Thesis Project Portfolio

Data-Driven Event Sequence Visualization of Rectal Cancer Outcomes

(Technical Report)

The Mechanisms and Conditions of End-User Manipulation in For-Profit Mobile Health Applications

(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

Andrew Balch

Spring, 2025

Department of Computer Science

Table of Contents

Sociotechnical Synthesis

Andrew Balch

April 20, 2025

While my technical report and my STS research project may seem disconnected, I argue that the development of any algorithmic system that guides and informs human decisions should also involve a critical analysis of how such systems encode and reproduce normative understandings of the world. My technical report introduces a data visualization and analysis tool to aid clinical decision-making over the course of rectal cancer treatment. Developed in collaboration with a stakeholder in rectal cancer surgery, the tool is based on a Sankey diagram and features integrated statistical tests and machine learning model-building. Conversely, my STS research paper seeks to understand how for-profit mobile health applications may unjustly influence end-users through their central features. To do this, I analyzed the descriptions of six top-grossing health apps through the mechanisms and conditions framework for affordances.

Effectively treating rectal cancer requires attentive consideration of hundreds of details about the patient, and how each may influence patient outcomes and negative side-effects. The US Rectal Cancer Consortium has compiled a one-of-a-kind dataset that contains fine-grain details about individual patient treatment paths from 6 different institutions over a decade. In my technical report, I worked with the Human-AI Technology Lab at UVA to develop a data-driven event sequence visualization of the incidence of a major postoperative complication (anastomotic leakage). Each data point in the RCC was represented as a node in a flowchart that culminates in the incidence of an anastomotic leak. The user can choose which treatment features they would like to investigate in the context of the complication, but because there are hundreds of treatment variables that may be relevant, we also implemented a novel chained feature selection machine

learning approach to help uncover and present which variables are powerful predictors of an anastomotic leak. We found that this approach was more effective for highlighting a variety of variables from different stages in the treatment process, compared to the pre-existing Sequential Feature Selection algorithm. A rectal cancer surgeon stakeholder was consulted throughout development and provided feedback on the final tool. Importantly, the tool is generalizable to other event sequence datasets, and has been made publicly available at https://github.com/HAI-lab-UVA/RCC-Project.

Mobile health (mHealth) smartphone applications are a rapidly growing field, and one which is dominated by large, for-profit corporations. These apps employ theories of digital behavioral technology, gathering user data as a means to understand and ultimately influence user behavior. Due to the fact that corporations aggregate deeply sensitive personal health data and actively promote shifts in behavior towards normalized ideals of health, a power imbalance is created between mHealth app users and the corporations that control them. My STS research paper applied feature analysis to understand the mechanisms and conditions through which the features of the six top-grossing mHealth apps on the Google Play Store afford actions to their users. App store descriptions were a primary source of data for my analysis, allowing me to identify both the core features of the app and how the developers motivated feature use. I found that app features overwhelmingly required habitual self-tracking, which was motivated by appeals to affect (specifically, being "healthier"). However, the nature of app features themselves indicated that concrete, positive impacts on health outcomes were largely secondary to the continued collection and analysis of user data by the app. Therefore, I argued that whether a user can actually realize a perceived benefit from the app is dependent on whether the user is one of a normative, privileged few who can participate in the healthy activity defined by the app and

navigate its patterns of manipulation without experiencing negative impacts or being exploited for commercial gain.

Working on my technical project and my STS research paper cemented my interests at the intersection of algorithmic systems and the social impacts of technology, which I am excited to pursue further in my PhD. Although the two are quite distinct from each other, my technical project provided me with useful insider knowledge to apply in my STS research paper. In turn, my STS research paper prompted careful reflection upon my technical project, considering how the project itself encodes normative understandings of the world and projects them upon individual end-users. Overall, I valued the opportunity to engage in both technical and social science research over the course of my thesis. I think that in doing so I was able to form a greater understanding of how both of these aspects intersect with each other, despite the tendency within engineering to consider the social aspect as something entirely separate.