

## An IoT-empowered air quality monitoring system integrated with a machine learning framework to detect and predict defined air quality parameters

### Problem statement

It is undoubtedly vital to state that the balanced air quality is utmost important to the environmental existence including the survival of humankind. Nonetheless, the gradual deterioration of global air homeostasis urges the need for efficient novel approaches to motivate greater levels of engagement in the air quality debate by empowering citizens around the world to demand cleaner air and better policy formulations and to effectively use the air quality information by collecting and processing impartial data.

### High-level conceptual block diagram

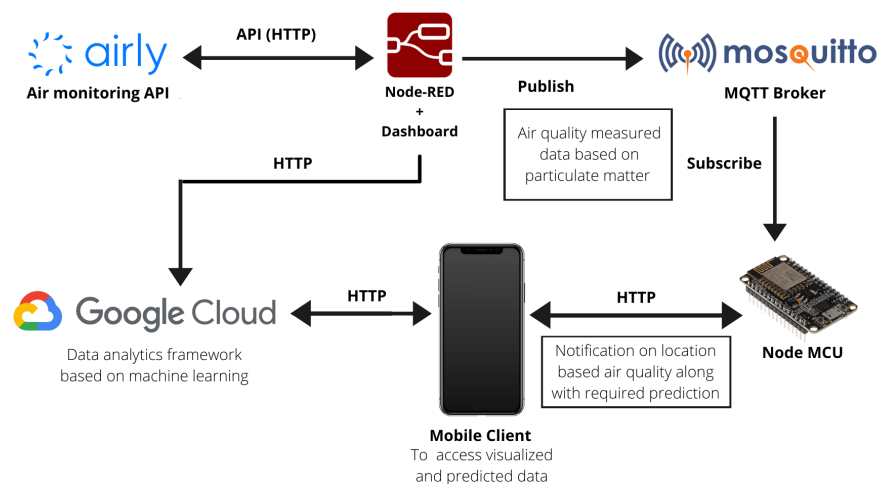


Figure 1: The proposed high-level block diagram for the system

### The Open-source data framework to be used

It is planned to use “airly”, an open-source air quality monitoring API, which offers precise, hyper-local data about air pollution in a given community, giving a better understanding of the health risks and make improvements. This source API carries the ability to obtain up to 100 daily calls for real-time air quality measurements, up to 5000 daily calls for real-time air quality measurements, and millions of daily calls for real-time air quality measurements.

### Planned usage of each of the main block/component

- “airly” air monitoring API: To obtain the real-time raw data of defined air quality parameters
- Node-red: Initial data processing and virtualization obtained form airly API
- Mosquitto: To establish the publish-subscribe network application layer protocol between node-red and MCU
- Node-MCU: To obtain the processed data form node-red via the subscription through mosquitto to communicate with mobile (via http)
- Cloud Storage: To facilitate the data analysis framework based on machine learning
- Mobile: To access the visualized-predicted data and notifications by the client

### A list of system functions and features

- Node-red dashboard representation
  - Location and time-based raw data virtualization for defined air quality parameters: particulate matter,  $SO_2$  and  $CO$
  - Visualization of one-day forested data with one-hour resolution time
  - Simplified decision making options
  - Settings for adjusting axis resolutions and visual representations
- To the client through mobile
  - Via Node-MCU: Location and time-based data requesting and representation, Decisive notifications for a user input location
  - Via cloud storage: Access and visualize time series raw data and modeled data of a location, obtained through the machine learning model