CSC8499: Project and Dissertation for
MSc in Advanced Computer Science

A Website Providing Metro Insights

Interim Report

By
Aditi Jain

1. Introduction

The Newcastle Metro is a complex system that requires meticulous management to ensure smooth and efficient operations. Metro staff have numerous responsibilities, including scheduling rotas, maintaining the cleanliness of stations and trains, monitoring for defects, and ensuring the overall safety and reliability of the metro services. The metro passengers also require a lot of information to plan their journey accordingly and have a good travel experience.

The goal of this project is to deliver actionable insights that will enable metro employees in their day-to-day work and passengers have a pleasant metro journey experience by utilising the abundance of current Newcastle Metro data, including metro timetables and passenger flow statistics. By utilizing Machine Learning, I aim to transform raw data into meaningful, interactive statistics presented through a user-friendly website. The website will feature dynamic tables, graphs, and images, enabling staff and passengers to visualize and interpret data with ease.

Understanding the significance of data in today's digital age is crucial. Proper data management can drive significant improvements in operational efficiency and service quality [1]. This research highlights the critical role that data plays in improving business processes by showing how data-driven decision-making can result in better resource allocation, optimised timetables, and enhanced passenger experiences.

Human-computer interaction (HCI) is a study of how humans use computers to perform certain tasks and use it in such a way that the interaction is being enjoyed and effective [2]. This project highlights the positive impacts of HCI. By designing interfaces that are intuitive and engaging, the aim is to bridge the gap between complex data sets and the metro staff and passengers who rely on them. Improving the user experience for the employees, currently having to figure out their tasks using complex data sets, they will be able to see visualizations and valuable information just by opening a website. Effective HCI principles ensure that the website is not only functional but also enhances the user experience, making it easier for staff and passengers to access the information they need when they need it.

2. Aim and Objectives

The aim of this project is to develop a website to display visualizations and insights derived from metro data, tailored specifically for metro staff and passengers to enhance their operational efficiency and decision-making processes.

Objectives:

1. **Identify Relevant Data Sources:**
    - Determine and collect the types of data that can be leveraged to derive meaningful insights to improve user experience, including metro schedules, passenger counts, maintenance logs, and staff schedules.
2. **Investigate Existing Systems and Requirements:**
    - Conduct thorough investigations into existing systems and processes.

- Gather staff requirements through interviews and collect passenger feedback through surveys to understand pain points and areas for improvement.

3. **Website Development:**

   - Develop a user-friendly website using HTML, CSS, and Python.
   - Ensure the website is capable of fetching the output from the data pipeline and effectively displaying it in various formats such as tables, graphs, and interactive visualizations.

4. **Machine Learning Research and Implementation:**

   - Research and implement appropriate machine learning algorithms to train on the identified data.
   - Automate the derivation of meaningful insights, such as predicting peak travel times and identifying maintenance needs.

5. **Data Pipeline Creation:**

   - Develop a robust data pipeline that includes stages for data pre-processing, transformation, training, and visualization.
   - Ensure the pipeline is efficient and capable of handling real-time data updates.

6. **Address Staff and Passenger Issues:**

   - Explore data-driven approaches to address the issues and challenges faced by both staff and passengers.
   - Implement solutions that improve operational efficiency and passenger satisfaction.

7. **Evaluation**

   - Test the website and efficiency of the visualizations on the website. To achieve this, I plan to set up persona and create a user journey map to get fellow students to step into the intended user's shoes and provide valuable feedback.

3. Overview of Progress

3.1 Research

Through reviewing various research papers of existing systems out there for enhancing user metro journey experience, I learned several key insights validating that a user-centric, data-driven approach is an effective way to develop this project. One noteworthy finding was how the ‘spatial characteristics’ of underground metro stations influence passenger comfort and safety, as shown by research in Warsaw which identified key architectural features impacting user experience [3]. Additionally, study on travel mode choices in cities like Doha presented the factors affecting commuters’ decisions, such as socioeconomic characteristics and trip conditions, which is derived using advanced statistical and machine learning models [4]. Likewise, research on optimizing metro operations through timetable scheduling and passenger flow control, as conducted in Beijing, highlighted the importance of data management to reduce congestion and improve service quality [5]. Moreover, research on transit performance measures using data-driven methods for reliability improvement, as demonstrated in study

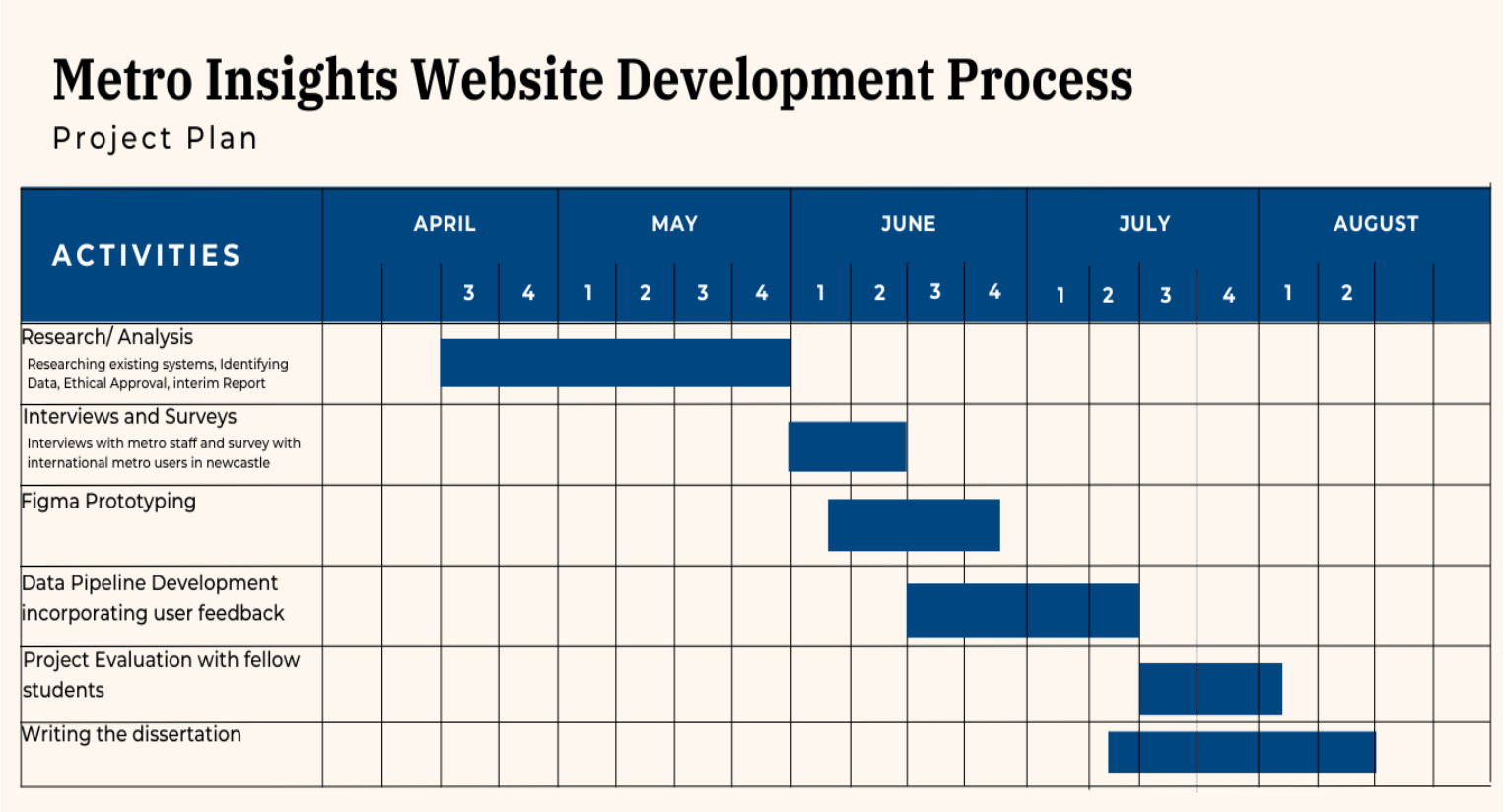
from Minnesota, highlighted the importance of addressing traffic delays and boosting operational reliability [6]. After going through the existing systems and existing researcher's works, I decided to develop a robust data-driven website aimed at enhancing metro operations and user satisfaction.

3.2 Development

I will be utilizing 'figma' for prototyping the website, which will help in evaluating the website and leaving time for me to focus on incorporating user feedback and working with data. Machine learning algorithms will be used to train the data to provide meaningful insights.

3.3 Progress till date

1) Received ethical approval for the project 'Metro Insights Website'.
2) I plan to interview metro staff and survey international people using Newcastle metro. The questionnaire and relevant information documents have been prepared for the same.
3) I have identified some data that can be useful to fetch insights from, will finalize after the interviews and make changes according to survey answers along the way.



Time spent on tasks is defined in weeks, which provide rough estimations which are not constrained by exact dates – allowing for a flexible, yet measured interpretation.

References

[1] B. Chaudhary, “Importance of Data (A Term Paper)”, doi: 10.13140/RG.2.2.23837.69602.

[2] H. Bansal and R. Khan, “A Review Paper on Human Computer Interaction,” *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 8, no. 4, p. 53, Apr. 2018, doi: 10.23956/ijarcsse. v8i4.630.

[3] K. Jasińska and D. Kłosek-Kozłowska, “Passengers’ experience in underground non-transfer metro stations: The impact of spatial characteristics,” *Tunnelling and Underground Space Technology*, vol. 143, Jan. 2024, doi: 10.1016/j.tust.2023.105482.

[4] A. Abulibdeh, “Analysis of mode choice affects from the introduction of Doha Metro using machine learning and statistical analysis,” *Transp Res Interdiscip Perspect*, vol. 20, Jul. 2023, doi: 10.1016/j.trip.2023.100852.

[5] Y. Hu, S. Li, Y. Wang, H. Zhang, Y. Wei, and L. Yang, “Robust metro train scheduling integrated with skip-stop pattern and passenger flow control strategy under uncertain passenger demands,” *Comput Oper Res*, vol. 151, Mar. 2023, doi: 10.1016/j.cor.2022.106116.

[6] E. Wong and A. Khani, “Transit Delay Estimation Using Stop-Level Automated Passenger Count Data,” *J Transp Eng A Syst*, vol. 144, no. 3, Mar. 2018, doi: 10.1061/jtepbs.0000118.