# **Week 2 Report: System Design and Architecture**

During Week 2 of the ‘**InsightNation – Government Data Analytics Platform for Citizen Opinion and Public Service Enhancement’** project, the focus shifted from planning to the Design Phase, where the primary objective was to lay down the architectural and functional blueprint of the data analytics platform. This phase was crucial in transforming conceptual ideas into a clear technical foundation that would guide the implementation in later stages.

The week began with a thorough re-evaluation of the project objectives and scope defined in Week 1. The platform’s core objective remains: to analyze and interpret citizen feedback for smarter and more responsive public service delivery. The refined scope now clearly targets municipal and local governance bodies as the key end-users. It includes functionalities such as data ingestion, preprocessing, exploratory data analysis, sentiment analysis using NLP, predictive modeling, and interactive dashboards for insight delivery. These functionalities are designed to empower stakeholders with timely, data-driven decisions.

The team then consolidated the high-level system requirements. Essential features were identified, including support for structured and unstructured feedback data, automated data cleaning modules, robust sentiment and topic modeling capabilities, and real-time visualization dashboards. The platform must handle diverse data formats (CSV, survey data, user comments) and provide modular capabilities for expansion. Requirements were mapped closely to the original project objectives to maintain alignment and ensure feasibility within the 12-week timeline.

Efforts were made to gather relevant datasets and development resources. Multiple sample datasets related to public service satisfaction and urban feedback were reviewed, and one primary dataset (cleaned\_citizen\_feedback.csv) was selected as the development and testing base. Resource requirements such as Python libraries (spaCy, pandas, scikit-learn, seaborn), visualization tools, Streamlit for front-end development, and the Gemini API for insight generation were finalized. The most significant deliverable this week was the creation of the system architecture and data flow design. A comprehensive diagram was developed to visualize the entire data lifecycle—from ingestion and preprocessing to analytics modeling and dashboard deployment. Key components include:

* Data Sources (CSV feedback data)
* Preprocessing Engine (cleaning, formatting)
* NLP Pipeline (SpaCy-based sentiment and topic analysis)
* ML Models (for predictive and classification tasks)
* Visualization Layer (Streamlit dashboard)
* AI Insight Engine (via Gemini API)

This architecture ensures modularity, reusability, and scalability while remaining aligned with the project’s analytical goals.

Parallel to the backend design, work also began on interface wireframes and dashboard mock-ups. Initial wireframes were created for the user dashboard, including elements such as data upload interfaces, dropdown filters (e.g., by city or age group), sentiment trend visualizations, topic summaries, and AI-generated insights. The wireframes focus on clarity and user experience, designed with decision-makers in mind. Feedback was noted for future iterations and refinements in upcoming weeks.

In summary, Week 2 concluded with a well-defined architecture, design mockups, and a solid understanding of how data and analytics will flow through the InsightNation platform. This design foundation will directly support the development and modeling work in the upcoming phases.