

# PROJECT ASSIGNMENT 1

## Team Information

### Team Member 1:

- **Name:** Daniel Evans.
- **Background:** Undergrad degrees in Linguistics & Russian Language, with coursework in mathematics and computer science. Strong background in the structure of natural language. Good understanding of some python ML, NLP, scraping tools. Management experience in the food service industry, using forecasting tools for sales and inventory management.
- **Proposed Role:** Data Preprocessing
- **Project Impact:** Cleaning, normalizing, and structuring data for the model. Ensure a smooth handoff between ingestion and the model.

### Team Member 2:

- **Name:** Joel Vinas
- **Background:** Former Business, Financial and Operations Analyst with experience building ETL Systems, and back-end databases for shipping and retail merchandising logistics. Currently a Solutions Architect, developing solutions with Jira & Workfront for T-Mobile. Proficient in [VB.NET](#) and SQL.
- **Proposed Role:** Data Engineer, developing the database structure and ingestion tools.
- **Project Impact:** In coordination with Daniel and Tony, I will assist in the development of the ETL system, and in improving the AI/ML solution to achieve optimal results.

### Team Member 3:

- **Name:** Tony Nguyen.
- **Background:** Graduate student in Data Science and Analytics with a background in supply chain operations. Proficient in Python (Pandas/NumPy), SQL, and Power BI. Experience with logistics software (YMS, Relay, SCC) and operational efficiency.
- **Proposed Role:** AI/ML Engineer (Domain logic & Modeling OR Spatio-Temporal Modeling) or Data Engineer
  - **Project Impact (AI/ML)**
    - I will build the development of KG+LLM architecture. My role is to bridge the gap between technical data and real-world logistics; I will define the risk parameters for “delay predictions” based on my experience with high-volume shipment disruptions and bottlenecks and adjust a shipment’s delivery path in real-time disruptions like weather events, traffic, or site congestion.

- I will design the multimodal forecasting model. I will integrate data sources (weather, sensors, events) and ensure the predictive outputs are accurately visualized for city planning.
  - **Project Impact (Data):** I will be responsible for the "Live Data Ingestion" and "Streaming" components of the project. This includes building the ETL (Extract, Transform, Load) pipelines to pull raw data from sources like the World Bank, US DOT, or OpenWeather API. I will ensure the data is normalized, cleaned, and structured into a high-performance database so that the AI/ML model can access accurate, real-time features for prediction.
- 

### Project Idea 1: Smart Supply Chain Optimization System

- **Objective & Impact:** The goal is to **predict delays and reroute shipments**. This impacts logistics efficiency by minimizing disruptions using global data.
  - **Models & GenAI Methods:** We will use Knowledge Graph (KG) to map a complex supply chain dependencies and an LLM to interpret unstructured disruption reports.
    - **RAG (Retrieval-Augmented Generation):** To provide a real-time, context-aware rerouting recommendations based on historical disruption patterns.
  - **Planned Deliverable:** A system featuring **live data ingestion, graph-based routing**, and an **optimization dashboard**.
    - **Graph-based routing logic:** To help design the nodes and edges of the Knowledge Graph (KG) based on real-world supply chain nodes (hubs, ports, and transit lanes).
    - **Operational Validation:** To develop a "validation framework" to ensure the AI's rerouting suggestions are compliant with standard logistics safety and efficiency protocols.
  - **Data Sources:**
    - **World Bank Logistics Data:** for global performance indicators
    - **OpenSupplyHub:** for accurate mapping of supply chain facilities and nodes.
- 

### Project Idea 2: Urban Mobility and Traffic Analytics

- **Objective & Impact:**
  - **Problem:** To optimize traffic flow and predict congestion using sensor, weather, and event data.
  - **Relevance:** This system supports city planning agencies, autonomous vehicle companies, and infrastructure AI platforms.
- **Models and GenAI Methods:**
  - **Multimodal:** To process weather, sensor, and event data simultaneously.
  - **Fine-tuning:** To adapt forecasting models to specify city topologies and local traffic patterns.

- **Planned Deliverable:**
  - **Streaming ingestion** of sensor data.
  - **Spatio-temporal forecasting model.**
  - **Interactive city dashboard.**
- **Data Sources:** US DOT Traffic Data, OpenWeather API, and City Open Data Portals.