COSC545 Project Proposal

Tyler Duckworth[1] , Connor Riley[1] , Jake Seaton[1] , Ryan Trenner[1]

I. INTRODUCTION

Air pollutants are a significant health hazard, and automobiles are one of the highest emitters of those. The EPA has programs to study the emissions near roads, but the data for those programs is often hard to find. We want to fix that and create a tool that allows for faster and more efficient data analysis.

II. MOTIVATION

Millions of people in the United States live close to busy roads. As automobiles emit a number of air pollutants, there is a risk of negative health impacts on those who are regularly exposed to the pollutants emitted from these vehicles. Airborne pollutants have already been linked to increases in lung and heart issues.

In 2010, as part of a review of the National Ambient Air Quality Standards (or NAAQS), the U.S. Environmental Protection Agency began monitoring emissions near high-traffic roads in populous cities in the United States to determine what kinds of pollutants were being emitted and at what levels. This Near Road monitoring program involved the creation of numerous sites at key roadways, with 75 sites currently active at the time of writing this proposal.

While the program has been running for many years, accessing and visualizing the data from the Near Road program is limited to visualizations created by the EPA. Furthermore, assembling data from different sites and different pollutants can prove difficult as researchers need to navigate a fairly complex REST API or ArcGIS maps to collate this data.

III. OBJECTIVE

Our goal for this project is to make it easier for researchers to collate and analyze data from multiple Near Road sites in this program. We aim to create a dashboard that displays an interactive map to be used to perform data analysis on the historical pollutant data for the sites maintained by the EPA for its Near Road monitoring project.

IV. DATA TO BE USED

For this project, we will use the publicly available data from the EPA in their AirData repository. This data comes in many forms, from pre-generated comma-delimited files, to interactive maps, to a REST API that can be used to retrieve site metadata in addition to actual concentrations. In terms of the form of this data, this will primarily be the concentrations of different pollutants. Some examples of pollutants are as follows:

- $PM_{2.5}$ - Particulate matter with a diameter of 2.5 micrograms or smaller, meaning it can penetrate the deepest parts of the human lungs.
- O_3 - Ozone, a common pollutant that is a byproduct of chemical reactions between other air pollutants.
- NO_2 - Nitrogen Dioxide, a compound linked to negative respiratory health effects, such as infection and asthma.
  - Similar compounds, such as NO and NO_x are also measured due to their similar effects

Each of these pollutants is frequently measured in different units such as $\mu g/m^3$, parts per million, parts per billion, and more.

V. RESPONSIBILITIES

Based on some internal discussion, we decided to split into two groups: a frontend team and a backend team. The frontend team will work on the front-facing portion of the application. This includes the map, dashboard, and other systems to allow for the visual display of the data. The backend team will work on creating a system to collect the data from the AirData API, store it, and host an API that the frontend can use to access that data quickly and build complex queries.

- **Frontend Team**
  - Tyler Duckworth
  - Jake Seaton
- **Backend Team**
  - Connor Riley
  - Ryan Trenner

VI. TIMELINE

The timeline for this project can be split into a three phases: Ideation, Iteration, and Polishing.

A. Ideation

The first phase, Ideation, will take the first few weeks after this proposal is submitted. This phase will consist of assembling a rough outline of the application and working to build a very bare-bones prototype. This prototype will be stripped down to a few key features, such as a simple map representation and basic data processing on the backend. We will also be choosing what tools and technologies we are using and working on outlining as much of this application as possible so we can spend most of our time developing.

---

[1]All the authors are students in the Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN

B. Iteration

The second phase, Iteration, will be where the majority of development takes place. We will be working in our teams and iterating on the frontend and backend respectively. Our end goal is to have a fleshed out proof-of-concept that hits every major feature we have outlined. This app will be mostly ready to submit. This phase will take up the majority of the semester and be divided up into sprints.

C. Polishing

The final phase, Polishing, will mostly consist of bugfixes and the creation of the final deliverables. This includes finalizing the paper and creating a slide deck to showcase the application. This phase will last the last few weeks of class and end once we give our final presentation and turn in the application.

VII. EXPECTED OUTCOME

We expect that we will finish the proof-of-concept for the application that, with some slight bugs, is ready to present to the class. This app will consist of a fleshed out backend that hosts a REST API and processes data from the EPA and a frontend that allow for sophisticated data analysis. We would like the ability to show data for all the Near Road sites in the U.S., however to ensure our goals are more attainable, we have selected three sites as a floor in case time gets ahead of us. We may also limit the pollutants we allow for display and analysis depending on how many are available.

VIII. CONCLUSION

Overall, we seek to create a dashboard to display the air pollutant data from the EPA's Near Road monitoring program in order to more effectively allow for that data to be analyzed. We have split into two subteams to each address the frontend and backend of this application respectively.