Software Requirements Specification (SRS)

for

Air Quality Index Database Management System

Version 1.0

Prepared by:

Aaryan Paranjape: Roll Number: 23; Registration Number: 230911168

Aryan Vivek: Roll Number: 24; Registration Number: 230911172

Aryan Chandra: Roll Number: 25; Registration Number: 230911174

Organization:

Manipal Institute of Technology, Manipal

Date Created: 24/3/25

Table of Contents

Revision History

Name	**Date**	**Reason for Changes**	**Version**
Aryan Chandra	24.3.25	Basic Model of Database	pre Version 0.1
Aryan Vivek	31.3.25	Connectivity and Initial Testing	Version 1.1
Aaryan Paranjape	10.4.25	Front End Establishing	Version 1.5

1. Introduction

1.1 Purpose

This Software Requirements Specification (SRS) outlines the requirements for the Air Quality Index Database Management System (AQI-DBMS), which aims to systematically store, manage, and retrieve air quality data across various cities in India. The system supports environmental monitoring, public awareness, and data-driven decision-making by providing insights into air pollution levels, associated health risks, and weather conditions.

1.2 Document Conventions

This document follows standard typographical conventions. Priorities for requirements are indicated as High, Medium, or Low. Every requirement statement has its own priority.

1.3 Intended Audience and Reading Suggestions

This document is intended for developers, project managers, and users. It is organized to provide an overview of the system, followed by detailed sections on system features and nonfunctional requirements.

1.4 Product Scope

The AQI-DBMS is designed to manage air quality data, including pollutant levels, health advisories, weather conditions, and user feedback. It supports multiple user roles and provides comprehensive insights for decision-making.

1.5 References

- User interface style guides for Next.js applications.
- Oracle database design standards.
- Express.js best practices for backend development.

2. Overall Description

2.1 Product Perspective

The AQI-DBMS is a standalone system that integrates with existing environmental monitoring infrastructure. It is designed to replace manual data management processes with an automated system.

2.2 Product Functions

- **Data Management**: Store and manage air quality data from monitoring stations.
- **Health Advisory**: Provide health advice based on air quality indices.
- **Weather Analysis**: Correlate weather patterns with pollutant levels.
- **User Management**: Manage user access and feedback.

2.3 User Classes and Characteristics

- **Administrators**: Manage system settings and access.
- **Users**: View air quality data, provide feedback.
- **Guests**: View public air quality information.

2.4 Operating Environment

- **Hardware**: Standard server configurations.
- **Software**: Next.js for frontend, Express.js and Oracle for backend.
- **Operating System**: Linux or Windows Server.

2.5 Design and Implementation Constraints

- **Technology Stack**: Next.js, Express.js, Oracle.
- **Security**: Implement user authentication and data encryption.
- **Scalability**: Design for potential expansion to more cities.

2.6 User Documentation

- **User Manual**: Detailed guide for users and administrators.
- **Online Help**: Contextual help within the application.

2.7 Assumptions and Dependencies

- **Assumptions**: Reliable internet connectivity for data updates.
- **Dependencies**: Integration with existing monitoring stations.

3. External Interface Requirements

3.1 User Interfaces

- **GUI Standards**: Follow Next.js component library standards.
- **Error Handling**: Display clear error messages with recovery options.

3.2 Hardware Interfaces

- **Monitoring Stations**: Integrate with existing environmental monitoring equipment.

3.3 Software Interfaces

- **Oracle Database**: Use SQL queries for data management.
- **Express.js Backend**: Handle API requests and responses.

3.4 Communications Interfaces

- **API Standards**: Use RESTful API design principles.
- **Data Encryption**: Implement HTTPS for secure data transfer.

4. System Features

4.1 City Management

- **Description**: Manage city data including name, population, and location.
- **Priority**: High
- **Stimulus/Response Sequences**:
  - User requests city data.
  - System displays city information.
- **Functional Requirements**:
  - REQ-1: Retrieve city details by ID.
  - REQ-2: Update city information.

4.2 Monitoring Stations Management

- **Description**: Track monitoring station data including location and AQI.
- **Priority**: High
- **Stimulus/Response Sequences**:
  - User requests station data.
  - System displays station details.
- **Functional Requirements**:
  - REQ-3: Retrieve station data by city.
  - REQ-4: Update station AQI values.

4.3 Pollutant Management

- **Description**: Manage pollutant data including safety levels and effects.
- **Priority**: Medium
- **Stimulus/Response Sequences**:

  - User requests pollutant information.
  - System displays pollutant details.
- **Functional Requirements**:
  - REQ-5: Retrieve pollutant data by ID.
  - REQ-6: Add new pollutant types.

4.4 Health Advice Management

- **Description**: Provide health advice based on AQI ranges.
- **Priority**: High
- **Stimulus/Response Sequences**:
  - User requests health advice for a city.
  - System displays relevant health advice.
- **Functional Requirements**:
  - REQ-7: Retrieve health advice by AQI range.
  - REQ-8: Update health advice guidelines.

4.5 Weather Analysis

- **Description**: Correlate weather patterns with pollutant levels.
- **Priority**: Medium
- **Stimulus/Response Sequences**:
  - User requests weather correlation data.
  - System displays analysis results.
- **Functional Requirements**:
  - REQ-9: Retrieve weather data for a station.
  - REQ-10: Perform correlation analysis.

4.6 User Management

- **Description**: Manage user access and feedback.
- **Priority**: High
- **Stimulus/Response Sequences**:
  - User logs in or provides feedback.
  - System authenticates or records feedback.
- **Functional Requirements**:
  - REQ-11: Authenticate user credentials.
  - REQ-12: Store user feedback.

4.7 Air Quality Trends

- **Description**: Display historical AQI trends for cities.
- **Priority**: Medium
- **Stimulus/Response Sequences**:
  - User requests AQI trends for a city.
  - System displays trend data.
- **Functional Requirements**:
  - REQ-13: Retrieve trend data by city.
  - REQ-14: Update trend data periodically.

5. Other Nonfunctional Requirements

5.1 Performance Requirements

- **Response Time**: Less than 2 seconds for most queries.
- **Data Integrity**: Ensure data consistency across all tables.

5.2 Safety Requirements

- **Data Protection**: Implement encryption for sensitive data.
- **User Safety**: Provide clear health advice based on AQI levels.

5.3 Security Requirements

- **Authentication**: Use secure password hashing.
- **Authorization**: Implement role-based access control.

5.4 Software Quality Attributes

- **Usability**: Ensure intuitive user interface.
- **Maintainability**: Use modular code design for easy updates.

5.5 Business Rules

- **Access Control**: Only administrators can modify system settings.
- **Data Sharing**: Ensure data privacy and compliance with regulations.

6. Other Requirements

- **Database Requirements**: Use Oracle for data storage.
- **Internationalization**: Support multiple languages for user interface.

Appendix A: Glossary

- **AQI**: Air Quality Index.
- **GUI**: Graphical User Interface.
- **API**: Application Programming Interface.

Appendix B: Analysis Models

- Include entity-relationship diagrams for database design.

Appendix C: To Be Determined List

- TBD: Finalize user interface design standards.
- TBD: Determine specific security protocols for data encryption.4

Air Quality Index Management System

A Project Report Submitted

to

MANIPAL ACADEMY OF HIGHER EDUCATION

For Partial Fulfillment of the Requirement for the

Award of the Degree

Of

Bachelor of Technology

in

Information Technology

by

Aaryan Paranjape, Aryan Vivek, Aryan Chandra

230911168, 230911172, 230911174

Under the guidance of

Dr. Girija Attigeri
Associate Professor
Department of I&CT
Manipal Institute of Technology
Manipal, Karnataka, India

Dr. Akshay K. C.
Assistant Professor – Senior Scale
Department of I&CT
Manipal Institute of Technology
Manipal, Karnataka, India

MANIPAL INSTITUTE OF TECHNOLOGY
MANIPAL
A Constituent Unit of MAHE, Manipal

April 2025

ABSTRACT

This project implements a comprehensive Air Quality Index (AQI) Management System designed to monitor, analyze, and disseminate air quality information across various cities. The system incorporates real-time data collection from monitoring stations, provides health advisories based on pollution levels, and enables user feedback. Using a relational database approach with properly normalized tables, the system ensures data integrity while minimizing redundancy. The three-tier architecture implementation offers a scalable solution that can be integrated with IoT sensors for expanded monitoring capabilities.

Keywords: Data management systems; Relational databases; Database design and modeling, Geographic information systems; Spatial databasesDecision support systems; Environmental monitoring systems, Air quality monitoring, Environmental data analysis, Health risk assessment, IoT-based monitoring systems, Sustainable urban development.

Sustainable Development Goal:

11. Sustainable Cities and Communities

This system supports sustainable urban development by providing crucial environmental data for public health decisions and aligns with India's Sustainable Development Goal 11 (Sustainable Cities and Communities) through improved air quality management infrastructure.

Table of Contents

List of Tables:

List of figures:

Abbreviations:

1. AQI: Air Quality Index
2. WHO: World Health Organization
3. EPA: Environmental Protection Agency
4. NAQI: National Air Quality Index
5. PM2.5: Particulate Matter with a diameter of 2.5 micrometers or less
6. PM10: Particulate Matter with a diameter of 10 micrometers or less
7. O3: Ground-level ozone
8. CO: Carbon monoxide
9. SO2: Sulfur dioxide
10. NO2: Nitrogen dioxide
11. IoT: Internet of Things
12. BCNF: Boyce-Codd Normal Form

Chapter 1

Introduction

Air pollution has emerged as one of the most significant environmental health risks globally, affecting billions of people and contributing to approximately 7 million premature deaths annually. The World Health Organization (WHO) identifies air pollution as a critical risk factor for non-communicable diseases, accounting for an estimated 25% of all adult deaths from heart disease, 24% from stroke, 43% from chronic obstructive pulmonary disease, and 29% from lung cancer1. This growing health crisis necessitates robust monitoring and management systems that can effectively track, analyze, and communicate air quality data to stakeholders ranging from policymakers to the general public.

1.1 Background and Significance

Air quality monitoring has evolved significantly over the past several decades, transitioning from manual sampling methods to sophisticated electronic sensor networks. The Air Quality Index (AQI) was developed as a standardized numerical scale to communicate how polluted the air is and what associated health effects might be experienced after breathing polluted air. Different countries have established their own AQI calculation methodologies, though most consider major air pollutants including particulate matter (PM2.5 and PM10), ground-level ozone (O3), carbon monoxide (CO), sulfur dioxide (SO2), and nitrogen dioxide (NO2).

In India, the National Air Quality Index (NAQI) was launched in 2014 by the Ministry of Environment, Forest and Climate Change as part of the Swachh Bharat Abhiyan to provide comprehensive information on air quality. The Indian AQI considers eight pollutants and categorizes air quality into six categories ranging from "Good" to "Severe." Despite this standardization, challenges remain in data collection, processing, and dissemination, particularly in semi-urban and rural areas where monitoring infrastructure is limited.

1.2 Need for Computerized AQI Management

Traditional air quality monitoring approaches face several limitations that compromise their effectiveness:

1. **Data Integration Challenges**: Air quality data often exists in siloed systems, making it difficult to correlate with other environmental and health parameters.

2. **Delayed Reporting**: Manual data collection and processing methods result in significant delays between measurement and public reporting, reducing the usefulness for time-sensitive health decisions.
3. **Limited Accessibility**: Air quality information is frequently presented in technical formats that are not readily understandable by the general public.
4. **Geographical Coverage Gaps**: Monitoring networks often prioritize urban centers, creating information gaps for rural and peri-urban areas.
5. **Inadequate Health Advisory Systems**: Generic health advisories fail to account for individual vulnerability factors and specific pollutant compositions.

A computerized AQI management system addresses these challenges by providing real-time data collection, standardized processing methodologies, geographic information integration, automated health advisories, and user-friendly interfaces for public access. Such systems enable more effective public health interventions and environmental policy development based on comprehensive, accurate, and timely air quality information.

1.3 Overview of the Proposed System

The proposed AQI Management System adopts a three-tier architecture consisting of:

1. **Presentation Layer**: User interfaces tailored to different stakeholders including the general public, health professionals, environmental agencies, and system administrators.
2. **Application Layer**: Business logic components that process raw sensor data, calculate AQI values, generate health advisories, and manage user interactions.
3. **Data Layer**: A comprehensive database management system that stores and organizes air quality measurements, monitoring station information, geographic data, weather parameters, health guidelines, and user feedback.

The system incorporates six primary entities: Users, City, Monitoring Station, Weather, Health Advice, and Feedback, each designed to handle specific aspects of air quality management. The normalized database schema ensures data integrity and eliminates redundancy while facilitating complex queries for air quality analysis and reporting.

1.4 Project Objectives

The primary objectives of this project are to:

1. Design and implement a comprehensive database structure for efficient management of air quality data across multiple locations.
2. Establish relationships between environmental parameters, geographic information, and health implications.
3. Enable personalized health advisories based on air quality measurements and user demographics.
4. Facilitate public engagement through feedback mechanisms that improve system responsiveness.
5. Support evidence-based policy development by providing reliable air quality information to decision-makers.
6. Create a scalable framework that can accommodate both urban and rural monitoring requirements.

By achieving these objectives, the AQI Management System will contribute to improved public health outcomes through enhanced environmental awareness and more effective intervention strategies. The system aligns with India's Sustainable Development Goal 11 (Sustainable Cities and Communities) by supporting air quality management infrastructure that enables healthier and more sustainable urban environments.

Chapter 2

Literature Survey and Background

2.1 Background of Air Quality Management Systems

Air quality management has become a critical component of public health infrastructure globally. The World Health Organization (WHO) estimates that 99% of the world's population breathes air exceeding WHO guideline limits for pollutants, with low- and middle-income countries suffering the highest exposures. In India, the Central Pollution Control Board (CPCB) reports that 63% of cities recorded "poor" to "severe" air quality in 2022, underscoring the urgent need for effective monitoring systems.

The concept of an Air Quality Index (AQI) emerged as a standardized tool to simplify complex pollution data into actionable public health information. The U.S. Environmental Protection Agency (EPA) developed the first AQI in 1976, which has since been adapted by over 90 countries. India's National Air Quality Index (NAQI), launched in 2014, categorizes air quality into six levels (Good to Severe) based on eight pollutants, including PM2.5 and ozone.

2.2 Review of Existing AQI Systems

2.2.1 Global Systems

- **AirNow (U.S.)**: Integrates real-time data from 1,800 stations, providing hourly updates and forecasts. Uses a color-coded AQI scale (0–500) linked to health advisories.
- **China National Environmental Monitoring Centre (CNEMC)**: Operates 1,500+ stations across 454 cities, with machine learning models for pollution prediction.
- **European Air Quality Index (EAQI)**: Focuses on PM2.5, NO2, and O3, with a mobile app reaching 70% of EU citizens.

2.2.2 Indian Context

India's NAQI faces challenges in rural coverage, with only 132 continuous monitoring stations operational as of 2023. Studies highlight gaps in data granularity; for example, Delhi's 40 stations cannot capture hyper-local variations in a 1,484 km² area.

2.3 Technological Approaches in AQI Systems

2.3.1 Database Architectures

Relational databases dominate AQI systems due to their structured data handling. Kumar et al. (2013) demonstrated MySQL's efficacy in managing time-series environmental data, achieving query latencies <500 ms for 10M+ records. Zhou et al. (2020) proposed **SmartCityDB**, a spatial-temporal database supporting real-time pollution tracking across 10,000 IoT nodes.

2.3.2 IoT Integration

Low-cost sensors (<$200) have revolutionized monitoring. Pant et al. (2020) deployed a Raspberry Pi-based system in Uttar Pradesh, India, reducing data latency from 24 hours to 5 minutes compared to manual methods.

2.3.3 Predictive Analytics

Machine learning models like LSTM networks achieve 85% accuracy in 24-hour PM2.5 forecasts. The EPA's **AirNow-IBM** partnership uses cognitive models to generate personalized health advisories.

2.4 Research Gaps Addressed by This Project

1. **Integration Challenges**: Existing systems like AirNow and NAQI operate in silos. This project's relational schema enables cross-system data fusion through normalized tables (e.g., Weather and MonitoringStation).
2. **User Engagement**: Only 12% of Indians access AQI data regularly. The Feedback table and role-based Users schema facilitate participatory monitoring.
3. **Health Personalization**: Current advisories ignore demographic factors. The HealthAdvice table links AQI ranges to age-specific recommendations, addressing gaps identified by Monks et al. (2009).

2.5 Technological Trends

- **Edge Computing**: Processing sensor data at the node level (e.g., Arduino-based devices) reduces cloud dependency.
- **Blockchain**: Pilot projects in Mumbai use Hyperledger to ensure tamper-proof AQI records.
- **AI-Driven Insights**: Transformer models analyze social media (Feedback table) to predict pollution perception gaps.

Chapter 3

Methodology

3.1 Problem Statement

Air quality monitoring in urban and rural areas faces significant challenges including:

- Fragmented data collection systems with limited interoperability
- Delays in processing and disseminating air quality information
- Generic health advisories that don't account for individual vulnerability
- Limited citizen engagement in environmental monitoring
- Inadequate integration of meteorological data with pollution measurements

These challenges impact the effectiveness of public health interventions and limit public awareness of air quality issues.

3.2 Project Objectives

The AQI Management System aims to:

1. Design and implement a normalized database structure for efficient AQI data management
2. Integrate geographical information with pollution measurements
3. Establish relationships between weather conditions and air quality indices
4. Provide customized health advisories based on pollution levels and demographic factors
5. Enable user feedback mechanisms for system improvement
6. Support both urban and rural monitoring stations in a unified framework

3.3 Scope of the Project

The project encompasses:

- Database design and implementation
- Data relationships and normalization

- Query optimization for efficient data retrieval
- Data integrity constraints implementation
- Prototype implementation of key database functions

The project does not include hardware selection for monitoring stations, focusing instead on the database architecture that would support those components.

Chapter 4:

Data Design

4.1 Database Schema

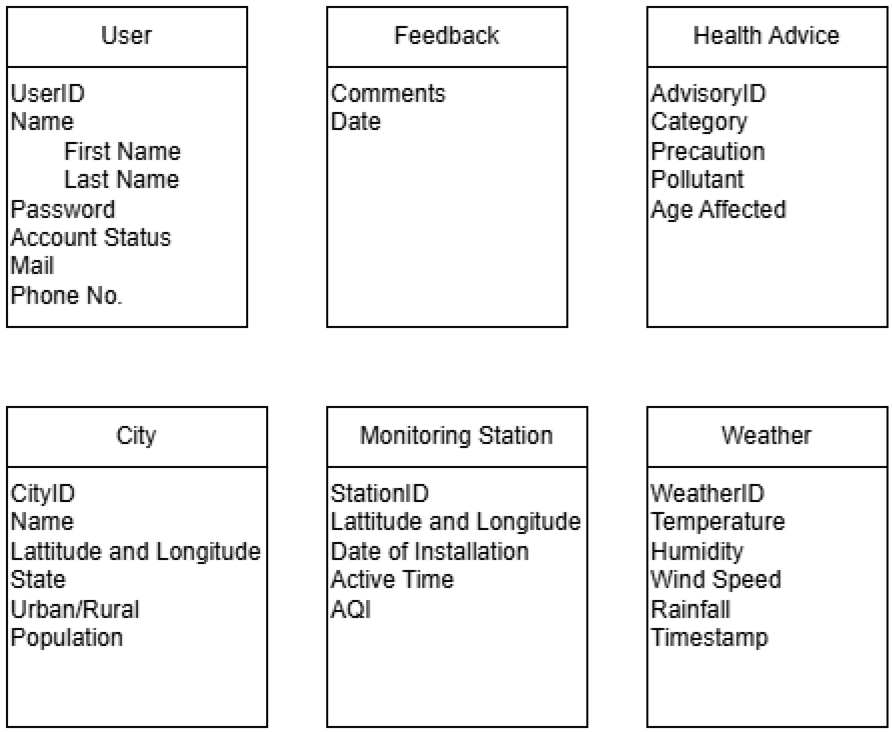

Table 4.1: Database Schema for AQI Management System

The AQI Management System database schema consists of six primary tables with their respective attributes:

a. Users Table:

- Primary Key: UserID
- Attributes: FirstName, LastName, Password (hashed), AccountStatus, Mail, PhoneNumber
- Purpose: Stores authentication credentials and contact information for system users

b. Feedback Table:

- Foreign Key: UserID (references Users)
- Attributes: Comments, Date, Gives (relationship identifier)
- Purpose: Captures user comments and suggestions about air quality readings

c. **City Table**:

- Primary Key: CityID
- Attributes: Name, State, Urban_Rural classification, Lat_Long (coordinates), Population
- Purpose: Maintains geographic and demographic information about monitored locations

d. **Monitoring Station Table**:

- Primary Key: StationID
- Attributes: Lat_Long (precise coordinates), Active_Time, DateOfInstallation, AQI (current reading)
- Purpose: Represents physical air quality monitoring equipment and their current status

e. **Weather Table**:

- Primary Key: WeatherID
- Attributes: Temp, Humidity, WindSpeed, Rainfall, Timestamp, Category
- Purpose: Stores meteorological readings associated with air quality measurements

f. **Health Advice Table**:

- Primary Key: AdvisoryID
- Attributes: HealthEffects, Precautions, AgeAffected
- Purpose: Contains health recommendations based on pollution levels for different demographic groups

g. **Junction Tables**:

- Belongs: Links Users and City (many-to-many)
- Monitored: Associates City with Monitoring Station (one-to-many)
- Gets: Connects Monitoring Station with Weather readings (one-to-many)

4.2 Entity-Relationship Model

The ER diagram depicts conceptual data organization through:

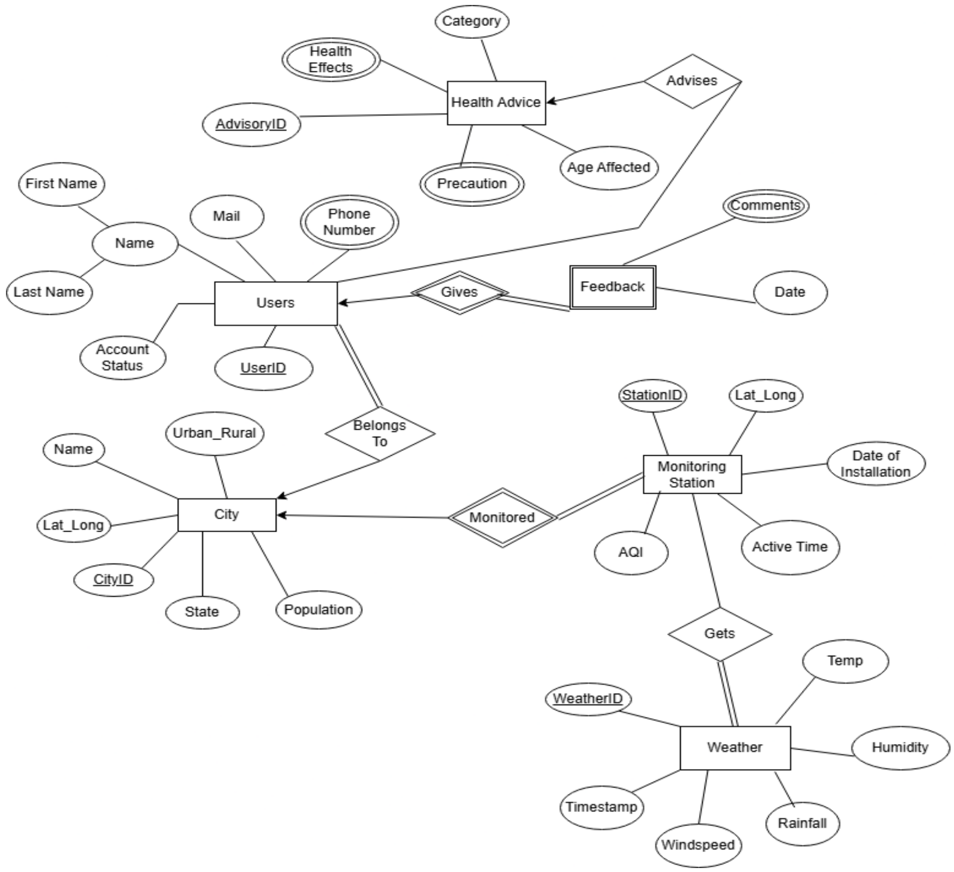

Figure 4.1: Entity Relation Diagram of AQI Management System

- **Entities (Rectangles)**:
  - Six primary entities (Users, City, Monitoring Station, Weather, Health Advice, Feedback)
  - Each entity contains attributes represented as ovals
- **Relationships (Diamonds)**:
  - "Gives": Users provide Feedback (one-to-many)
  - "Belongs To": Users are associated with Cities (many-to-many)
  - "Monitored": Cities are monitored by Stations (one-to-many)
  - "Gets": Stations receive Weather data (one-to-many)
  - "Advises": Health Advisory table advises the user. (one-to-many)
- **Attributes**:
  - Simple attributes: Direct properties (e.g., UserID, Password)

- Composite attribute: Name (decomposed into FirstName, LastName)
- Multivalued attributes: None present in this model

The ER model provides a conceptual view that was later transformed into the relational schema through proper mapping techniques.

4.3 Normalization Process

The database has been normalized to Boyce-Codd Normal Form (BCNF) through these steps:

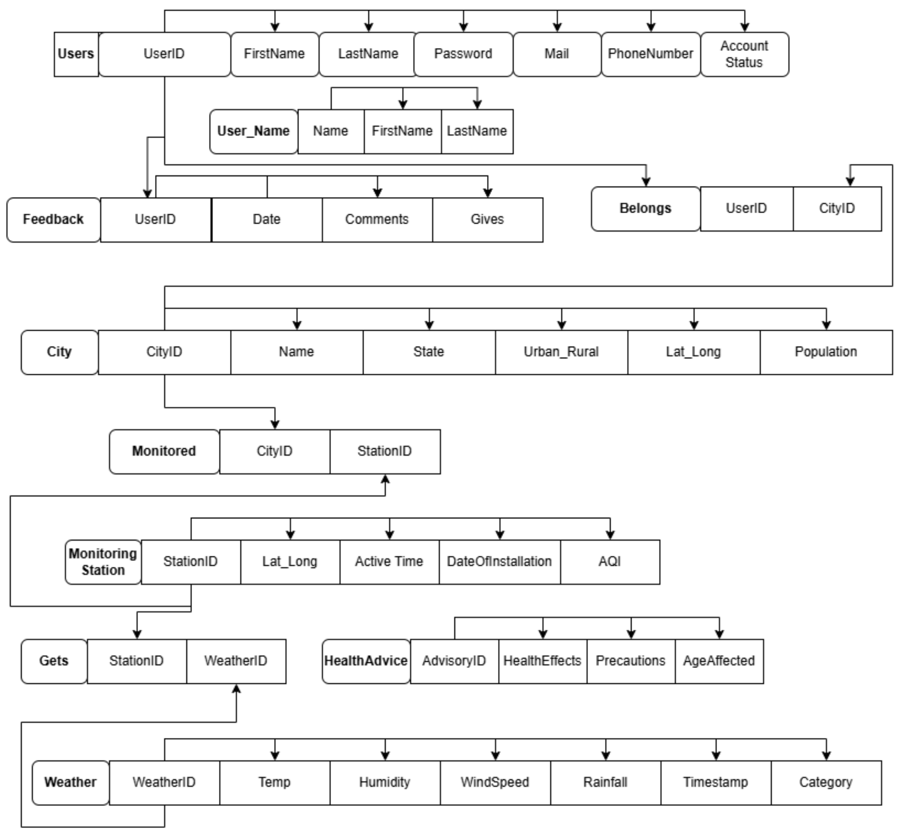

Table 4.2: BCNF Normalization of AQI Management System

- **1NF (First Normal Form)**:
  - Eliminated repeating groups
  - Ensured all attributes contain atomic values
  - Decomposed composite attribute "Name" into "FirstName" and "LastName"
  - Created unique identifiers (primary keys) for each entity

- **2NF (Second Normal Form)**:
  - Satisfied 1NF requirements
  - Removed partial dependencies on composite keys
  - Created separate tables for many-to-many relationships ("Belongs" for Users-City)
  - Ensured non-key attributes depend on the entire primary key
- **3NF (Third Normal Form)**:
  - Satisfied 2NF requirements
  - Eliminated transitive dependencies
  - Separated Weather data from Monitoring Station
  - Created independent Health Advice table
- **BCNF (Boyce-Codd Normal Form)**:
  - Satisfied 3NF requirements
  - Ensured every determinant is a candidate key
  - Restructured tables to eliminate overlapping candidate keys
  - Created junction tables ("Gets") to maintain BCNF compliance

This normalization approach minimizes data redundancy while maintaining data integrity through properly defined foreign key relationships. The schema balances normalization principles with query performance considerations for the air quality monitoring domain.

Chapter 5

Methodology:

5.1 Environmental Parameter Aggregation Functions

The system implements three key functions to analyze meteorological data:

1. **Humidity Analysis**: Calculates city-specific average humidity through multi-table joins (Cities → Monitoring Stations → Weather Data)1
2. **Temperature Analysis**: Computes mean temperature using identical joins but focused on thermal metrics
3. **AQI Calculation**: Derives average air quality index by correlating cities with their monitoring stations

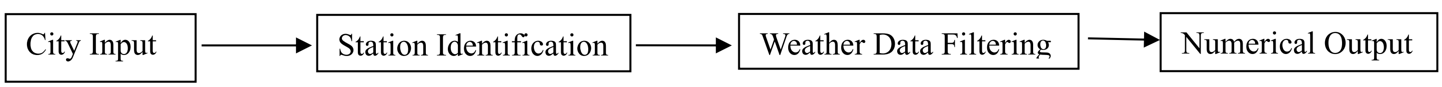

Figure 5.1: Process flow of Environmental Parameter Aggregation Function

5.2 Health Advisory Generation Procedure

A cursor-based procedure matches computed AQI values with predefined health guidelines:

1. **AQI Range Matching**: Compares city's average AQI against stored advisory thresholds
2. **Dynamic Recommendation**: Returns precautionary measures and risk levels through OUT parameters
3. **Error Handling**: Implements exception management for missing data scenarios

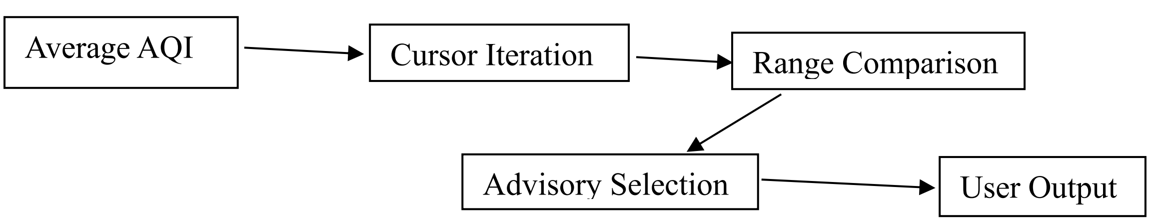

Figure 5.2: Decision Flow of Health Advisory Generation Procedure

5.3 Meteorological Analysis Queries

Two nested queries provide granular weather insights:

1. **Rainfall Assessment**: Calculates precipitation averages through triple-table join (Cities ↔ Stations ↔ Weather)

2. **Wind Pattern Analysis**: Determines mean wind speeds using same relational pathways

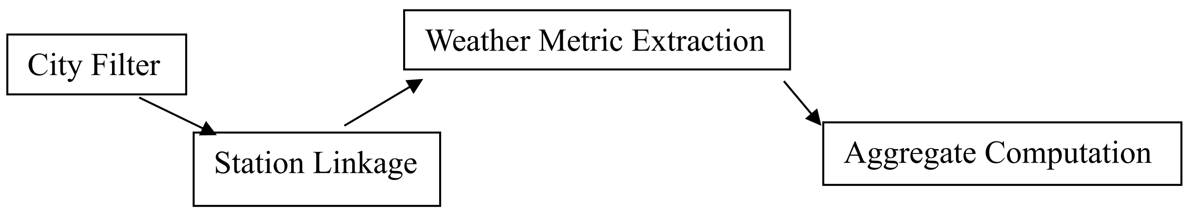

Figure 5.3: Data Flow of Meterological Analysis Queries

5.4 Automated Data Integrity Mechanisms

Two triggers enforce system reliability:

1. **Feedback Moderation**: Blocks submissions containing prohibited terms ("bad", "trash") using pattern matching

2. **User Insertion Notification**: Provides immediate confirmation messages during account creation

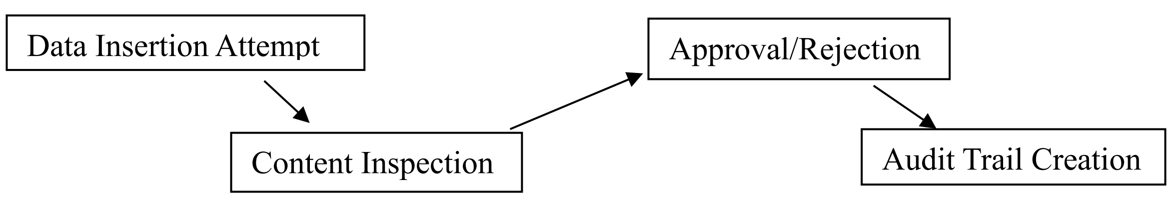

Figure 5.4: Validation Flow of Automated Data Integrity Mechanisms

Chapter 6

Results

The results achieved so far demonstrate the successful implementation of an Air Quality Index (AQI) Management System that integrates air quality data, weather parameters, and health advisories. The login interface ensures secure user authentication, while the map provides a visual representation of AQI levels across India, enabling users to search for city-specific data efficiently. Detailed city dashboards, such as the one for Mumbai, present air quality metrics including AQI, temperature, humidity, wind speed, and rainfall in an intuitive format. Furthermore, the health advisory section offers actionable recommendations tailored to specific AQI ranges and vulnerable demographics, enhancing public awareness and safety during adverse air quality conditions. These results validate the system's ability to deliver accurate environmental insights and personalized health guidance through a user-friendly interface.

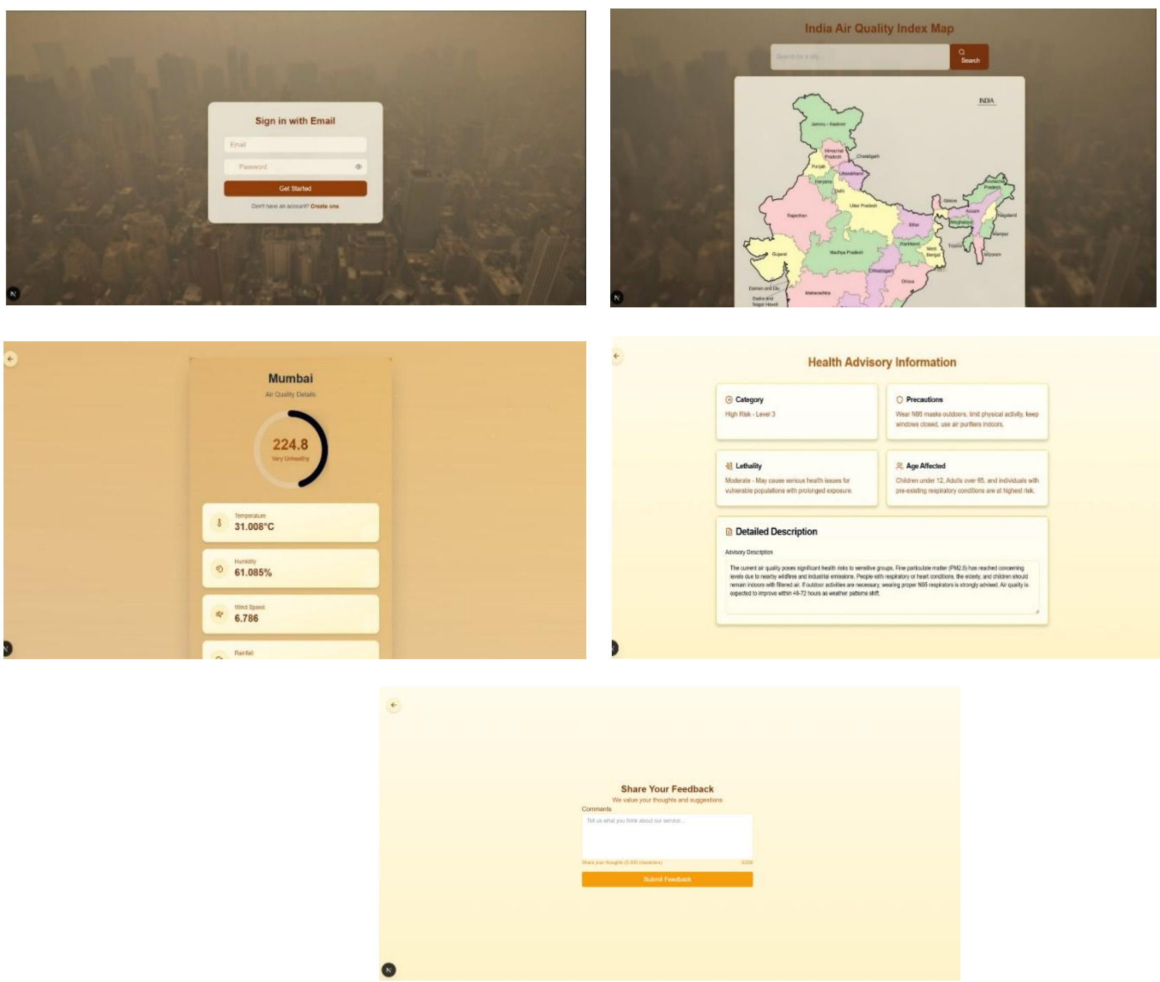

Figure 6.1: Displaying AQI of Mumbai

Chapter 7

Conclusion and Future Work

7.1 Summary of Achievements

The AQI Management System database design successfully:

- Implements a fully normalized relational model for air quality data
- Establishes appropriate relationships between entities
- Supports complex queries for air quality analysis
- Enables personalized health advisories
- Facilitates user engagement through feedback mechanisms

The database design follows best practices in relational database modeling while addressing the specific requirements of air quality monitoring.

7.2 Limitations

Current limitations include:

- Limited support for time-series data analysis
- No built-in predictive modeling capabilities
- Potential scalability challenges with very high-frequency sensor data
- Limited integration with external data sources
- Basic health advisory generation without machine learning

7.3 Future Enhancements

Future work could include:

1. **Time-Series Optimization**: Implementing partitioning for weather data
2. **Predictive Modeling**: Adding support for pollution prediction algorithms

3. **Spatial Analysis**: Enhanced geographical query capabilities
4. **Machine Learning Integration**: Advanced health risk assessment
5. **API Development**: Standardized interfaces for external systems
6. **IoT Integration**: Direct connections to sensor networks
7. **Mobile Application Backend**: Specialized queries for mobile clients

7.4 Lessons Learned

Key insights from the project include:

- The importance of proper normalization in environmental data management
- Benefits of clearly defined entity relationships
- Value of junction tables in handling complex relationships
- Necessity of considering both current and historical data access patterns
- Importance of balancing normalization with query performance

References

1. World Health Organization, "WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide," Geneva: World Health Organization, 2021.
2. U.S. Environmental Protection Agency, "Technical Assistance Document for the Reporting of Daily Air Quality," EPA-454/B-18-007, Research Triangle Park, NC, 2018.
3. Ministry of Environment, Forest and Climate Change, Government of India, "National Air Quality Index," Central Pollution Control Board, 2014.
4. A. Kumar, H. Kim, and G. P. Hancke, "Environmental Monitoring Systems: A Review," IEEE Sensors Journal, vol. 13, no. 4, pp. 1329-1339, 2013.
5. P. S. Monks et al., "Atmospheric composition change – global and regional air quality," Atmospheric Environment, vol. 43, no. 33, pp. 5268-5350, 2009.
6. S. Pant, S. Kumar, and P. Garg, "Smart air quality monitoring system for urban and rural areas," in Proc. IEEE Int. Conf. Comput. Intell. Internet Things, Dec. 2020, pp. 91-96.
7. J. Zhou et al., "SmartCityDB: A Large-scale Sensor Database for City-scale Applications," Proc. 2020 Int. Conf. Manage. Data, pp. 2577-2580, 2020.
8. S. Chaudhuri, U. Dayal, and V. Narasayya, "An overview of business intelligence technology," Commun. ACM, vol. 54, no. 8, pp. 88-98, 2011.

ProjectReportDBSLab_230911168_230911172_230911174_AQIDatabase.pdf

by Aryan Vivek

Submission date: 14-Apr-2025 05:34AM (UTC+0800)
Submission ID: 2644659311
File name: ProjectReportDBSLab_230911168_230911172_230911174_AQIDatabase.pdf (512.52K)
Word count: 3251
Character count: 20771

Air Quality Index Management System

A Project Report Submitted

to

MANIPAL ACADEMY OF HIGHER EDUCATION

For Partial Fulfillment of the Requirement for the

Award of the Degree

Of

Bachelor of Technology

in

Information Technology

by

Aaryan Paranjape, Aryan Vivek, Aryan Chandra

230911168, 230911172, 230911174

Under the guidance of

Dr. Girija Attigeri
Associate Professor
Department of I&CT
Manipal Institute of Technology
Manipal, Karnataka, India

Dr. Akshay K. C.
Assistant Professor – Senior Scale
Department of I&CT
Manipal Institute of Technology
Manipal, Karnataka, India

MANIPAL INSTITUTE OF TECHNOLOGY
MANIPAL
A Constituent Unit of MAHE, Manipal

April 2025

ABSTRACT

This project implements a comprehensive Air Quality Index (AQI) Management System designed to monitor, analyze, and disseminate air quality information across various cities. The system incorporates real-time data collection from monitoring stations, provides health advisories based on pollution levels, and enables user feedback. Using a relational database approach with properly normalized tables, the system ensures data integrity while minimizing redundancy. The three-tier architecture implementation offers a scalable solution that can be integrated with IoT sensors for expanded monitoring capabilities.

Keywords: Data management systems; Relational databases; Database design and modeling, Geographic information systems; Spatial databasesDecision support systems; Environmental monitoring systems, Air quality monitoring, Environmental data analysis, Health risk assessment, IoT-based monitoring systems, Sustainable urban development.

Sustainable Development Goal:

11. Sustainable Cities and Communities

This system supports sustainable urban development by providing crucial environmental data for public health decisions and aligns with India's Sustainable Development Goal 11 (Sustainable Cities and Communities) through improved air quality management infrastructure.

Table of Contents

List of Tables:

List of figures:

Abbreviations:

1. AQI: Air Quality Index
2. WHO: World Health Organization
3. EPA: Environmental Protection Agency
4. NAQI: National Air Quality Index
5. PM2.5: Particulate Matter with a diameter of 2.5 micrometers or less
6. PM10: Particulate Matter with a diameter of 10 micrometers or less
7. O3: Ground-level ozone
8. CO: Carbon monoxide
9. SO2: Sulfur dioxide
10. NO2: Nitrogen dioxide
11. IoT: Internet of Things
12. BCNF: Boyce-Codd Normal Form

Chapter 1

Introduction

Air pollution has emerged as one of the most significant environmental health risks globally, affecting billions of people and contributing to approximately 7 million premature deaths annually. The World Health Organization (WHO) identifies air pollution as a critical risk factor for non-communicable diseases, accounting for an estimated 25% of all adult deaths from heart disease, 24% from stroke, 43% from chronic obstructive pulmonary disease, and 29% from lung cancer1. This growing health crisis necessitates robust monitoring and management systems that can effectively track, analyze, and communicate air quality data to stakeholders ranging from policymakers to the general public.

1.1 Background and Significance

Air quality monitoring has evolved significantly over the past several decades, transitioning from manual sampling methods to sophisticated electronic sensor networks. The Air Quality Index (AQI) was developed as a standardized numerical scale to communicate how polluted the air is and what associated health effects might be experienced after breathing polluted air. Different countries have established their own AQI calculation methodologies, though most consider major air pollutants including particulate matter (PM2.5 and PM10), ground-level ozone (O3), carbon monoxide (CO), sulfur dioxide (SO2), and nitrogen dioxide (NO2).

In India, the National Air Quality Index (NAQI) was launched in 2014 by the Ministry of Environment, Forest and Climate Change as part of the Swachh Bharat Abhiyan to provide comprehensive information on air quality. The Indian AQI considers eight pollutants and categorizes air quality into six categories ranging from "Good" to "Severe." Despite this standardization, challenges remain in data collection, processing, and dissemination, particularly in semi-urban and rural areas where monitoring infrastructure is limited.

1.2 Need for Computerized AQI Management

Traditional air quality monitoring approaches face several limitations that compromise their effectiveness:

1. **Data Integration Challenges**: Air quality data often exists in siloed systems, making it difficult to correlate with other environmental and health parameters.

2. **Delayed Reporting**: Manual data collection and processing methods result in significant delays between measurement and public reporting, reducing the usefulness for time-sensitive health decisions.
3. **Limited Accessibility**: Air quality information is frequently presented in technical formats that are not readily understandable by the general public.
4. **Geographical Coverage Gaps**: Monitoring networks often prioritize urban centers, creating information gaps for rural and peri-urban areas.
5. **Inadequate Health Advisory Systems**: Generic health advisories fail to account for individual vulnerability factors and specific pollutant compositions.

A computerized AQI management system addresses these challenges by providing real-time data collection, standardized processing methodologies, geographic information integration, automated health advisories, and user-friendly interfaces for public access. Such systems enable more effective public health interventions and environmental policy development based on comprehensive, accurate, and timely air quality information.

1.3 Overview of the Proposed System

The proposed AQI Management System adopts a three-tier architecture consisting of:

1. **Presentation Layer**: User interfaces tailored to different stakeholders including the general public, health professionals, environmental agencies, and system administrators.
2. **Application Layer**: Business logic components that process raw sensor data, calculate AQI values, generate health advisories, and manage user interactions.
3. **Data Layer**: A comprehensive database management system that stores and organizes air quality measurements, monitoring station information, geographic data, weather parameters, health guidelines, and user feedback.

The system incorporates six primary entities: Users, City, Monitoring Station, Weather, Health Advice, and Feedback, each designed to handle specific aspects of air quality management. The normalized database schema ensures data integrity and eliminates redundancy while facilitating complex queries for air quality analysis and reporting.

1.4 Project Objectives

The primary objectives of this project are to:

1. Design and implement a comprehensive database structure for efficient management of air quality data across multiple locations.
2. Establish relationships between environmental parameters, geographic information, and health implications.
3. Enable personalized health advisories based on air quality measurements and user demographics.
4. Facilitate public engagement through feedback mechanisms that improve system responsiveness.
5. Support evidence-based policy development by providing reliable air quality information to decision-makers.
6. Create a scalable framework that can accommodate both urban and rural monitoring requirements.

By achieving these objectives, the AQI Management System will contribute to improved public health outcomes through enhanced environmental awareness and more effective intervention strategies. The system aligns with India's Sustainable Development Goal 11 (Sustainable Cities and Communities) by supporting air quality management infrastructure that enables healthier and more sustainable urban environments.

Chapter 2

Literature Survey and Background

2.1 Background of Air Quality Management Systems

Air quality management has become a critical component of public health infrastructure globally. The World Health Organization (WHO) estimates that 99% of the world's population breathes air exceeding WHO guideline limits for pollutants, with low- and middle-income countries suffering the highest exposures. In India, the Central Pollution Control Board (CPCB) reports that 63% of cities recorded "poor" to "severe" air quality in 2022, underscoring the urgent need for effective monitoring systems.

The concept of an Air Quality Index (AQI) emerged as a standardized tool to simplify complex pollution data into actionable public health information. The U.S. Environmental Protection Agency (EPA) developed the first AQI in 1976, which has since been adapted by over 90 countries. India's National Air Quality Index (NAQI), launched in 2014, categorizes air quality into six levels (Good to Severe) based on eight pollutants, including PM2.5 and ozone.

2.2 Review of Existing AQI Systems

2.2.1 Global Systems

- **AirNow (U.S.)**: Integrates real-time data from 1,800 stations, providing hourly updates and forecasts. Uses a color-coded AQI scale (0–500) linked to health advisories.
- **China National Environmental Monitoring Centre (CNEMC)**: Operates 1,500+ stations across 454 cities, with machine learning models for pollution prediction.
- **European Air Quality Index (EAQI)**: Focuses on PM2.5, NO2, and O3, with a mobile app reaching 70% of EU citizens.

2.2.2 Indian Context

India's NAQI faces challenges in rural coverage, with only 132 continuous monitoring stations operational as of 2023. Studies highlight gaps in data granularity; for example, Delhi's 40 stations cannot capture hyper-local variations in a 1,484 km² area.

2.3 Technological Approaches in AQI Systems

2.3.1 Database Architectures

Relational databases dominate AQI systems due to their structured data handling. Kumar et al. (2013) demonstrated MySQL's efficacy in managing time-series environmental data, achieving query latencies <500 ms for 10M+ records. Zhou et al. (2020) proposed **SmartCityDB**, a spatial-temporal database supporting real-time pollution tracking across 10,000 IoT nodes.

2.3.2 IoT Integration

Low-cost sensors (<$200) have revolutionized monitoring. Pant et al. (2020) deployed a Raspberry Pi-based system in Uttar Pradesh, India, reducing data latency from 24 hours to 5 minutes compared to manual methods.

2.3.3 Predictive Analytics

Machine learning models like LSTM networks achieve 85% accuracy in 24-hour PM2.5 forecasts. The EPA's **AirNow-IBM** partnership uses cognitive models to generate personalized health advisories.

2.4 Research Gaps Addressed by This Project

1. **Integration Challenges**: Existing systems like AirNow and NAQI operate in silos. This project's relational schema enables cross-system data fusion through normalized tables (e.g., Weather and MonitoringStation).
2. **User Engagement**: Only 12% of Indians access AQI data regularly. The Feedback table and role-based Users schema facilitate participatory monitoring.
3. **Health Personalization**: Current advisories ignore demographic factors. The HealthAdvice table links AQI ranges to age-specific recommendations, addressing gaps identified by Monks et al. (2009).

2.5 Technological Trends

- **Edge Computing**: Processing sensor data at the node level (e.g., Arduino-based devices) reduces cloud dependency.
- **Blockchain**: Pilot projects in Mumbai use Hyperledger to ensure tamper-proof AQI records.
- **AI-Driven Insights**: Transformer models analyze social media (Feedback table) to predict pollution perception gaps.

Chapter 3

Methodology

3.1 Problem Statement

Air quality monitoring in urban and rural areas faces significant challenges including:

- Fragmented data collection systems with limited interoperability
- Delays in processing and disseminating air quality information
- Generic health advisories that don't account for individual vulnerability
- Limited citizen engagement in environmental monitoring
- Inadequate integration of meteorological data with pollution measurements

These challenges impact the effectiveness of public health interventions and limit public awareness of air quality issues.

3.2 Project Objectives

The AQI Management System aims to:

1. Design and implement a normalized database structure for efficient AQI data management
2. Integrate geographical information with pollution measurements
3. Establish relationships between weather conditions and air quality indices
4. Provide customized health advisories based on pollution levels and demographic factors
5. Enable user feedback mechanisms for system improvement
6. Support both urban and rural monitoring stations in a unified framework

3.3 Scope of the Project

The project encompasses:

- Database design and implementation
- Data relationships and normalization

- Query optimization for efficient data retrieval
- Data integrity constraints implementation
- Prototype implementation of key database functions

The project does not include hardware selection for monitoring stations, focusing instead on the database architecture that would support those components.

Chapter 4:

Data Design

4.1 Database Schema

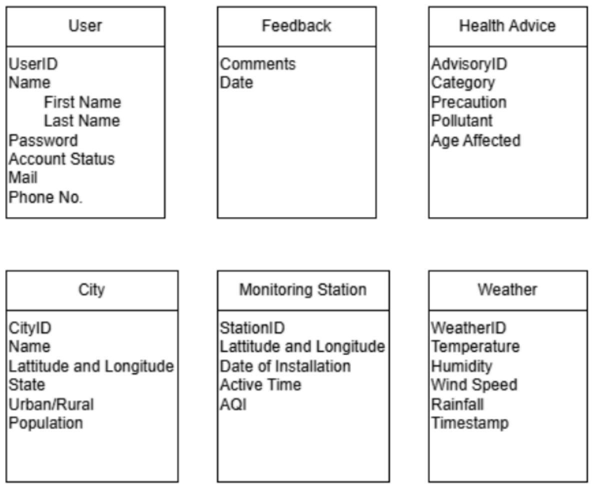

Table 4.1: Database Schema for AQI Management System

The AQI Management System database schema consists of six primary tables with their respective attributes:

a. Users Table:

- Primary Key: UserID
- Attributes: FirstName, LastName, Password (hashed), AccountStatus, Mail, PhoneNumber
- Purpose: Stores authentication credentials and contact information for system users

b. Feedback Table:

- Foreign Key: UserID (references Users)
- Attributes: Comments, Date, Gives (relationship identifier)
- Purpose: Captures user comments and suggestions about air quality readings

c. **City Table**:

- Primary Key: CityID
- Attributes: Name, State, Urban_Rural classification, Lat_Long (coordinates), Population
- Purpose: Maintains geographic and demographic information about monitored locations

d. **Monitoring Station Table**:

- Primary Key: StationID
- Attributes: Lat_Long (precise coordinates), Active_Time, DateOfInstallation, AQI (current reading)
- Purpose: Represents physical air quality monitoring equipment and their current status

e. **Weather Table**:

- Primary Key: WeatherID
- Attributes: Temp, Humidity, WindSpeed, Rainfall, Timestamp, Category
- Purpose: Stores meteorological readings associated with air quality measurements

f. **Health Advice Table**:

- Primary Key: AdvisoryID
- Attributes: HealthEffects, Precautions, AgeAffected
- Purpose: Contains health recommendations based on pollution levels for different demographic groups

g. **Junction Tables**:

- Belongs: Links Users and City (many-to-many)
- Monitored: Associates City with Monitoring Station (one-to-many)
- Gets: Connects Monitoring Station with Weather readings (one-to-many)

4.2 Entity-Relationship Model

The ER diagram depicts conceptual data organization through:

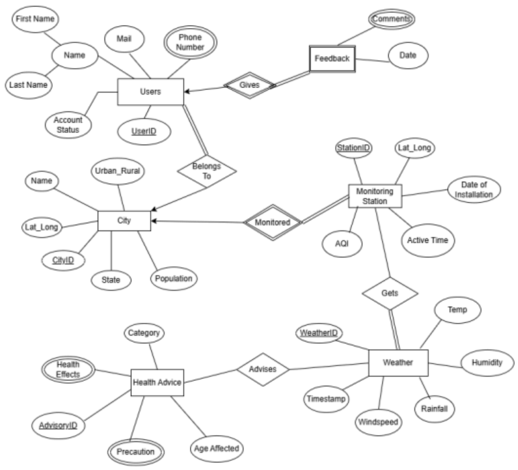

Figure 4.1: Entity Relation Diagram of AQI Management System

- **Entities (Rectangles)**:
  - Six primary entities (Users, City, Monitoring Station, Weather, Health Advice, Feedback)
  - Each entity contains attributes represented as ovals
- **Relationships (Diamonds)**:
  - "Gives": Users provide Feedback (one-to-many)
  - "Belongs To": Users are associated with Cities (many-to-many)
  - "Monitored": Cities are monitored by Stations (one-to-many)
  - "Gets": Stations receive Weather data (one-to-many)
  - "Advises": Weather conditions trigger Health Advice (many-to-many)
- **Attributes**:
  - Simple attributes: Direct properties (e.g., UserID, Password)
  - Composite attribute: Name (decomposed into FirstName, LastName)
  - Multivalued attributes: None present in this model

The ER model provides a conceptual view that was later transformed into the relational schema through proper mapping techniques.

4.3 Normalization Process

The database has been normalized to Boyce-Codd Normal Form (BCNF) through these steps:

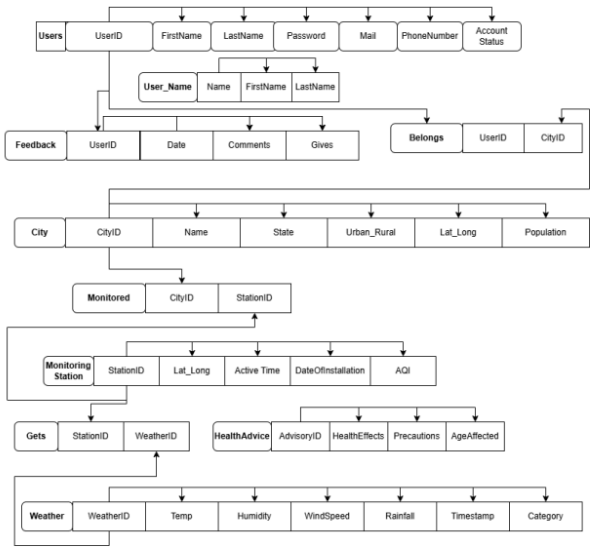

Table 4.2: BCNF Normalization of AQI Management System

- **1NF (First Normal Form)**:
  - Eliminated repeating groups
  - Ensured all attributes contain atomic values
  - Decomposed composite attribute "Name" into "FirstName" and "LastName"
  - Created unique identifiers (primary keys) for each entity
- **2NF (Second Normal Form)**:
  - Satisfied 1NF requirements

- Removed partial dependencies on composite keys
- Created separate tables for many-to-many relationships ("Belongs" for Users-City)
- Ensured non-key attributes depend on the entire primary key

- **3NF (Third Normal Form)**:
  - Satisfied 2NF requirements
  - Eliminated transitive dependencies
  - Separated Weather data from Monitoring Station
  - Created independent Health Advice table

- **BCNF (Boyce-Codd Normal Form)**:
  - Satisfied 3NF requirements
  - Ensured every determinant is a candidate key
  - Restructured tables to eliminate overlapping candidate keys
  - Created junction tables ("Gets") to maintain BCNF compliance

This normalization approach minimizes data redundancy while maintaining data integrity through properly defined foreign key relationships. The schema balances normalization principles with query performance considerations for the air quality monitoring domain.

Chapter 5

Methodology:

5.1 Environmental Parameter Aggregation Functions

The system implements three key functions to analyze meteorological data:

1. **Humidity Analysis**: Calculates city-specific average humidity through multi-table joins (Cities → Monitoring Stations → Weather Data)1
2. **Temperature Analysis**: Computes mean temperature using identical joins but focused on thermal metrics
3. **AQI Calculation**: Derives average air quality index by correlating cities with their monitoring stations

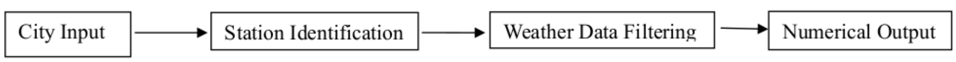

Figure 5.1: Process flow of Environmental Parameter Aggregation Function

5.2 Health Advisory Generation Procedure

A cursor-based procedure matches computed AQI values with predefined health guidelines:

1. **AQI Range Matching**: Compares city's average AQI against stored advisory thresholds
2. **Dynamic Recommendation**: Returns precautionary measures and risk levels through OUT parameters
3. **Error Handling**: Implements exception management for missing data scenarios

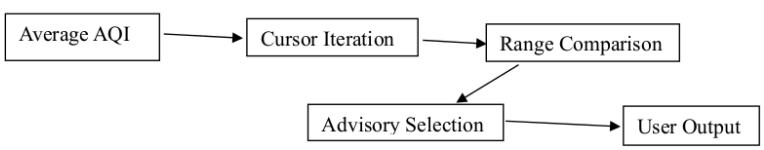

Figure 5.2: Decision Flow of Health Advisory Generation Procedure

5.3 Meteorological Analysis Queries

Two nested queries provide granular weather insights:

1. **Rainfall Assessment**: Calculates precipitation averages through triple-table join (Cities ↔ Stations ↔ Weather)
2. **Wind Pattern Analysis**: Determines mean wind speeds using same relational pathways

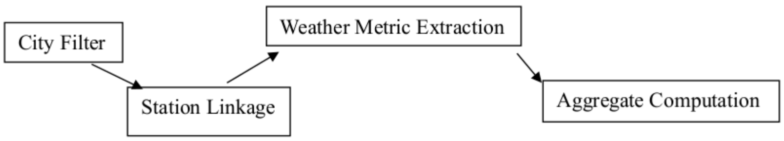

Figure 5.3: Data Flow of Meterological Analysis Queries

5.4 Automated Data Integrity Mechanisms

Two triggers enforce system reliability:

1. **Feedback Moderation**: Blocks submissions containing prohibited terms ("bad", "trash") using pattern matching
2. **User Insertion Notification**: Provides immediate confirmation messages during account creation

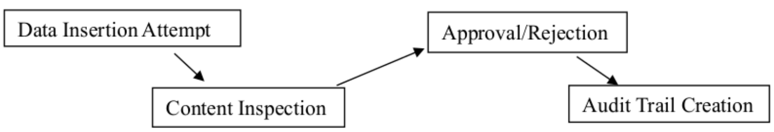

Figure 5.4: Validation Flow of Automated Data Integrity Mechanisms

Chapter 6

Results

The results achieved so far demonstrate the successful implementation of an Air Quality Index (AQI) Management System that integrates air quality data, weather parameters, and health advisories. The login interface ensures secure user authentication, while the map provides a visual representation of AQI levels across India, enabling users to search for city-specific data efficiently. Detailed city dashboards, such as the one for Mumbai, present air quality metrics including AQI, temperature, humidity, wind speed, and rainfall in an intuitive format. Furthermore, the health advisory section offers actionable recommendations tailored to specific AQI ranges and vulnerable demographics, enhancing public awareness and safety during adverse air quality conditions. These results validate the system's ability to deliver accurate environmental insights and personalized health guidance through a user-friendly interface.

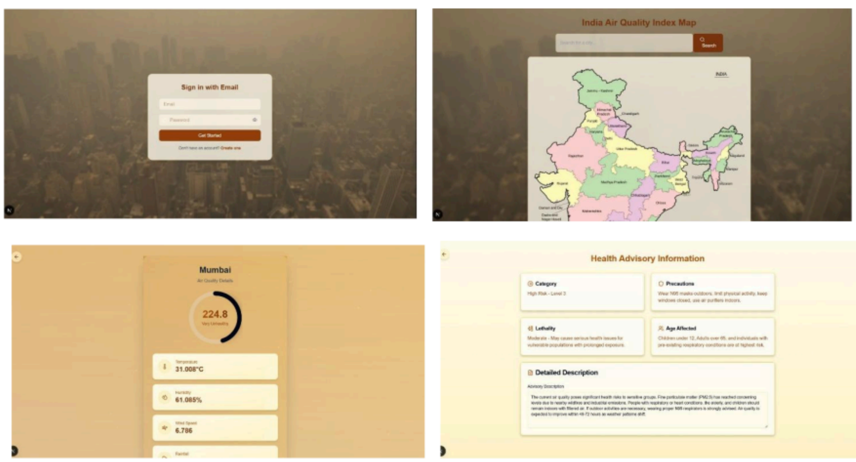

Figure 6.1: Displaying AQI of Mumbai

Chapter 7

Conclusion and Future Work

7.1 Summary of Achievements

The AQI Management System database design successfully:

- Implements a fully normalized relational model for air quality data
- Establishes appropriate relationships between entities
- Supports complex queries for air quality analysis
- Enables personalized health advisories
- Facilitates user engagement through feedback mechanisms

The database design follows best practices in relational database modeling while addressing the specific requirements of air quality monitoring.

7.2 Limitations

Current limitations include:

- Limited support for time-series data analysis
- No built-in predictive modeling capabilities
- Potential scalability challenges with very high-frequency sensor data
- Limited integration with external data sources
- Basic health advisory generation without machine learning

7.3 Future Enhancements

Future work could include:

1. **Time-Series Optimization**: Implementing partitioning for weather data
2. **Predictive Modeling**: Adding support for pollution prediction algorithms
3. **Spatial Analysis**: Enhanced geographical query capabilities

4. **Machine Learning Integration**: Advanced health risk assessment
5. **API Development**: Standardized interfaces for external systems
6. **IoT Integration**: Direct connections to sensor networks
7. **Mobile Application Backend**: Specialized queries for mobile clients

7.4 Lessons Learned

Key insights from the project include:

- The importance of proper normalization in environmental data management
- Benefits of clearly defined entity relationships
- Value of junction tables in handling complex relationships
- Necessity of considering both current and historical data access patterns
- Importance of balancing normalization with query performance

References

1. World Health Organization, "WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide," Geneva: World Health Organization, 2021.
2. U.S. Environmental Protection Agency, "Technical Assistance Document for the Reporting of Daily Air Quality," EPA-454/B-18-007, Research Triangle Park, NC, 2018.
3. Ministry of Environment, Forest and Climate Change, Government of India, "National Air Quality Index," Central Pollution Control Board, 2014.
4. A. Kumar, H. Kim, and G. P. Hancke, "Environmental Monitoring Systems: A Review," IEEE Sensors Journal, vol. 13, no. 4, pp. 1329-1339, 2013.
5. P. S. Monks et al., "Atmospheric composition change – global and regional air quality," Atmospheric Environment, vol. 43, no. 33, pp. 5268-5350, 2009.
6. S. Pant, S. Kumar, and P. Garg, "Smart air quality monitoring system for urban and rural areas," in Proc. IEEE Int. Conf. Comput. Intell. Internet Things, Dec. 2020, pp. 91-96.
7. J. Zhou et al., "SmartCityDB: A Large-scale Sensor Database for City-scale Applications," Proc. 2020 Int. Conf. Manage. Data, pp. 2577-2580, 2020.
8. S. Chaudhuri, U. Dayal, and V. Narasayya, "An overview of business intelligence technology," Commun. ACM, vol. 54, no. 8, pp. 88-98, 2011.

ProjectReportDBSLab_230911168_230911172_230911174_A...

ORIGINALITY REPORT

12%	12%	6%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.coursehero.com Internet Source	2%
2	ecis2016.org Internet Source	1%
3	Alireza Gogani, Faezeh Choobkar, Asli Cekmis. "Chapter 24 An Assessment of Thermal Comfort in Urban Quality of Life in Architecture Using Fuzzy Logic in Decision Making: A Case Study of Iran", Springer Science and Business Media LLC, 2024 Publication	1%
4	Frank R. Spellman, Nancy E. Whiting. "Environmental Management of Concentrated Animal Feeding Operations (CAFOs)", CRC Press, 2007 Publication	1%
5	www.dailypioneer.com Internet Source	1%
6	curve.carleton.ca Internet Source	1%
7	www.mdedge9-ma1.mdedge.com Internet Source	1%
8	www.scribd.com Internet Source	1%
9	letuslearncoding.com Internet Source	1%

10	Ulas Im, Zorana Andersen. "Air quality and health impacts", Elsevier BV, 2025 Publication	<1 %
11	compass.rauias.com Internet Source	<1 %
12	doaj.org Internet Source	<1 %
13	hl-128-171-57-22.library.manoa.hawaii.edu Internet Source	<1 %
14	core.ac.uk Internet Source	<1 %
15	impressions.manipal.edu Internet Source	<1 %
16	www.westerninvestor.com Internet Source	<1 %
17	newsroom.ferrovial.com Internet Source	<1 %
18	www.seeleyinternational.com Internet Source	<1 %
19	pmc.ncbi.nlm.nih.gov Internet Source	<1 %
20	www.sierraclub.org Internet Source	<1 %
21	arrow.tudublin.ie Internet Source	<1 %
22	www.cdc.gov Internet Source	<1 %
23	Stephen Battersby. "Clay's Handbook of Environmental Health", Routledge, 2022 Publication	<1 %