

Thesis Project Portfolio

Knowledge Graphs: Powering Intelligent Search

(Technical Report)

From Profit to People: Reimagining Data Governance Beyond the Corporate Model

(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

Trishal Muthan

Spring, 2025

Department of Computer Science

Table of Contents

Sociotechnical Synthesis

Knowledge Graphs: Powering Intelligent Search

From Profit to People: Reimagining Data Governance Beyond the Corporate Model

Prospectus

Sociotechnical Synthesis

In both my STS research and technical project, I set out to better understand the methods by which the data we produce is created, controlled, and used. While both efforts differed in their approaches, they were motivated by the same central focus: data's immense power to inform and influence society and the need for responsible systems to guide its use. I was drawn to this topic because of the growing gap I began to see between society's technological capabilities and how thoughtfully we can manage the consequences of those capabilities. In my technical work, I explored how knowledge graphs could be used to store and analyze real-world data relationships. In my STS research, I investigated how current approaches to data governance fall short ethically and structurally, and how they might need to be reimagined to serve the public interest better. Together, these projects allowed me to examine not only the technology behind modern data systems but also the values embedded within them, an approach I believe is essential to responsible engineering practice today.

The technical portion of my thesis produced a full-stack web application built on top of a custom knowledge graph that allows users to search for entities extracted from web-scraped news articles and view their relationships with other people, places, and organizations. I worked as part of a team of other interns to develop this system using AWS cloud services for data ingestion, Spring Boot for the back-end API, and Angular for the front-end user interface. We implemented a real-time pipeline to properly extract, structure, and load entity data into a graph database, enabling users to interact with rich, interconnected data through a visually accessible application. Our goal was to demonstrate the value of knowledge graphs in organizing and contextualizing information, especially for use in large language model pipelines. By the project's conclusion, we had delivered a working tool that was not only technically robust but aligned with the company's focus on scalable, intelligent data services.

In my STS research, I explored the ethical and societal implications of modern data governance through the lens of classical moral theory. I examined how data practices, particularly when driven by corporate interests, can undermine individual autonomy, erode public trust, and reinforce existing inequalities. The research emphasized that data is not just a technical resource but a deeply political one, shaped by the values of the entities that manage it. I argued for a shift away from corporate-centric approaches and towards governance models that prioritize democratic control and public accountability. By framing data as a collective resource rather than a commodity, we can begin to imagine systems where ethical values are embedded into the infrastructure of data practices. The goal of this research was not simply to critique current practices but to envision alternatives that better reflect the needs and rights of individuals in our digital society.

Working on these two projects in parallel has deepened my understanding of data, not just as a technical resource but as something inherently political and social. Building the knowledge graph made clear how powerful data can be when efficiently structured and linked. However, that same power can be used to surveil, manipulate, or harm. Looking at the issues at hand through STS perspectives taught me that data is never neutral and that it is always shaped by the social, technical, and institutional systems in which it is produced and used. What gets collected, how it is processed, who has access, and what purposes it serves are all questions that carry ethical weight. Engineering does not exist in a vacuum and every design decision we make has the potential to reinforce or resist unjust systems. These projects showed me that to be a responsible engineer is not just to write good code, but to understand and care about the systems that code fits into. In a world increasingly shaped by data, engineers play a critical role in

deciding how that power is used. It's our responsibility to design systems that remain fair and accountable to the communities they impact.