AIRNOMINAL

AN OPEN PLATFORM FOR COLLECTION AND COLLABORATIVE SHARING OF AIR QUALITY DATA

AUTHORS: FILIP ŠTAMCAR, JAKOB KRALJ

SUPERVISOR: ALENKA MOZER, M. SC.

GIMNAZIJA VIČ

LJUBLJANA, MARCH 2021 – MARCH 2022

Table of Contents

Table of Figures

ABSTRACT

The problem of air pollution is one of the most acutely dangerous ecological problems we face today. According to studies made by World Health Organization, almost all of the global population breathe air that exceeds WHO guideline limits containing high levels of pollutants. Even though sensors and other measurement equipment is already available, data that researchers use usually comes from two to three stations in an area, and large datasets of high spatial resolution are not available. To test if any solution addressing air quality is in fact effective, researchers need to be able to obtain many higher quality data in a standardized format. To address this problem, we decided to build an open platform of air quality measuring stations that enables people to connect and exchange air quality measurements and design an open measuring station that anyone can build and connect to the platform, thus providing valuable information for the research and limiting of the main causes of high PM particles concentration.

Keywords:

- Air pollution
- Air monitoring
- Air quality
- Public health
- Climate change
- Ecology
- Network
- DIY

INTRODUCTION

Of all the ecological problems we face today, very few are to more acutely dangerous than the problem of the air pollution. According to studies made by World Health Organization, almost all of the global population breathe air that exceeds WHO guideline limits containing high levels of pollutants. Air pollution is recognized as a risk factor for many diseases, including ischaemic heart disease, stroke, chronic obstructive pulmonary disease, asthma and cancer. It is estimated that diseases attributed to air pollution are now on par with other health risks such as unhealthy diets and tobacco smoking, and believed that this problem alone accounts for more than 7 million death annually (that is more people than the population of 12 EU member states). [1, p. 7] [2]

The problem has been known for a number of years, and sensors and other measurement equipment is cheap and readily available. For a lot of children, us included, collecting air measurement data may have been the first Arduino project we had done. Even though, data that researches use still usually comes from two to three stations in an area, and large datasets of high spatial resolution are not available. To test if any solution addressing air quality is in fact effective, researchers need to be able to obtain many higher quality data in a standardized format.

With that problem in mind, we decided to build an open platform of air quality measuring stations that enables people to connect and exchange air quality measurements, including humidity, temperature, PM particle concentration and NO_x compounds, and design an open measuring station that anyone can build and connect to the platform, thus providing valuable information for the research and limiting of the main causes of high PM particles concentration.

Additionally, this project provides initiative for young people to learn programming, soldering and construction skills, and improve their knowledge about the air pollution and air quality measuring.

AIR POLLUTION

The air pollution is one of the most acutely dangerous environmental problems. Most of the global population breathes air that exceed WHO guideline limits, with the low- and middle-income countries being the most affected. It is recognized as an important source of morbidity and mortality, and contributes as a risk factor for many diseases, including heart diseases, chronic obstructive pulmonary diseases and acute respiratory infections. Indoor and ambient air pollution combined are estimated to cause around 8 million premature deaths worldwide per year. [3] [4]

Air pollution is caused by any chemical, physical or biological agents, which modify air characteristics. Main sources of air pollution are domestic combustion devices, industrial facilities, and motor vehicles. Major pollutants include particulate matter (PM), carbon monoxide (CO), nitrogen dioxide (NO_2), sulfur dioxide (SO_2) and ozone (O_3).

Particulate matter (PM) is a common indicator for air pollution and affects more people than any other pollutant. It mainly consists of sulfate, nitrates, ammonia, sodium chloride, black carbon, mineral dust and water, of a complex mixture of solid and liquid particles suspended in the air. It is commonly categorized into different PM_x categories, where "x" defines the maximum particle diameter in µm. PM_{10} particles most commonly affect eyes and throat, $PM_{2.5}$ particles are even more dangerous as they can travel through the lungs into the blood system. When more sensitive measurement tools are available, particles are commonly divided into additional categories, enabling a better particle analysis. The smaller they are, the deeper into respiratory and blood system can they penetrate, causing bigger

and bigger risk. Chronic exposure to particles contributes to cardiovascular and respiratory diseases and lung cancer. [4] [5]

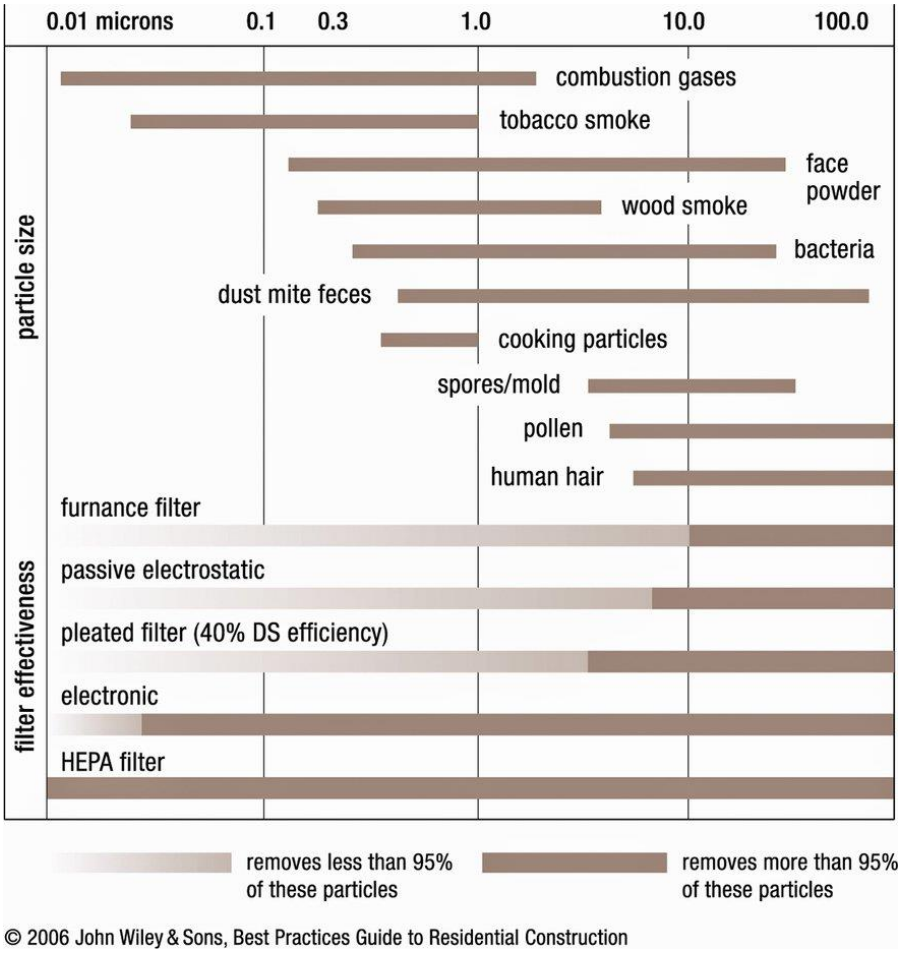

Figure 1: Size range of common pollutant sources[6]

PROJECT OBJECTIVES

WHAT IS OUR MAIN PURPOSE?

We want to design an open platform of air quality measuring stations that allows researchers to exchange and study air quality data. The platform alone does not impose any restrictions on specific measuring device used, allowing usage with a wide range of sensors and making it more accessible. It is possible, for example, to connect with a normal stationary measuring device, or cheaper and lighter portable devices that people can take alone and thus enable a wider data coverage. In the Environmental Pollution Research in Schools Erasmus+ project, which the Airnominal platform is part of, we also want to design a lightweight aircraft that will be able to measure air quality on different altitudes, and encourage local communities in installing measuring stations at their homes and public places. With the help of experts from the University of Nova Gorica and National Environmental Agency of Slovenia, we will compare our data with the professional measuring stations and evaluate their accuracy and reliability.

With an open platform for sharing air quality data, we hope to enable researchers to test if any solution addressing air quality is really effective, allowing them to focus only on working solutions and improving the local quality of air, protecting and improving the health of people.

To allow people to easily take part in the project, we want to provide a reference implementation of the measuring station both as a finished product that people can buy and as a DIY project. The station will collect current location and air quality measurements and sent them to our platform using Wi-Fi module, or, where Wi-Fi is not available, through other means of wireless communication. Our buy version will allow even people with little knowledge about the topic to take part and help, while the DIY version will be able to improve development of fine motor skills as well as programming and construction knowledge.

WHAT PROBLEMS DO WE ADDRESS?

Air pollution is one of key factors that affect public health and climate changes, however, as the professional measuring equipment is very expensive, it is hard to measure it on a very large scale. With our platform, anyone can measure and exchange measurements into a public dataset, allowing hobbyists and researchers to easily collect obtain a large quantity of data to analyze them. As our platform is open source, researchers can not only control and deploy it, but also customize it to suit their own needs. It does not restrict users to any specific measuring station and sensors (although our pre-built version is provided as a reference), allowing them to add their own sensors, measure more reliable data or reduce the cost and install more stations, depending on their needs. The service is designed to work with affordable microcontrollers such as Arduino Uno, ESP8266 and similar, allowing anyone to contribute to science for a very small cost. The system is free and easy to use for novice and experienced programmers alike. We provide a large number of ways to access data, so less time needs to be spent cataloging.

Existing professional measuring installations can provide more reliable data than our sensors, but only if we account for data from a single measurement device. However, their installations commonly only provide areal data and do not have complete area coverage. In comparison, our measuring data allows individuals to build and place measuring systems at their homes or other places, choose the location according to their needs or install multiple stations if needed. This is especially important for PM measurements as their can differ drastically even on short distances. As the project progresses, more and more stations will be installed, increasing the density and coverage. With bigger density, our platform can provide more reliable and useful data than the existing solutions.

Measuring stations can then be installed to public places and institutions, such as schools, libraries and bus stations, allowing us to collect a large amount of air quality data. Public institutions such as libraries can also take part in the project and raise awareness of the air quality problems and their effects on our health. It can allow researchers and people in charge of addressing the air pollution problem to use large public datasets to test if their solutions are effective and take actions accordingly. The product also allows people to have better knowledge about their environment, encourages them to take part in solving the air quality problem in local communities, and educates them about the air pollution problem.

Additionally, the openness of our platform with support for custom stations encourages people to design their own products, learn about relevant topics and improve their programming, soldering, construction, and fine motor skills. Our project allows schools and other educational institutions to work actively on improving these skills.

Existing Alternatives

There are already many consumer-grade weather stations on the market to measure basic air parameters, however, most of them cannot collect advanced air quality measurements such as PM particles and connect to the internet. There also exist some commercial solutions, which are rather expensive and use proprietary collection systems, making harder for researchers to use the data.

Current alternatives for measuring air quality include:

- WAQI (https://waqi.info/)/AQICN (https://aqicn.org/): Use only data from the official sources, providing more reliable data yet with more limited coverage. It is not open source.
- PurpleAir (https://www2.purpleair.com/): Manufacture their own professional sensors that are very expensive for normal people. Do not provide complete public data and their platform is not open source.
- EnviroMonitor (https://enviromonitor.github.io/): An open source project for measuring and collecting data, however, not yet launched. Only supports a single type of microcontroller (ESP).

Other services are either only for enterprise or non-public use, or do not collect air measurements primarily. In contract to most existing alternatives, our project is completely open source and extendable, allowing anyone to connect and exchange data. It can be used with affordable microcontrollers, without requiring access to expensive measurement equipment. As the project also allows using custom DIY measuring station, it provides an initiative for people and local communities to learn about and take part in air quality measuring, as well as improve their programming and fine motor skills.

Our project is also unique in a way that it encourages individuals to install measuring devices where they want. As the project progresses, more and more people will install their devices, therefore increase the density. In comparison to the existing solutions on the market which either only provide very few stations per area or only allow access to data to specific customers, our platform provides both an open access to data and accurate and dense data.

Air quality sensors, specifically PM particles detectors, are already accessible on the market. However, most of sensors do not come with a full measuring device and exchange data to the public, limiting reach of the data. In contract, our platform allows easy set up to connect various sensors and globally exchange data. As more and more sensors are added to the platform, data coverage is increased, mitigating the accuracy issues caused by less accurate while more accessible sensors in comparison to very few expensive professional devices.

How It Works

Arduino-based Measurement Station

Our measurement stations currently consist of Arduino and ESP microcontrollers, onto which temperature, humidity and PM sensors, as well as GPS module for obtaining station's location, are connected. The station collects the measurements from sensors and GPS module and uses our library to sends them to our server. Initially, we also wanted to measure NO_x and other air quality characteristics, but, unfortunately, we were unable to obtain the required sensors due to

unpredictable nature of shopping and shipping from foreign online stores. However, as our project is very extensible, it is easily possible to also collect those data once we acquire sensors.

Our users are also encouraged to design and build their own measurement stations using Arduino-compatible microcontrollers and our development library to suit their needs.

Arduino-compatible Library

The project provides a library for Arduino-compatible microcontrollers that allows a wide range of supported measuring devices and sensors. It is designed to be lightweight and extensible, so it can run on many different platforms.

When using an official DIY kit or pre-built product, the library is pre-configured to collect data using the default sensors (temperature, humidity, PM and location), and send them to the official public instance of the data server through Wi-Fi network (by default) or other wireless communication protocols (if configured by user). For cases where no wireless connection is available, data are stored locally on the microSD card and can later be manually imported to the service.

The library also allows makers of custom measurement stations to extend them with their own sensors and hardware through a simple-to-use code interface, allowing a wider range of data and providing benefits for researchers and other individuals with specific needs. It is also possible to set up other means of wireless communication, such as mobile network or radio signals to allow usage where Wi-Fi is not available or reliable. Users can also set up library to deploy data to their private instances, however, we encourage everyone to connect to our public service so more people can benefit from it.

Communication HTTP API

The layer connecting all parts together is a well-documented and easy-to-use HTTP API which exchanges data from measuring devices to our server and provides the measurements and other information to the user.

The API is internally used by the Arduino library to send measurements from sensors, and by the website to display them in a human-friendly format, but it can also be used directly by advanced users to suit their needs. The API is already designed to be very extensible, support any measurement and sensor types and allow exchanging data both in a MessagePack (https://msgpack.org/) – very efficient, mainly suitable for less powerful microcontrollers – and JSON (https://www.json.org/) – very popular, human-readable, suitable for usage by the website or other viewing platforms – serialization formats.

The majority of the API will be publicly accessible, except for calls for creating and managing measuring stations, which will require creating a free account, and sending measurements to a server, which will require providing a station-specific API token to prevent abuse.

Server

The official backend server is created in Python using the Flask (https://flask.palletsprojects.com/) micro web framework and SQLAlchemy (https://www.sqlalchemy.org/) ORM. The server collects data from stations, stores them in a database and provides them to a frontend website, and allows registration, validation and management of the equipment. Because we use SQLAlchemy database toolkit, our server can use a wide range of SQL databases, from simple file-based databases such as

SQLite to powerful database servers such as PostgreSQL. The server is lightweight and can also be self-hosted from equipment as cheap as Raspberry Pi, using any WSGI-compatible web server.

Website

Our website is written in TypeScript (https://www.typescriptlang.org/) using Vue (https://vuejs.org/) framework and Vuetify (https://vuetifyjs.com/) theme which implements Material Design guidelines. Using Vue allows better modularity as it provides a support to develop reusable components as custom HTML elements.

The website provides user instructions for setting up own measurement devices and registering them to a server. As the project is still in the early stages, registering a station does not require an account, and users currently cannot modify or remove the station after its creation. After registering it, they just obtain a station config which needs to be manually configured at the station. Other features of the website already work as expected.

Once the website is fully developed, users will be able to create an account for free and then register their own measurement devices or manage their existing ones (renaming or deleting them, adding new measurement types, etc.). When registering a device, they will obtain a provided station-specific API token and configure it at the station. It is planned to also provide an easy-to-use configuration program that allows users with less technical knowledge to automatically configure their stations (setting up Wi-Fi network, adding station key, etc.). Through a dashboard, users will also be able to manually import data that were stored on a microSD card, in case network was unavailable when they were running measurements.

Data are public and do not require user registration. They are obtained from the backend server through an API and displayed in various formats. Users can view a map of all currently active stations or pick a specific one to view all historical data. The station-specific page also displays current measurements, historical data using easy-to-understand charts as well as current and all past locations on a map. It is possible to also select multiple stations to compare their data.

Researchers that need to analyze a large amount of data can directly download historical data through an API. The server also allows restricting data to a specific date range, station or measurement type, so they do not have to worry about filtering data and can focus on their research. This functionality can be very useful for researchers or other individuals or organizations who need to locally perform more advanced analyses.

Source Code

The project is open source and available on GitHub:
https://github.com/ChristofferNorgaard/Airnominal.

The official instance is currently deployed on the school's servers: https://zrak.gimvic.org/.

How To Use

Registering Stations

Users can access the station registration page with a button in the top-right corner of the website. On the registration page, they can enter details about the station (name and description), as well as specify which sensors the station provides. After submitting the form, they obtain a station config with the API key which needs to be configured at the station.

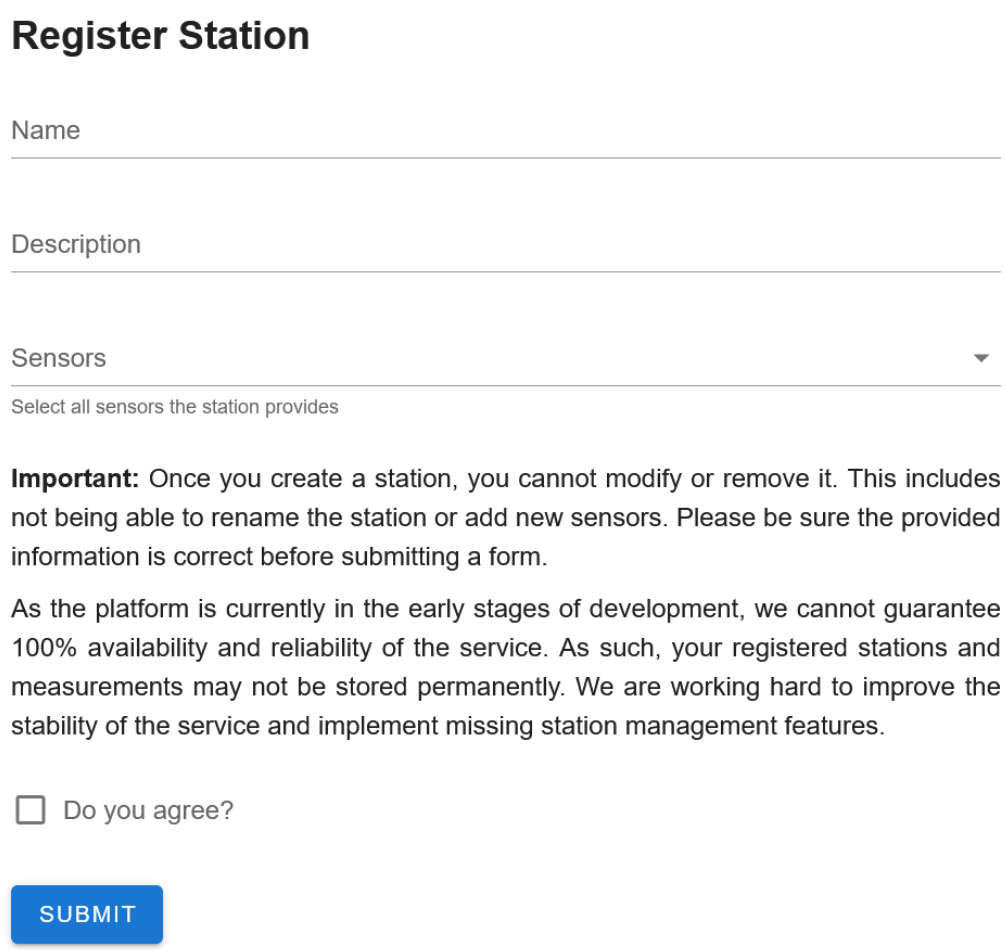

Figure 2: View of the station registration form

Configuring Stations

Users can configure their Arduino-based measurement stations with the help of C++ header file which contains definitions for Wi-Fi network details and the API key. The file currently needs to be manually edited to include the correct data, compiled using Arduino-compatible compiler and uploaded to the board.

As the project develops, we will also provide an easy-to-use configuration program that will allow less technical users to automatically perform those steps, as well as add setup instructions and programs for non-Arduino-based stations.

Selecting Stations

Users can select stations which they want to view from a map or a list of all registered stations. Once the website is fully developed, they will also be able to filter stations based on various characteristics, such as provides sensors, their locations, data history, etc., as well as organize them into groups.

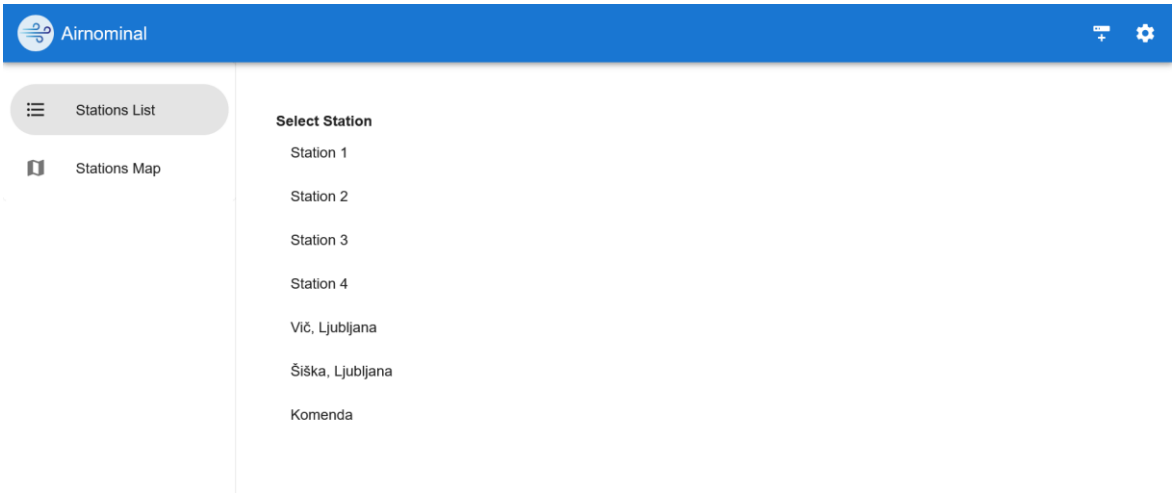

Figure 3: A list of connected stations

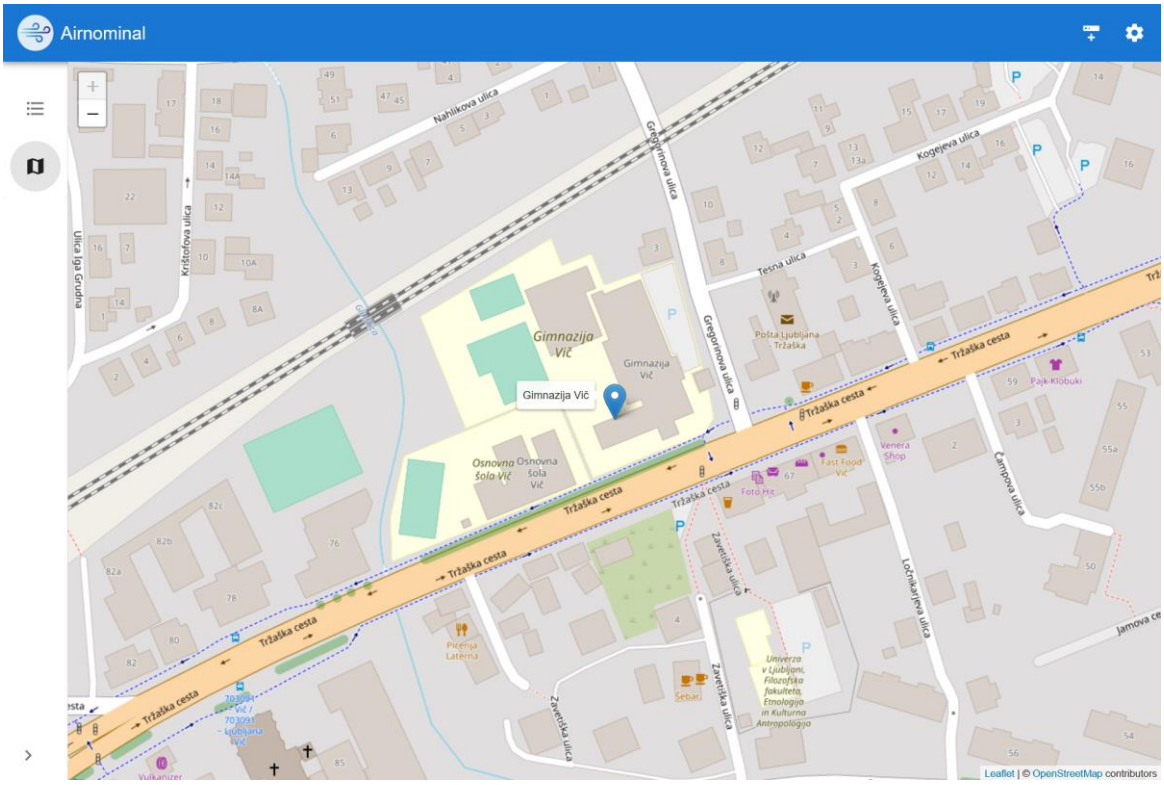

Figure 4: A map of connected stations

Viewing Measurements

One a user selects a station, they can view current measurements, charts with all historical data and map with station's current and past location. It is possible to disable displaying current data, charts or

map through the settings page. Users can also compare data from multiple stations, although this currently requires manually editing the page URL.

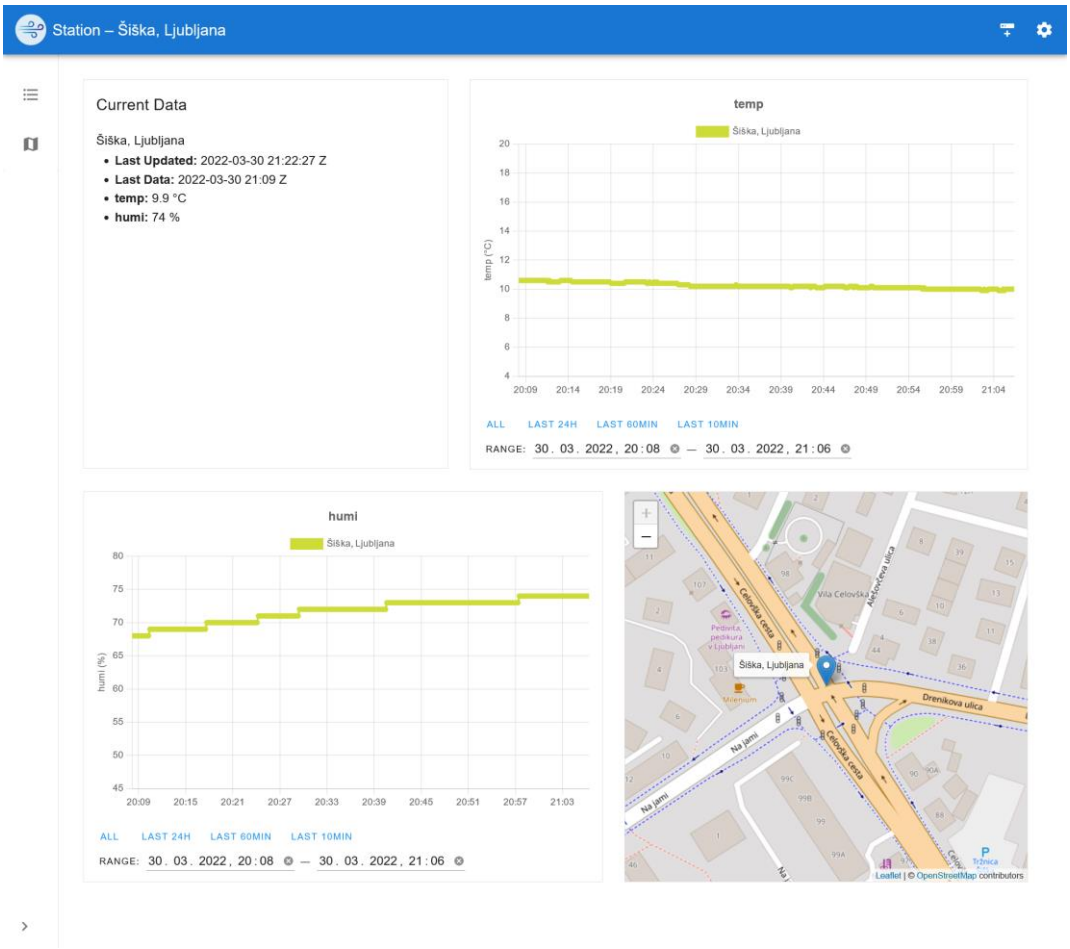

Figure 5: A view of station measuring temperature and humidity

PROMOTION PLAN

We will encourage existing users of the platform to recommend our project to their friends and spread the word in our school and work. As more and more people start using our service, it will start to advertise itself, while also increasing its research value as more and more data will become available on the platform. When the project is shown to be working on a smaller scale and is ready to extend its usage, we will start attending seminars and events about the air pollution, spreading awareness about the topic and promoting our project.

We can collaborate with local public institutions such as schools, universities and libraries to install measurement devices, educate their visitors about the importance of air quality and spread awareness of the danger of air pollution. We can encourage local communities and organizations to start using our platform and educating them about the air pollution. We can connect with hiking clubs and their members to take our portable measuring device with them to the mountains, allowing us to collect data from wider area and research the air pollution on more remote locations.

As part of our main Erasmus+ project, we also plan on creating a lightweight aircraft with sensors attached. We expect this will both provide us and researchers useful data about air quality on different

locations and altitudes, as well as drastically increase the media coverage about the project and promote it to more people.

We will promote both our DIY kit and pre-built product. The DIY kit will mainly be targeted at schools and other educational institutions as it allows them to educate children about dangers of air pollution and improve their knowledge and skills, individuals with interest in programming, soldering and related topics, as well as research organizations that need to customize the station with specific sensors and hardware. The pre-built will be targeted at other individuals to allow them to easily participate in the project without them needing to have specific skills, and other organizations which may be interested to install our product.

We could connect with municipalities and organizations that manage public infrastructure to install our equipment to public places such as bus stops, road junctions and parks. This would increase density of our stations and provide better and more reliable data for the researchers.

Initially, we want to focus on setting up the project in our home country, Slovenia. There are more than 250 high schools, around 90 faculties and more than 30 hospitals and health centers which are our potential customers. In Ljubljana, the capital and the biggest city of Slovenia, there are more than 30 public libraries where our stations could be installed and where we could host seminars about the topic. Additionally, there are more than 450 bus stops in Ljubljana where the measuring equipment could be installed in collaboration with the public company Ljubljanski potniški promet and Javni holding Ljubljana.

Once the project establishes a strong base in Slovenia, we are planning on spreading our work abroad. As the problem of air pollution is global, we believe there is potential both in more and less developed countries. According to WHO, it is especially problematic in low- and middle-income countries, mainly in South-East Asia and Western Pacific regions. [4] Spreading the project to those countries could raise awareness about the air pollution and enable researchers to study the effects and possible solutions to restore healthy air quality. In more developed countries, the project can still take part in various educational and awareness-spreading activities, while simultaneously providing broader data coverage for researchers.

Financial Plan

The main cost of our project comes from hosting a public instance of data collection and viewing server. As our starting budget is low and we expect not to have a lot of users at the start of the project, we estimate we can continue using our school-provided server for free for at least a year. As the project progresses and more and more users start sending measurements to the server and viewing the website, we will need to acquire more powerful server and invest more into reliability and backups in case the server malfunctions.

Money will also have to be spent on advertising the service. We plan to, at least initially, advertise our platform to our friends and invite them to invite their friends, which lets us cut on the expenses. Other advertisements may need to be done using online advertising services as well as achieving media coverage. We also hope we could connect with various research and education institutions and various government and ecological organizations to help as advertise the project and cover some of the project-related costs.

We plan to cover those expenses by selling DIY kits and pre-built products of our reference station. The kit and the product will be using affordable ESP microcontrollers (around $8), PM sensors (around

$10), temperature and humidity sensors (around $4) and GPU modules (around $7). Such, materials for one station cost around $30. While adding our own production costs and accounting for income for our service, the pre-built station will cost around $45, while the DIY version will be cheaper. Considering many sensors manufactures and sellers offer discounts for mass sells, actual material cost may be lower. Thus, specific cost will be determined based on our other expenses and incomes, and availability and price of the required sensors and equipment on the market. When the project goes large-scale, we will need to optimize our production. We may also offer pre-built customizable stations for individuals and organizations with specific requirements.

Data Analyses

We wanted to include and analyze data obtained in few locations in Slovenia. Unfortunately, we were currently not able to measure all types of measurements we wanted, due to unpredictable situation regarding shipping sensors from China, but we were able to obtain at least temperature, humidity and PM measurements.

Figure 6: Data from station measuring PM particles

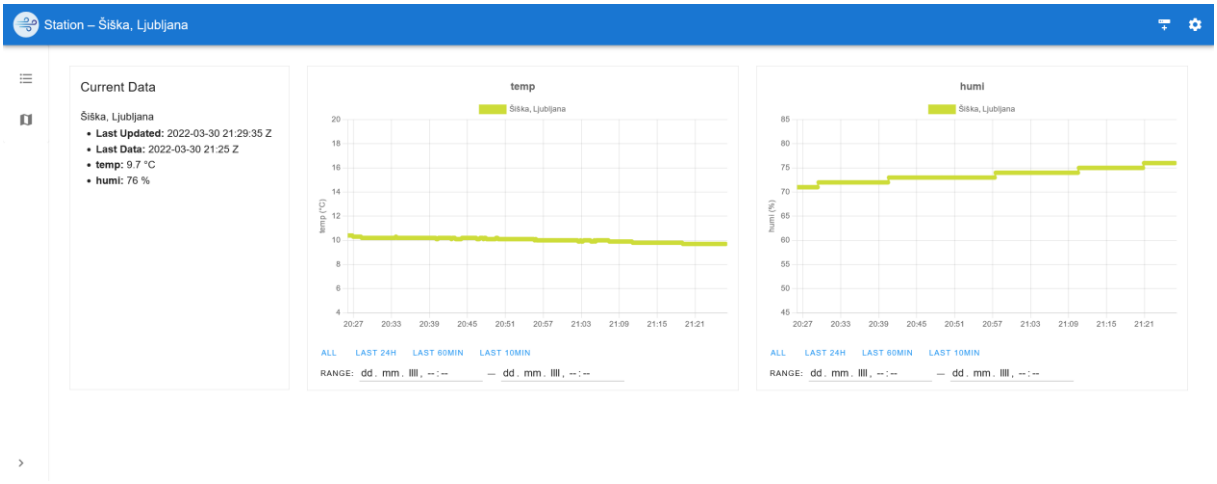

Figure 7: Data from station measuring temperature and humidity

Future Development

As more and more people start using our project, we will have to improve our manufacturing process and find reliable suppliers for sensors and other production materials. We will also have to start producing the product at a larger scale and find a reliable delivers service to bring it to customers.

For the software part, we plan to heavily improve code quality to make it possible for others to understand and contribute to it more easily. In the very near future, we also plan to implement some specific features that we were not able to implement yet due to limited time and make the API more open and well documented. If necessary, we will optimize the platform to allow a bigger number of simultaneous connections from all measurement stations.

As more and more stations will be added to the platform, displaying all of them at once may become impenetrable for researchers that need to focus only on a specific area. To address this issue, we will implement better controls on the website to allow users to only show stations that meet specific criteria and to store and organize stations into different projects/groups. We think these features will become very useful as more stations get connected to a platform. To make it easier to perform analytics on our data, we can integrate existing open scientific and analyses tools directly to our website. For more advanced analyses, we can provide easy-to-use API clients in various programming languages and other tools. We also plan to connect with researchers to see which specific data they need the most and in which format to then optimize the platform to their needs.

In collaboration with research and educational institutions, we can develop educational website to educate our users about the problem of air pollution. In addition to our main website (which is mostly targeted at existing users of the platforms that own a station and researchers), we can create a website that allows less experienced people and general public to view air quality measurements in their local area. This will enable to alert people in case of exceeded concentrations of air pollutants (above the recommended WHO guidelines), educate them about the air pollution, spread awareness on air quality in connection to health issues, as well as promoting our service.

To address the problems of too much load on the main platform servers, we could implement a system of multiple distributed servers, including self-hosted ones and semi-public servers hosted by research institutions, which would regularly synchronize the data to all other servers in the network. With such system, we could keep load on our main servers relatively load while still achieving our goal of public and accessible data for all researchers.

Once the project takes off, there is no reason why it should be limited just to air quality measurements. Our extensible API design already allows usage of non-air-related measurements. If needed, we can also optimize our website for display of those measurements. We will connect with other research institutions to see where the platform can also be used.

Conclusion

The main unique aspect of our project is that is does not restrict users to any specific measurement station, allowing them to customize the equipment completely to suit their needs. However, for users that just want to take part of the platform, we also provide a pre-built product and a DIY kit, while still maintaining a low cost. This helps the project to attract users of all kinds and make it grow more easily.

Initially, our product will be targeted towards individuals and specific public institutions. As the platform grows, more and more data will become available for researchers, increasing both scientific

and market value of the project. Taking feedback from individuals and research organization will allow us to continuously improve our platforms to make it better and even more useful for different users.

We will connect with local communities and clubs to help us promote the project and take measurement devices on remote locations, increasing the data coverage. As part of the main Erasmus+ project, we will also create other environmental measuring systems and connect them to our platform.

BIBLIOGRAPHY

[1] World Health Organization, "WHO global air quality guidelines," 2021. [Online]. Available: https://apps.who.int/iris/rest/bitstreams/1371692/retrieve.

[2] World Health Organization, "Air pollution," [Online]. Available: https://www.who.int/health-topics/air-pollution.

[3] World Health Organization, "Household air pollution and health," 22 09 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health.

[4] World Health Organization, "Ambient (outdoor) air pollution," 22 09 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health.

[5] L. Lattanzio, "Particulate Matter Sensing for Air Quality Measurements," 27 12 2018. [Online]. Available: https://www.fierceelectronics.com/components/particulate-matter-sensing-for-air-quality-measurements.

[6] S. Bliss, Best Practices Guide to Residential Construction, John Wiley & Sons, 2006.