Digitizing a Lifetime of Genealogy: Creating a Searchable NLM Dataset
(Technical Paper)

Trust in the Machine: An Actor-Network Analysis of Privacy in Apple Intelligence
(STS Research Paper)

An Undergraduate Thesis

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

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

Riley Immel
Spring 2025

Department of Computer Science

Portfolio Table of Contents

Sociotechnical Synthesis

Introduction

My capstone and STS research project are mildly connected as they both involve artificial intelligence (AI) as a key component; however, their similarities end there. My capstone focuses on an app that I plan to build whose purpose is to digitize and upgrade my grandfather's genealogical work of our family tree and history. I chose this project as currently, his life's work exists only as paper copies in a small attic in Wyoming and is therefore hard to access for the rest of the family; additionally, I wanted to make it understandable to all of us by including natural language processing. My STS research project focuses on Apple's new in-house AI called Apple Intelligence and focuses on an STS analysis of this innovative technology through the lenses of privacy, data, and user trust.

I chose these topics because both involve transforming human experiences—whether ancestral memories or everyday device interactions—through technical design, raising ethical questions about data handling and user empowerment. STS informs engineering practice by highlighting how human and non-human actors co-construct systems of trust, guiding engineers to anticipate social impacts alongside technical functionality. In this synthesis, I briefly describe my STS inquiry into Apple Intelligence's hybrid AI architecture and my technical work on an OCR-driven, natural-language genealogy search tool.

STS Research Project

In my STS research, I applied the Actor-Network Theory to examine how Apple's design decisions—processing AI tasks on-device versus offloading them to a "Private Cloud Compute" (PCC) service—shape user trust and privacy expectations. By treating users, regulatory bodies, hardware

modules, and cloud servers as actors in a sociotechnical network, I traced how on-device processing reduces data exposure while the usage of the PCC introduces points of vulnerability and dependency on Apple's opaque safeguards. I analyzed official documentation, independent security audits, and privacy-advocacy reports to uncover tensions between Apple's "no peeking" promises and the realities of third-party partnerships (e.g., OpenAI integration) and proprietary code validation. My findings highlight that robust privacy by design demands not only encryption and ephemeral data handling, but also transparent verification and ongoing independent review to sustain user trust.

Technical Project

The technical portion of my project produced an overview of a modular pipeline-based app that I could implement that converts my grandfather's handwritten genealogical research into a structured, searchable database with a natural-language query interface. It uses image preprocessing (noise reduction, contrast adjustment) to enhance OCR accuracy using Tesseract, then parses text into normalized records capturing names, dates, and relationships. These records would populate a relational schema database that is designed for hierarchical family-tree queries. Finally, I would integrate a lightweight NLP layer—leveraging Stanford CoreNLP components—to translate conversational questions (e.g., "Who emigrated in the 1800s?") into SQL queries. Preliminary benchmarks demonstrated a noticeable reduction in transcription errors after preprocessing, and planning talks with early family members indicated that a conversational interface would significantly lower the barrier to exploring complex ancestral links.

Conclusion

Working on both the STS research and the technical project underscored that engineering solutions cannot be solved without considering their social, and indeed ethical, contexts. While the genealogy app teaches the power of technical precision and user-centered design to preserve and democratize access to historical data, the STS study reminds me that even well-intentioned features must be scrutinized for hidden privacy trade-offs and trust dependencies. Together, these projects have taught

me to design with a dual lens: ensuring functionality and reliability while proactively mapping the network of stakeholders and values that will judge, adopt, or challenge my work. This synthesis of technical innovation and sociotechnical reflection will guide my future practice as an engineer committed to ethically informed, human-centered technology development.