serve as input for a joint funding application on the development of digital tools to support victims and mitigate the potential impact of online risks.

4. Project Timeline

Activity	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
Start traineeship										
Theory development										
<i>Study 1</i>										
Data collection (2012-22)										
Manual coding										
Training and testing ML algorithm										
Data collection (2023)										
Data analysis (machine learning)										
<i>Study2</i>										
Data analysis (network analysis)										
Classification of influential accounts										
Write journal manuscript										
Deliverables			D1 D2			D3 D4	D5		D6	D7 D8

Deliverables

D1: Submit abstract to Etmaal van de Communicatiewetenschap conference*

D2: Submit extended abstract to International Communication Association conference*

D3: Present paper at Etmaal

D4: *Diggit* magazine article

D5: Present findings to peers (DCU monthly seminar)

D6: Present paper at ICA

D7: Submit manuscript to *New Media and Society*

D8: Present findings at Research Traineeship symposium

** Both conferences bring together scholars from cultural studies, communication science as well as other disciplines relevant to this study, such as computational science and social psychology.*

5. Research Trainee Profile

General tasks. The research trainees will form a research team, including both research trainees and the coordinators. Both research trainees will dive into the literature and write a theoretical framework, in collaboration with the coordinators. As we propose a mixed-methods approach, one research trainee will take the lead in collecting, coding, and analyzing data for machine learning, while the other will take the lead in analyzing data for social network analysis. The whole team will work together on reporting the research. We want to involve our research trainees in all steps involved in conducting a scientific study in order for them to gain as much experience as possible, while being supported by us, senior researchers, who are experienced in the proposed research topic as well as the proposed methodologies.

Trainee expectations. We are looking for two excellent and enthusiastic students. We would prefer to hire students with training and/or experience in using computational research methods. However, students who lack such experience but are interested in computational methods may apply as well, so they can learn from working with the project coordinators. Both bachelor, master, and research master students with at least a 7.5 grade average are eligible to apply. Both trainees should be available for 1 day a week between September 2022 and June 2023 and should have excellent communication skills.

Application information. Applications, including a motivation letter, resume and an example showcasing the writing skills of the applicant (e.g., course paper), should be sent to both project coordinators: dr. Saif Shahin (s.s.shahin@tilburguniversity.edu) and dr. Sara Pabian (s.j.r.pabian@tilburguniversity.edu). In the motivation letter, clearly indicate your interest in the topic of the traineeship as well as your experience with data collection from Twitter, machine learning and network analysis. Please do not hesitate to contact us with any questions using the above email addresses.

6. References

1. Stop Pesten Nu. (n.d.). Online bullying, cyber bullying, knowledge center. <https://www.stoppestennu.nl/online-pesten-cyberpesten-kenniscentrum>
2. Global Witness. (2022). Facebook approves adverts containing hate speech inciting violence and genocide against the Rohingya.

<https://www.globalwitness.org/en/campaigns/digital-threats/rohingya-facebook-hate-speech/>

3. Pierri, F. et al. (2022). Online misinformation is linked to early COVID-19 vaccination hesitancy and refusal. *Scientific Reports*. <https://www.nature.com/articles/s41598-022-10070-w>
4. Glasgow Caledonian University. (2021). Echo-chambers on social media radicalise young people, according to study. <https://www.gcu.ac.uk/theuniversity/universitynews/2021-echo-chambers-social-media-radicalise/>
5. Kane, L.D. (2021). 1 in 2 people of colour have experienced online racism in Canada, survey says. *Global News*. <https://globalnews.ca/news/7709742/online-racism-canada-survey/>
6. Flores-Yeffal, N., Vidales, G., & Plemons, A. (2011). The Latino cyber-moral panic process in United States. *Information Communication and Society*, 14(4), 568-589.
7. Farkas, J., Schou, J., & Neumayer, C. (2018). Cloaked Facebook pages: Exploring fake Islamist propaganda in social media. *New Media & Society*, 20(5), 1850-1867.
8. Women's Media Center. (n.d.). WMC Speech Project. <https://womensmediacenter.com/speech-project/research-statistics>
9. Safer Internet Day. (n.d.). Together for a better internet. <https://www.saferinternetday.org/home>
10. Safer Internet Day. (n.d.). About Safer Internet Day. <https://www.saferinternetday.org/about>
11. <https://developer.twitter.com/en/docs/tutorials/choosing-historical-api>
12. Association of Internet Researchers. (2019). Internet Research: Ethical Guidelines 3.0. <https://aoir.org/reports/ethics3.pdf>