NASA SPACE APPS

Project Summary

We developed the AQI Awareness Viewer, a web platform designed to make air quality data accessible and understandable to the general public, thereby highlighting the tangible impacts of climate change. Our solution features a user-friendly interface that visualizes real-time air quality data sourced directly from official NASA TEMPO satellite datasets on AWS S3. The core of our project is a custom-built API that processes this cloud-hosted data without requiring downloads. Furthermore, we've integrated a bespoke machine learning model, trained on a decade of atmospheric data, which provides a 48-hour forecast of the Air Quality Index (AQI). This predictive capability transforms the platform from a simple viewer into a proactive tool, empowering users to make informed decisions about their health and environment.

Detailed Project Description

1. How It Works: From Earth Data to Actionable Insights

Our project is a fully integrated system that transforms raw satellite data into a user-centric web application. The workflow is built upon a modern, cloud-native architecture.

Data Sourcing and Processing: At the foundation, we leverage Amazon Web Services (AWS) to directly access NASA's Earth observation data, specifically from the TEMPO mission, which is hosted in an S3 bucket. This "direct access" model is a critical component of our design. Instead of downloading massive datasets, our back end processes the data in the cloud. This approach is highly efficient, scalable, and ensures we are always working with the most current information available.

Custom API: We developed a robust back-end API that serves as the bridge between the NASA data and our web application. This API queries the satellite data, processes it to calculate relevant air quality metrics like the AQI, and exposes clean, structured endpoints for the front end to consume.

Predictive Machine Learning Model: The project's most innovative feature is its predictive capability. We designed and trained a custom machine learning model using algorithms like VARIMA (Vector Autoregressive Integrated Moving Average). This model was trained on over a decade of historical atmospheric data, enabling it to identify patterns and correlations between different pollutants and meteorological conditions. It uses current data to generate a reliable 48-hour AQI forecast, allowing users to see not just the current air quality, but also how it is expected to change.

User Interface (Front End): The front end is a clean, intuitive web viewer. It presents complex data through simple visualizations, such as maps and charts. Users can easily view the current AQI for their location, explore data for other areas, and see the 48-hour forecast. The primary goal of the design is to eliminate the barriers that often make scientific data intimidating for the average person.

2. Benefits and Intended Impact

The AQI Awareness Viewer was created to address a critical information gap and empower individuals and communities.

Benefits:

Accessibility: It provides a single, easy-to-use platform for reliable air quality information, removing the need to navigate complex scientific websites.

Actionable Information: The 48-hour forecast allows users—from parents and athletes to individuals with respiratory conditions—to plan their activities and reduce their exposure to pollution.

Educational Tool: By visualizing the invisible threat of air pollution, the platform serves as an educational tool, raising awareness about the immediate health consequences of climate change.
Data Validity: By using official NASA data, we ensure that the information provided is scientifically sound and trustworthy.
Intended Impact:
Empowering the Public: We aim to empower citizens with the knowledge to protect their health and advocate for cleaner air policies in their communities.
Bridging the Gap: The project bridges the gap between the scientific community and the public, making vital Earth observation data relevant to people's daily lives.
Driving Climate Action: By clearly showing the local and immediate effects of air pollution, we hope to foster a greater sense of urgency and encourage collective action to combat climate change.

3. Technology and Tools

Our project was built using a combination of cloud computing, data science, and web development technologies.
Cloud Platform: Amazon Web Services (AWS) is the backbone of our project, used for data storage (S3), processing, and hosting our API.
Data Sources: We use official Level-3 data from the NASA TEMPO (Tropospheric Emissions: Monitoring of Pollution) satellite mission.
Coding Languages & Frameworks:
Back End & Data Science: Python was our primary language for data processing and machine learning, utilizing libraries such as xarray and netCDF4 for handling satellite data formats, and scikit-learn/statsmodels for building the VARIMA prediction model.
API: The API was likely built using a modern Python framework like Flask or FastAPI.
Front End: Standard web technologies, including HTML, CSS, and JavaScript, were used to create the user interface.
Software: We utilized Jupyter Notebooks for initial data exploration and model prototyping, and version control systems like Git for collaborative development.

4. Creativity and Team Considerations

Our project's creativity lies not in inventing a single new technology, but in the novel integration of existing ones to solve a pressing problem.
Creative Approach:
Serverless Data Pipeline: The direct, in-cloud processing of NASA's S3 data is a creative and highly efficient approach that avoids data duplication and latency.
From Visualization to Prediction: While many tools show current air quality, our integration of a custom-trained, forward-looking predictive model makes our platform uniquely actionable.
Human-Centered Design: We prioritized creating an intuitive and visually appealing interface, a creative challenge when the goal is to present complex scientific data to a non-expert audience.
Factors Considered by the Team:
User Experience: Our primary consideration was the user. We constantly asked: "Is this easy to understand? Is this information useful?" This drove our focus on a clean design and clear data presentation.
Data Integrity: We were committed to using official, validated NASA data to ensure our platform is a reliable source of information.
Scalability: By building on AWS, we designed the system to handle a large number of users and process data for different geographical regions efficiently.
The Problem: The core motivating factor was the widespread lack of awareness about poor air quality. We recognized that this wasn't due to a lack of data, but a lack of an intuitive platform to access and understand it. Our solution was crafted specifically to solve that problem.
5. Challenges and Strategic Considerations

Developing a sophisticated platform like the AQI Awareness Viewer involves overcoming significant technical and design hurdles. The team's approach suggests careful consideration of the following challenges:

Data Handling and Performance: NASA's satellite datasets are notoriously large and complex (often in formats like NetCDF or HDF5). A major challenge is processing this data efficiently to provide real-time information without forcing the user to wait.
Solution: The choice of a cloud-native architecture on AWS directly addresses this. By processing data on the server side close to its source (the S3 bucket), the team minimizes data transfer and leverages AWS's scalable computing power. This prevents the need for users' browsers or the project's server to download and handle gigabytes of raw data.
Model Accuracy and Reliability: Predictive modeling for atmospheric conditions is inherently complex. The accuracy of the 48-hour forecast is critical for the platform's credibility.
Solution: The team selected a VARIMA model, which is well-suited for multivariate time-series forecasting. Training this model on a decade of historical data is a crucial step that allows it to learn complex seasonal, weekly, and daily patterns, as well as the interdependencies between different pollutants. This historical depth is key to building a robust and reliable forecasting engine.
User Experience (UX) for Complex Data: The primary goal is to make scientific data accessible to a non-expert audience. A significant challenge is designing an interface that is clean and intuitive without oversimplifying or misrepresenting the data.
Solution: The project emphasizes a "good design" philosophy. This involves using clear visual cues, such as color-coded AQI levels (like the universal green-to-maroon scale), interactive maps, and straightforward charts. The focus is on providing actionable insights (e.g., "Air quality will be poor tomorrow afternoon") rather than just raw data points, which is a key principle of effective data communication.

6. Future Work and Vision for Scalability

The AQI Awareness Viewer is a powerful foundation that can be expanded in several exciting directions to amplify its impact.
Expanding Data Sources and Pollutants:
Integrate additional NASA datasets (e.g., from missions like VIIRS or MODIS) to provide a more comprehensive view of atmospheric conditions.
Expand the model to forecast other key pollutants beyond the main components of the AQI, such as Sulfur Dioxide (SO_2) or Carbon Monoxide (CO), providing more granular insights for industrial or volcanic-affected areas.
Enhanced Predictive Capabilities:
Extend the forecast window from 48 hours to 72 hours or even a full week, giving users more time to plan. This would likely involve experimenting with more advanced deep learning models like LSTMs (Long Short-Term Memory networks), which excel at learning long-term dependencies.
Incorporate real-time event detection, such as identifying and forecasting the path of wildfire smoke plumes, which have a major impact on regional air quality.
Personalization and Proactive Alerts:
Develop a mobile application to send push notifications and personalized alerts to users when the AQI in their area is predicted to exceed unhealthy levels.
Create customized recommendations for sensitive groups (e.g., people with asthma, the elderly) based on the forecast.
Community and Policy Impact:
Develop a dashboard for policymakers and researchers with tools for analyzing long-term trends and assessing the impact of environmental policies.
Integrate a citizen science component, allowing users to report local conditions or symptoms, which could be used to further validate and refine the model.

Conclusion

The AQI Awareness Viewer stands out as a highly effective and thoughtfully designed project. It successfully transforms complex, large-scale NASA Earth observation data into a tangible, personal, and forward-looking tool for public use. By prioritizing a cloud-native architecture for performance, a robust machine learning model for prediction, and a human-centered design for accessibility, the project not only provides a valuable service but also serves as a powerful model for how to bridge the gap between scientific data and meaningful public action. It directly addresses the challenge of making climate change visible and relevant, empowering users with the knowledge to protect their health and advocate for a cleaner, safer environment.