Thesis Project Portfolio

Knowledge Graphs: Gaining Deeper Insight into Open-Source Data

(Technical Report)

Exploring How WMATA's Metrorail System Can Be Improved

(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

Emmitt Sun James

Spring, 2025

Department of Computer Science

Table of Contents

Sociotechnical Synthesis

Today, a vast amount of open-source data is generated every single second, presenting a significant challenge in its collection and management. To address this, my team and I implemented a knowledge graph based solution, providing a structured representation of entities and their complex relationships. This solution was then hooked up to a web application to allow human users to have an easier time analyzing open-source data. The human and social dimensions of this technology are very important for a couple of reasons. For one, data security and privacy are very important, and it would be wrong to ingest unethical data such as data found in leaks. This technology would also be used by humans, meaning it would have to present data in a way that is easy for a human to understand.

While working on this project, my team and I had to commute to our company's office in Northern Virginia. This led several of us to rely on the public transit options provided by the Washington Metropolitan Area Transit Authority (WMATA), particularly the Metrorail system. Given that the Metro is essentially a large piece of infrastructure, I found it fitting to analyze the system using Susan Star's properties of infrastructure. The research consisted of applying these properties of infrastructure to the Metro to discover how WMATA could improve ridership across the system. To conduct this research, I will interview WMATA employees. The questions I will ask will focus on specific properties of infrastructure as defined by Star. To supplement this, I will be analyzing publicly available agency reports such as the FY2026 Proposed Budget. Through my research, I found that WMATA must address three key areas to increase ridership. First, securing a more robust and reliable funding structure with dedicated annual funding is crucial for long-term financial stability and capital improvements. Second, enhancing connectivity by expanding and improving access to the Metro will encourage riders to utilize the system. Finally, improving service frequency and reliability by running more consistent and frequent trains will reduce wait times, making Metro a more competitive and convenient transportation option. My research and my capstone project are not directly related but they do have some connection to each other. My capstone project is about the work I did in my most recent internship, which I rode the Metro to get to. By improving the Metro, software engineers in the DC area will be able to have an easier commute to work and be more productive.