Teaching students to better manage their digital security
(Technical Paper)

Far-right radicalization through the "For You Page" on TikTok
(STS Paper)

A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

By
Claire Cofield

December 9, 2022

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISORS

MC Forelle, Department of Engineering and Society

Rosanne Vrugtman, Computer Science

Introduction

We live in a digital age, run by digital networks, on digital footprints, using digital relationships. But many people traversing this digital landscape do not fully understand it. In fact, according to the Pew Research Center, many American adults "struggle to understand the nature and scope of data collected about them" (2016). They also struggle to understand how this collected information can be used to affect their behavior without their conscious realization.

Much of the social media content a user consumes is not material the user sought, but rather, is material the platform itself decided to show the user. Platform-selected posts are meant to engage users, so they spend more time on the app, and the app makes more money from them. However, hours of interacting with this media - even just through passive consumption - is actually capable of changing users' thoughts and behaviors (Hernandez, 2022). Through displays of extreme content, and by showing users that social communities exist around extremist talking points, social media can push users towards extremism (Martineau, 2019).

Though the effects of social media-based radicalization appear in many ways, I will focus on far-right political radicalization, particularly through the lens of the platform-determined feed known as the For You Page (FYP) on TikTok. TikTok is a popular social media application that shows its users a short videos, selected for the user by the app's artificial intelligence (Taulli 2020). Its recommended content algorithms show moderately conservative individuals right wing extremist videos (Borbee, 2022). Peripheral persuasion and repeated exposure then further draw these users in, causing many of them to develop extreme, far-right positions (Borbee, 2022).

This radicalization is fueled by algorithmic models of behavior and identity, facilitated by the large amounts of data that companies such as TikTok gather, collect, and buy about users (Boeker & Urman, 2022). Buying data to use for targeted ads and other content displays is a

huge industry, yet many people have no idea of the sheer scale of data that can be aggregated against them (Kans, 2021). There are two related problems at issue: large-scale collection of personal data about users without their informed consent, and the misuse of that data in ways that pose risk to the individuals and their communities. I will investigate both problems, as they are deeply intertwined. In my technical research, I will explore the possible development of an educational curriculum that would teach students to better maintain their data privacy and information security, specifically by studying the ways the Privacy in the Internet Age and Introduction to Cybersecurity classes at the University of Virginia (UVA) address this problem, and propose a synthesis of the two classes to more effectively focus on these issues. In my STS research, I will examine the ways user data is applied to affect behavior - specifically by studying the ways TikTok's platform-chosen content can drive users towards far-right political extremist positions.

Technical Topic

Many people do not know how to protect their personal digital security, which allows their digital footprints to be used against them. A digital footprint is the collection of the digital traces left by an individual interacting with something digitally (NSW Department of Education, n.d.). Visiting the dentist, applying for a credit card, and registering a car all leave digital traces, adding to one's digital footprint (Lim, 2021). Social media use and internet browsing add further. Much of the information that forms these digital footprints is not information a user shared intentionally. For example, taking a picture on a smartphone can automatically store geolocation information; background apps can store, log, and track data about a user when not in use; and

accounts with different phone numbers, names, email addresses, or credit card numbers associated with them can nonetheless be traced back to one person (Cveticanin, 2022).

Additionally, individuals have little to no control over who has access to their personal data, despite the fact that a large digital footprint can be harmful to a user (Federal Trade Commission, 2014). Data brokering - the process of one business selling their customers' data to another or to governments - is a common practice (Kans, 2021). In addition to being sold, a user's data can be spread through illegal means: servers storing user data are commonly the targets of hackers (Kans, 2021). If an attacker has access to a user's digital footprint, it can make it easier to obtain user passwords, commit identity theft, and lock users out of their accounts (NSW Department of Education, n.d.). There have even been cases of stalkers using their victims' digital footprints to track them down, and various groups using digital footprints to track and punish suspected political dissenters and activists (Chandler, 2019; NSW Department of Education, n.d.).

Education on internet best practices can reduce the risk of harm. Simple acts like deleting old emails and social media accounts, using ad blockers and VPNs, or deleting the location information from pictures posted online, can offer significant protection (Lim, 2021). UVA offers such education, to a certain extent, in two courses: Introduction to Cybersecurity and Privacy in the Internet Age. Parts of both classes teach students how to better navigate the digital world. I propose a new course synthesizing the best practices taught in each course, to more effectively increase students' ability to minimize their footprint and protect their data.

In Introduction to Cybersecurity, students learn about the negative consequences of data storage being compromised, whether that be a company server with thousands of people's credit card information, or a singular individual's email account (Orebaugh, 2022a). They also learn the

shocking prevalence with which users tend to keep default passwords on their devices and networks; or use common, short, frequently reused passwords (Orebaugh, 2022b). In Privacy in the Internet Age, students learn about how data is transmitted between devices and over networks, and the various points of vulnerability along those paths (Sun, n.d.). They also learn about how mechanisms like cross-site tracking, privacy-averse default settings, and data aggregation compound to enlarge a user's digital footprint (Sun, n.d.). This class also teaches students how such systems can be used for political purposes, such as censorship, finding and arresting political dissenters, and carrying out governmental or corporate surveillance (Sun, n.d.).

STS Topic

TikTok's business model aims to increase user engagement. It relies on machine learning, which has reached a level of sophistication where models can parse the emotional signals in texts and the semantic similarity between media (Araque & Iglesias, 2020). The app is thus capable of identifying which posts are more likely to get emotional responses, and has an incentive to rank such posts higher when deciding which videos to display on the FYP (Araque & Iglesias, 2020). Political issues often get emotional responses, which means TikTok often shows political videos (Carson, 2021). When TikTok identifies a relatively politically moderate user leaning left or right, the app begins showing them content that is politically inclined in that direction, with this content getting more extreme over time (Carson, 2021). Extremist groups further this display of radical content by actively attempting to promote themselves on TikTok, using trending audio clips and hashtags to increase their chances of showing up on a user's FYP (O'Connor, 2021). Continued exposure to this extreme content normalizes and even banalizes the ideas presented

therein, which is dangerous as it idealizes and can thus encourage engaging in politically radical behavior (Boucher, 2022). Thus, the platform-selected content shown to users can end up posing a safety risk to users and their communities.

The STS framework I chose to rely most heavily upon for my project is technological citizenship, a framework established by Philip Frankenfeld at the University of Chicago. One of the basic ideas of technological citizenship is that citizens should have the rights to "the political resources they need to protect themselves from and verify their safety amid complex, concealable environmental hazards" (Frankenfeld, 1992, p. 459). The four main types of rights in technological citizenship are "1. rights to knowledge or information; 2. rights to participation; 3. rights to guarantees of informed consent; and 4. rights to the limitation on the total amount of endangerment of collectivities and individuals" (Frankenfeld, 1992, p. 465).

My project will investigate how far-right radicalization through the FYP violates these rights. The Federal Trade Commission has found that much of the customer data companies have access to were acquired without the customers' knowledge or consent (2014). Furthermore, even in cases where users legally consent to some of this data harvesting by agreeing to the privacy policies of various online entities, they lack the knowledge necessary for this legal consent to be considered informed consent, as the average user does not comprehend TikTok's complex algorithmic models or the vast amounts and varied types of data TikTok has and can feed into its models. Thus, users are denied the third category of rights of digital citizens (Boeker & Urman, 2022).

Additionally, the opaque nature of machine learning models and algorithms, and the fact that many legislators do not have the technical knowledge needed to understand these algorithms, makes it hard to place restrictions on them and legislate the use of algorithmic

mechanisms (Kans, 2021). These facts limit the ability of citizens to exercise their category two and category four rights as digital citizens as well. With the research I will conduct as part of my STS project, I will aim to discover more about the processes by which this far-right radicalization is made possible on TikTok, and summarize and report on it in such a way that will allow for users and legislatures to have more agency in these matters.

Research Question and Methods

Social media has a proven effect on users' moods and behaviors (Rodriguez, 2021). With over half of American teens regularly using TikTok, TikTok specifically has a significant effect (Rodriguez, 2021). With my research I will aim to answer the following question: how does consumption of platform-selected media on TikTok - namely, the FYP - steer users towards far-right radicalization? The methodological approach I will use to conduct my research is a literature review and synthesis. I will compile data from sources across various fields of scientific study, investigating the different perspectives each field brings.

I will use technical sources to determine how users are shown increasingly extreme content. I will read TikTok's privacy policy to determine what user data TikTok collects and whether the policies make that collection clear. To examine radicalization, I will read the reports of various researchers who fed information into dummy accounts, and recorded the resulting feeds TikTok produced. These studies will also be helpful in their discussion of how certain non-user-entered data is used by TikTok, such as location data.

I will use psychological/social sources to see how repeated exposure to content and community leads to radicalization - that is, how users are influenced by even passive consumption. I will also review articles by anti-terrorism organizations about social media-based

terrorism recruitment programs, to determine if these programs pose a credible danger to communities. Additionally, I will read the results from psychological studies on the effects of echo chambers on participants' actions and beliefs, and on the risk factors that make a user more likely to fall victim to online radicalization.

Conclusion

Large amounts of data about individuals are collected and distributed without their knowledge. These data are widely used to fuel behavior-based machine learning models, such as the content-display algorithms used by many online platforms. One such algorithm is TikTok's FYP video-selection algorithm, which shows users content that can radicalize them. Both this collection/distribution of data and the behavior-affecting models built on it violate users' rights to informed consent.

Through the technical research proposed in this prospectus, I will design an educational course to teach students about internet security hygiene and managing their digital footprint. Through the STS research proposed, I will provide a deeper understanding of the radicalizing mechanisms of TikTok FYP content-selection algorithm. The findings of this research will hopefully encourage changes on individual, corporate, and legislative levels. Awareness of the dangers posed by a digital footprint will encourage users to manage theirs better. Demonstrating corporate culpability in the radicalization process will hopefully start a shift toward greater corporate responsibility. Finally, determining the specific mechanisms that lead to online radicalization will allow legislatures to target those features when regulating social media platforms for the good of the public.

References

Araque, O., & Iglesias, C. A. (2020). An approach for radicalization detection based on emotion signals and semantic similarity. *IEEE Access*, *8*, 17877–17891. https://doi.org/10.1109/ACCESS.2020.2967219

Becker, A. (2021, August 24). *On TikTok, misogyny and white supremacy slip through 'enforcement gap.'* The 19th. https://19thnews.org/2021/08/tiktok-misogyny-and-white-supremacy-slip-through-enforcement-gap/

Boeker, M., & Urman, A. (2022). *An empirical investigation of personalization factors on TikTok*. University of Zurich, Switzerland. https://www.researchgate.net/publication/358233121_An_Empirical_Investigation_of_Personalization_Factors_on_TikTok (Original work published 2022)

Borbee, J. (2022). *#FRINGETOK: An Ethnographic Content Analysis of TikTok's Far-Right Radicalization Pipeline* [Degree of Master of Arts in Strategic Communication]. American University School of Communication.

Boucher, V. (2022). *Down the TikTok rabbit hole: Testing the TikTok algorithm's contribution to right wing extremist radicalization* [Master of Arts thesis, Queen's University]. https://qspace.library.queensu.ca/handle/1974/30197

Carson, D. (2021). *A content analysis of political discourse on TikTok*. University of Mary Washington. https://scholar.umw.edu/student_research/415

Chandler, S. (2019, October 11). *Social media is fostering a big rise in real-world stalking*. Forbes. https://www.forbes.com/sites/simonchandler/2019/10/11/social-media-proves-itself-to-be-the-perfect-tool-for-stalkers/

Cveticanin, N. (2022, April 11). *What is a digital footprint: everything you need to know*. Data Prot; DataProt. https://dataprot.net/articles/what-is-a-digital-footprint/

Federal Trade Commission. (2014). *Data Brokers: A Call for Transparency and Accountability*. Federal Trade Commission. https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf

Frankenfeld, P. J. (1992). Technological Citizenship: A Normative Framework for Risk Studies. *Science, Technology, & Human Values*, *17*(4), 459–484. https://www.jstor.org/stable/689737

French, R. B. et al. (2021, September 8). *How TikTok serves up sex and drug videos to minors*. WSJ; The Wall Street Journal. https://www.wsj.com/articles/tiktok-algorithm-sex-drugs-minors-11631052944

Hernandez, A. (2022). *TikTok's Strategic Influence on Political Communication and Campaigns*. California State Polytechnic University. https://scholarworks.calstate.edu/downloads/1v53k3627

Kans, M. (2021, April 29). *Data brokers and national security*. Lawfare; The Lawfare Institute. https://www.lawfareblog.com/data-brokers-and-national-security

Lim, J. (2021, August 11). *Why digital footprints are dangerous & how to erase it.* Defiel. https://defiel.com/how-to-erase-digital-foot-print/

Martineau, P. (2019, October 23). Maybe it's not YouTube's algorithm that radicalizes people. *Wired*. https://www.wired.com/story/not-youtubes-algorithm-radicalizes-people/

NSW Department of Education. (n.d.). *Leaving a digital footprint*. Digital Citizenship; NSW Government. Retrieved October 27, 2022, from https://www.digitalcitizenship.nsw.edu.au/articles/leaving-a-digital-footprint

O'Connor, C. (2021). *Hatescape: An In-Depth Analysis of Extremism and Hate Speech on TikTok* (pp. 1–57). Institute for Strategic Dialogue. https://www.isdglobal.org/isd-publications/hatescape-an-in-depth-analysis-of-extremism-and-hate-speech-on-tiktok/

Orebaugh, A. (2022a). *CS 3710: Introduction to Cybersecurity, Spring 2022*.

Orebaugh, A. (2022b, January 21). *CS3710 Week 1* [Class lecture and slides].

Pew Research Center. (2016, September 21). *The state of privacy in post-Snowden America* [Nonpartisan fact tank]. Pew Research; Pew Research Center. https://www.pewresearch.org/fact-tank/2016/09/21/the-state-of-privacy-in-america/

Rodriguez, S. (2021, November 18). *TikTok usage surpassed Instagram this year among kids aged 12 to 17, Forrester survey says*. CNBC. https://www.cnbc.com/2021/11/18/tiktok-usage-topped-instagram-in-2021-among-kids-12-to-17-forrester-.html

Sun, Y. (n.d.). *CS4501-001: Privacy in the internet age, fall 2022*. https://www.cs.virginia.edu/~ys3kz/courses/fall22/cs4501/

Taulli, T. (2020, January 31). *Tiktok: Why the enormous success?* Forbes; Forbes Media LLC. https://www.forbes.com/sites/tomtaulli/2020/01/31/tiktok-why-the-enormous-success/