

Thesis Portfolio

Demystifying Data Collection on Facebook
(Technical Report)

**Toward Broader Knowledge of the Dangers of Facebook: An Analysis of User Awareness
and Behaviors**
(STS Research Paper)

An Undergraduate Thesis

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Stephen Shamaiengar
Spring, 2021

Department of Computer Science

Table of Contents

Sociotechnical Synthesis

Demystifying Data Collection on Facebook

Toward Broader Knowledge of the Dangers of Facebook: An Analysis of User Awareness and Behaviors

Thesis Prospectus

Sociotechnical Synthesis

Facebook and its subsidiary platforms, including Messenger, Instagram, and WhatsApp, are the most populous social networking platforms, with over 3 billion monthly active users across them as of 2021. They serve as free hubs of communication and information for users while providing valuable advertising services for organizations, a lucrative business model that has turned the Facebook company into one of the most valuable in the world. But the manner by which Facebook has reached this level of financial success comes at costs to users and to society as a whole. These costs create a need for reform of the Facebook platform and reconsideration of our continued usage. This thesis delves into this issue on two fronts, in pursuit of the mitigation of Facebook's negative impacts and an improvement to our Internet-heavy existence. On one front, the technical thesis covers transparency around data collection and aggregation on Facebook. On the other, the STS thesis characterizes user awareness and perception of Facebook's inner workings and impacts broadly.

Technical Thesis

To power the behavioral predictions that drive Facebook's advertising targeting services, the platform mines copious amounts of data from users of the platform. Some of this is well known and documented in the privacy policy, but much of it happens behind the scenes, extracting subtleties like the exact time spent viewing each piece of content in the News Feed, or the actions one takes on third-party websites that display Facebook advertisements. Many users report feeling uneasy about the extent to which Facebook collects and uses their private data, but most of them are likely not aware of the full extent of these practices. In an effort to increase transparency around data collection on Facebook, the technical project of this thesis proposes a design for a mock social media platform that would clearly display the data points being

collected as users browse a Facebook-esque content feed. Future implementation of this proposal as an educational tool could cause people to reduce their usage of Facebook and thus help reduce its negative impacts.

STS Thesis

Many of these negative impacts have been widely publicized, contributing to the creation of a movement to quit Facebook. This movement is not mainstream, however, and has not grown enough to have much tangible benefit. In an effort to characterize the kinds of outreach and messaging that might be needed to grow the movement, the sociotechnical project of this thesis analyzes the public awareness and behaviors regarding Facebook. For example, it finds that many users cannot explain how the News Feed algorithm determines what content they see, and some users are not aware that such an algorithm exists at all. Overall, this thesis finds that the platform continues to grow despite its harms because of widespread lack of awareness of its inner workings and impacts. These results help to illuminate the path forward for the movement to boycott and reform Facebook.

My work on this thesis has provided critical lessons that will benefit me in my own career in the technology industry. Through learning about the rise of social media and its problems, I have realized the importance of making engineering and design decisions with the users and the broader public in mind. The Center for Humane Technology, an organization founded by former technology employees to re-envision digital technology without its current harms, puts it nicely, stating that “Technology is never neutral.” The social is influenced by these technologies in many ways, and to ignore that influence or dismiss it in light of economic motives is an error. I aim to keep these lessons in mind in order to help bring humanity back into technology going forward.