POLITECNICO
MILANO 1863

Requirement Analysis and Specification Document

Lorenzo Carlassara
Angelica Iseni
Emma Lodetti
Virginia Valeri

Software Engineering for Geoinformatics

Academic year 2022/23

Title:	Requirement Analysis and Specification Document
Authors:	Carlassara L. (10601118), Iseni A. (10668862), Lodetti E. (10619244), Valeri V. (10639607)
Date:	April 17, 2023
Professor:	Giovanni E. Quattrocchi

Contents

1 Introduction

The purpose of this document is to describe the requirements for the development of an interactive client-server application that enables users to access, query and visualize air quality and weather sensor data retrieved from existing public digital archives. The system consists of three main components: a database to ingest and store the selected data, a web server (backend) to expose a REST API for querying the database, and a dashboard to provide means for requesting, processing (descriptive statistics and forecasting), and visualizing data (e.g. maps, dynamic graphs, etc.).

1.1 Context and motivations

The European Commission's agreement on electronic fuels will enable the sale of heat-powered vehicles beyond 2035, when the ban on gasoline and diesel cars comes into effect. This is provided that these vehicles run on synthetic, climate-neutral fuels. Additionally, European energy ministers have approved the regulation to phase out gasoline and diesel engines by 2035 through a majority vote. Public authorities collect air quality and weather observations in near-real time from ground sensor stations and store them in digital archives. Ground sensors data are composed of long time series of observations and sensors metadata, including coordinates, type of measured variable, etc. Often, observations from different sensors and/or providers require different patterns for data accessing, harmonization, and processing. Interactive applications and dashboards, capable of facilitating such tasks, are key for supporting both public authorities as well as ordinary citizens in data processing and visualization.

1.2 Definitions, acronyms, abbreviations

- **Ordinary citizens:** refers to an average person who is a member of a particular community, society, or country and who is not distinguished by any special status, profession, or achievement. In the context of government and politics, the term is often used to describe the majority of the population who are not involved in the decision-making process or who do not hold any formal political power.
- **Air quality:** refers to the degree to which the air is clean, clear, and free from pollutants.
- **Ground sensors:** refers to the sensors installed on the ground that monitor air quality and weather conditions.
- **Public digital archives:** refer to online repositories or databases of digital information that are publicly accessible and maintained by governmental or public entities.

1.3 Solution overview

In our project, the application consists in a set of function that allows user to manipulate e visualize data.

- **Data harmonization:** the user has the possibility to manipulate different formats of data.
- **Data preprocessing:** the user has the possibility of query only data on which he is interested.
- **Data cleaning:** consists in outlier detection and removal.
- **Descriptive statistics:** the user can obtain mean, mode, median, range, standard deviation, variance, interquartile range of the data of interest.
- **Forecast:** the user can obtain predictions of air pollution in the next hours.

1.4 Scope and limitations

Observations from different sensors require different patterns for data accessing, harmonization, and processing. This can make it difficult to access and process data from different sources. Additionally, the system may not be able to handle large amounts of data or complex queries. It is also important to note that the system relies on existing public digital archives to retrieve air quality and weather sensor data. If these archives are not up-to-date or do not contain all necessary data, the effectiveness of the system could be limited.
It has been decided to analyze the air pollution data in some of the most popular capitals of Europe. Specifically:

- Paris
- Rome
- Oslo
- Madrid
- Berlin

This cities will be evaluated in order to discern the differences in the main pollution indeces:

$PM_{2.5}$: Particulate matter 2.5, refers to tiny particles or droplets in the air that are two and one half micron or less in width. Emissions from combustion of gasoline, oil, diesel fuel or wood produce much of the PM2.5 pollution found in outdoor air, as well as significant proportion of PM10.

PM_{10}: Particulate matter 10, refers to tiny particles with a diameter of 10 micron or less .PM10 also includes dust from construction sites, landfills and agriculture, wildfires and brush burning, industrial sources, wind-blown dust from open lands(pollen and fragments of bacteria).

O_3: Ozone, an highly reactive gas composed of three oxygen atoms. It occurs both in the Earth's upper atmosphere and at ground level.

SO_2: Sulfur Dioxide, a heavy, colourless, and poisonous gas composed of sulfur and oxygen. It is produced from the burning of fossil fuels(coal and oil) and the smelting of mineral ores.

NO_2: Nitrogen Dioxide, a gaseous air pollutant composed of nitrogen and oxygen. This gas forms when fossil fuels (coal, oil, gas,diesel) are burned at high temperatures .

CO: Carbon Monoxide, a colorless,odorless gas that can be harmful when inhaled in large amounts. The greatest sources of CO to outdoor air are cars,trucks and other vehicles or machinery that burn fossil fuels.

These pollutants can have negative impacts on human health, as they can be inhaled deep into the lungs and bloodstream and cause respiratory problems as asthma, bronchitis and cardiovascular problems. Long-term exposure can increase the risk of chronic diseases such as lung cancer and heart diseases. Furthermore,they can also affect visilibility, air quality, and the health of surronding environment.

2 Requirements

2.1 Stakeholders

- Ordinary citizens
- European Environmental Ministries
- Automotive production companies
- Public authorities

2.2 Actors

- European Ministries
- Public authorities

2.3 Domain Assumption

- The user can find the link of the project repository on GitHub.
- The users can send feedback to the team , to suggest improvements or express satisfiability with the service contacting us.
- The ground sensor stations are distributed geographically and collect data on a regular basis.
- The air quality and weather observations collected by the ground sensor stations are stored in databases.
- The data collected by different ground sensor stations may have different formats, structures, and levels of quality.

2.4 Uses cases

Sign up Users can create a profile in the application by using their email address and generating a password.

User login/logout The users can log in with their credentials.

Data PreProcessing The user selects the desired location and temporaln range for data visualization.

Data Visualization The user can visualize selected data in a graphical format or in tables. The user can interact with the data visualization, zooming in and out, selecting specific variables to view, and adjusting the time range of the data displayed.

Processing Results Visualization The user can visualize the results of the data processing, as descriptive statistics and forecasts.

Admin data management The admin logs into the application and is presented with a dashboard displaying the current data collection status, as well as the ability to update metadata, and adjust the data processing algorithms.

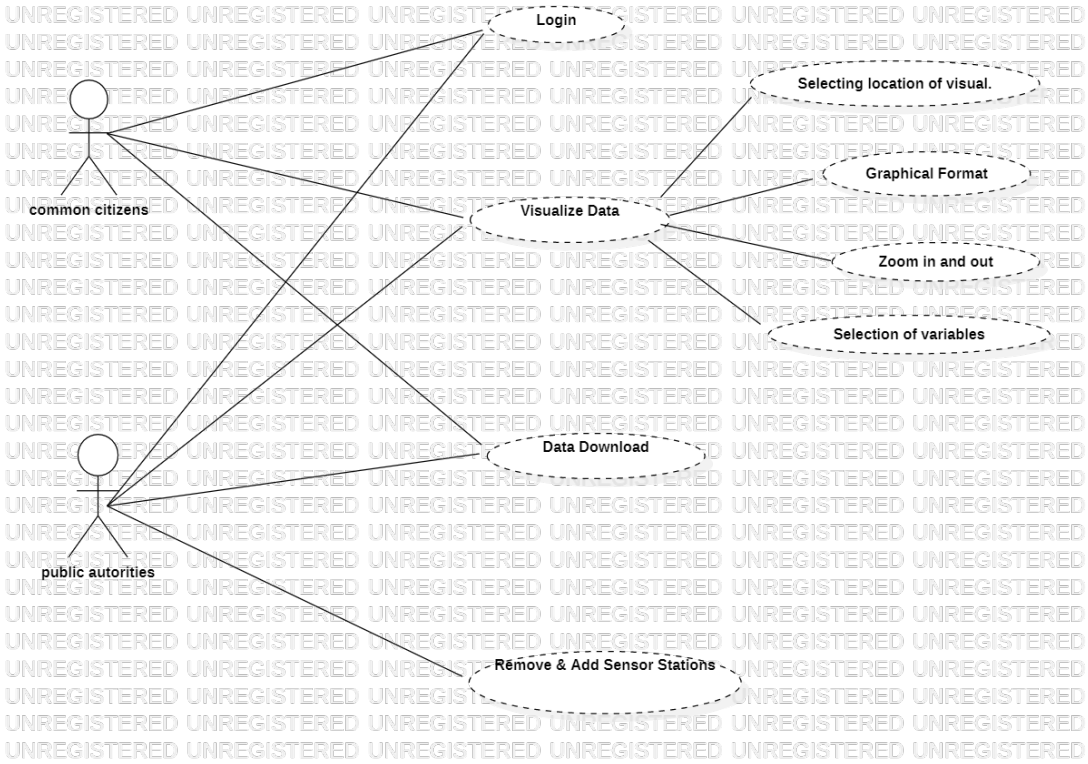

Figure 1: Uses cases diagram

2.5 User stories

User sign up

- As a user I need to sign up to the application in order to access to its functionalities.

User login/logout

- As a user, I want to be able to login/out to the application so that I can access air quality and weather data.

Data PreProcessing

- As a user, I want to be able to select the location for data visualization so that I can view data for a specific area.
- As a user, I want to be able to select the location for data visualization so that I can view data for a specific temporal range.

Data Visualization

- As a user, I want to be able to visualize data in tables.
- As a user, I want to be able to interact with the data visualization so that I can zoom in and out, select specific variables to view, and adjust the time range of the data displayed.

Processing Results Visualization

- As a user, I want that my requested data results without outliers.
- As a user, I want to have the possibility of visualize some descriptive statistics of the data.
- As a user,I want to visualize forecasts about the air pollution for a certain temporal range.

Admin data management

- As an admin, I want to be able to log in to the application so that I can manage the air quality and weather data stored in the database.
- As an admin, I want to be able to add or remove sensor stations so that I can update the data collection process.
- As an admin, I want to be able to update metadata so that users have accurate information about the data.
- As an admin, I want to be able to adjust the data processing algorithms so that users receive high-quality data.

2.6 Requirements

The following requirements must be met by the system:

2.6.1 Non-Functional Requirements

- The interface language will be the English
- The system should be available 24h a day
- The data should be displayed in a manner that is clear and easy to understand

2.6.2 Functional Requirements

- The application needs to be able to access and harmonize data from different sources and formats, using standard protocols and tools.
- The application needs to be able to perform basic data processing and analysis tasks, such as data filtering, aggregation, and visualization.
- The application needs to be scalable and able to handle large amounts of data, as well as multiple concurrent users.
- The application needs to be secure and protect sensitive data, such as personal information or confidential data.
- The application should easily let the users to log in/out

2.6.3 Technical Requirements

- The main backend implementation language must be Python

2.7 Constraints

- The system must comply with data privacy regulations and standards.
- The system should be developed using open-source software.
- The system must be developed using Python programming language.
- The system must be hosted on a cloud platform.