Project DRISHTI

A Data-Driven and AI-Powered Framework for Intelligent and Sustainable Transportation Planning

Team Name: The Ragnarok

Submission Date: September 8, 2025

1. Introduction & Team Details

This proposal identifies a holistic and forward-thinking solution aimed at transforming transportation data collection and usage in planning. **Emphasizing the capture of precise trip-related data like origin, destination, time, mode, and trip chains, the project addresses head-on the perennial shortcomings of conventional manual household surveys, which are labor-intensive, expensive, and only account for a minute portion of the population**. Seizing the potential of Artificial Intelligence and mobile-based data gathering, Project DRISHTI seeks to develop a scalable, reliable, and consent-driven ecosystem for producing high-quality mobility datasets. Through this project, NATPAC scientists and policymakers will be empowered with actionable insights to make smarter, evidence-backed decisions to create a safe, sustainable, and efficient transport system for society.

Team Information

Team Leader:	Akhil Astitva(23BCE11678)	**Role:** Project Architect & Lead Developer
Contact:	akhilastitva365@gmail.com	+91-9720590884

Team Members:

1. *Srajal Jain(23BCE10773)* - Backend Developer
2. *Sparsh Jain(23BCE10512)* - AI/ML Engineer
3. *Shri Ram Prince Mishra(23BCE11629)* – Mobile App Developer (Frontend)
4. *Palak Bisht(24BEY10053)* - Research & QA Specialist
5. *Shrey Pandey(23BCE10944)* - Data Analyst / Visualization Specialist

Mentor:

Dr. Ajeet Singh

Assistant Professor Grade 1

School of Computer Science and Engineering(SCOPE) **ajeetsingh@vitbhopal.ac.in**

2. Problem Statement Selection

Problem Statement ID/Code: SIH25082

Problem Statement Title:

Development of a travel related software app that can be installed on mobile phones that could capture trip related information

Organization: Government of Kerala

The selected problem statement fits ideally with our group's mission to leverage technology for the greater good. **It identifies the urgent demand for a precise, scalable, and privacy-aware mechanism for gathering mobility data, which forms the core of efficient transportation planning. Existing mechanisms like manual household surveys are disjointed, time-intensive, and sample only a minority of the population, resulting in poor data for policy-making and infrastructure planning**. Our solution addresses these issues head-on with the introduction of a mobile-based, AI-supported, and cloud-powered ecosystem that guarantees end-to-end reliable, real-time trip data collection. By equipping NATPAC with high-quality datasets, the project establishes the platform upon which smarter, safer, and more sustainable transport systems can effectively serve the needs of society.

3. Idea Details

Idea Title: Project DRISHTI

Idea Summary: Project **DRISHTI** is a broad, three-level solution aimed at transforming transport planning based on precise, large-scale mobility data capture. **The solution comprises a citizen mobile app, a NATPAC scientist-centralized analytics dashboard**, and a highly capable backend supported by AI and cloud technologies. **With real-time trip characteristics like origin, destination, mode, and purpose captured, the project allows for the creation of daily trip chains and mobility patterns**. The overall objective is to transition from restrictive, time-consuming human surveys to a scalable, automated, and privacy-aware system that facilitates wiser transport planning for everybody.

Detailed Solution Description:

Our proposed solution is a socio-technical system designed to bridge the existing gaps in mobility data collection and planning. It is architected for scalability, privacy, and reliability. The core components are:

1. **Citizen Mobile Application (iOS/Android):** A convenient app which auto-detects journeys with GPS and sensor on phone. Enables users to enter missing information like purpose of journey or fellow travellers.
2. **NATPAC Analytics Dashboard:** A web-based, centralized dashboard for NATPAC planners and scientists. Offers interactive visualizations of trip frequency, mode share, mobility patterns, and Origin-Destination (OD) matrices.
3. **Backend Ecosystem:** A secure, cloud-native backend for processing and storing trip data. AI/ML modules for autonomous travel mode detection and trip chain formation. Geo-analytics tools for mapping congestion areas, travel lanes, and trip densities

4. **Methodology**

We will adopt an Agile development methodology with a focus on iterative development and user feedback. The project will be divided into the following phases:

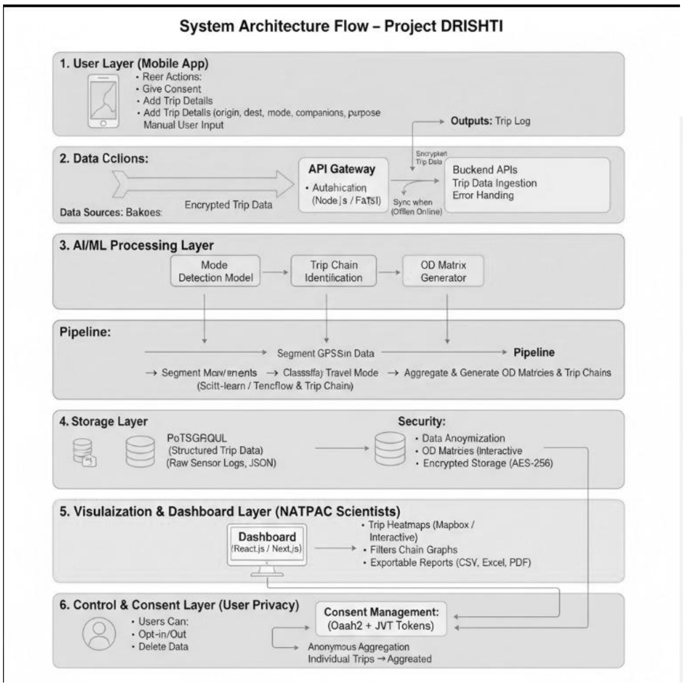
1. **Phase 1 (Proof-of-Concept):** Create a mobile app prototype where users can manually enter trip information (origin, destination, time, mode, companions). Simple GPS-based trip logging to identify motion and log trip start/stop events. Store data collected in a secure server-based/local database.
2. **Phase 2 (MVP):** Implement AI/ML models for trip segmentation (automated detection of trip boundaries). Implement mode of transport detection (walk, car, bus, etc.) using sensor data and APIs. Implement a simple web dashboard for NATPAC scientists to see aggregated trip data, O-D matrices, and simple visualizations. Implement consent-based nudging where users approve or modify auto-detected trip information.
3. **Phase 3 (Full-Scale Development):** Improve AI/ML models with enhanced datasets for greater precision in mode detection and trip chain analysis. Roll out a feature-rich NATPAC Dashboard with heatmaps, trip chain analytics, and downloadable reports. Introduce advanced privacy features like anonymization and secure data handling. Streamline the app for improved UX, scalability, and integration with policy-making tools.

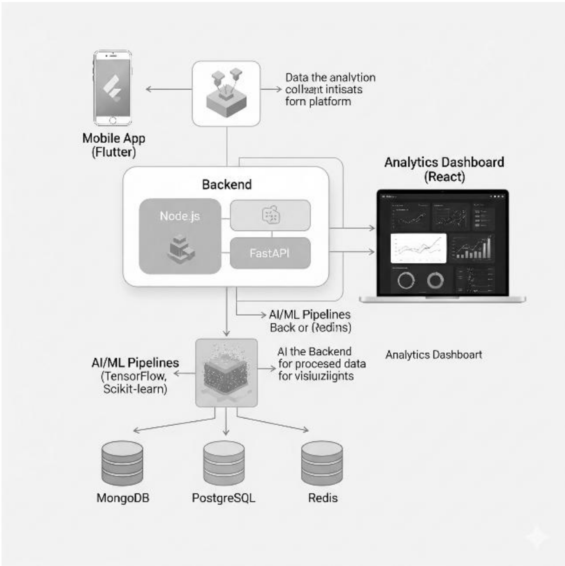
Key Features & Deliverables

- **Trip Recording & Logging:**
  A mobile application that captures trip details such as origin, destination, time, mode of transport, and companions. Trips can be recorded automatically via GPS and sensors or manually entered by the user.
- **AI-Powered Mode Detection:**
  Machine learning models analyze smartphone sensor data (GPS, accelerometer, speed) to predict the mode of travel (walk, car, bus, train, etc.) with minimal user intervention.
- **Trip Chain Analysis:**
  Automatic detection of linked trips (e.g., *home → work → shopping → home*) to establish a complete daily travel pattern for better planning insights.
- **Consent-Based Data Collection:**
  Users are always in control—data is collected only with explicit consent. Nudges are provided to confirm or correct automatically detected trips, ensuring accuracy and privacy.
- **NATPAC Scientist Dashboard:**
  A web-based visualization tool for transport planners to analyze aggregated trip data, view heatmaps, generate O-D matrices, and extract actionable insights for policy decisions.
- **Secure Data Management:**
  All data is anonymized and stored securely in a structured database, ensuring user privacy while enabling high-quality analytics for transportation research.

Technical Specifications

Component	Tech Stack / Tools
Frontend (Mobile)/(DashBoard)	Flutter, Google Activity Recognition, Google Maps SDK,React JS
Backend	Node.js(Express.js) ,FastAPI & JWT/OAuth2
AI/ML	Python, TensorFlow, Scikit-learn, Hugging Face (NLP models), Pytorch, Pre-trained datasets/apis, LLM integration
Security & Privacy	End-to-End Encryption (TLS/HTTPS), Data Anonymization Pipeline, Consent Management Module
Databases	NoSQL(MongoDB), PostgreSQL , Redis
Deployment	Local/NATPAC servers , NGINX, Docker





User Flow:

The user flow is designed to ensure **frictionless adoption** by households and commuters.

- Download & Installation: Consumers download the app, set up a lightweight profile (no extensive KYC required), and authorize GPS and activity recognition permissions.
- Passive Data Collection: The application identifies travel activities by default (walking, cycling, bus, car, train) and records origin, destination, date and time, mode, trip purpose (if asked).
- User Confirmation & Feedback: Consumers are allowed to verify occasionally detected trips by responding to notifications (e.g., "Did you just travel by bus?") in order to enhance model performance.
- Data Sync: Anonymized, encrypted trip data is synced regularly to the NATPAC central server.
- Planner Dashboard: Researchers and policy makers view the dashboard to see Origin-Destination matrices, mode share, chains of trips, and congestion hotspots.

AI/ML Algorithms & Agents:

Trip Detection & Mode Classification

- **Algorithm:** A **Hybrid ML Model** using **Random Forest/XGBoost** trained on smartphone sensors (GPS, accelerometer, gyroscope, Wi-Fi, cell tower data).
- **Inputs:**
  - GPS velocity & stop durations
  - Accelerometer patterns (walking vs. vehicle)
  - Route-matching with transit data (bus/train schedules, Google Transit feeds)
  - Time-of-day & trip chaining behavior
- **Outputs:**
  - Detected trip segments (home → work, work → shop, etc.)
  - Travel mode (walk, bike, car, bus, train, auto-rickshaw)
  - Trip purpose (optional user input to train future models)

AI Risk & Anomaly Detection (Optional Enhancement)

- Detect abnormal travel patterns such as **unusually long commutes, detours, or inconsistent modes**.
- Helps **flag unreliable data** and improve survey integrity.

NLP Agent (Optional AI Assistant)

- A **multilingual AI chatbot** integrated in the app to:
  - Guide users on survey participation.
  - Answer FAQs about data privacy.
  - Provide travel tips (eco-friendly options, congestion alerts).
- Built using **RAG (Retrieval-Augmented Generation)** + **preloaded NATPAC guidelines/FAQs**.

ML Pipeline Workflow

1. Data Collection (Mobile App)

- Raw sensor + GPS logs → locally stored in SQLite → periodic upload.

2. Preprocessing

- Noise filtering (GPS drift correction).
- Trip segmentation (stop-detection algorithm).

3. Feature Engineering

- Extract velocity, acceleration, time intervals, stop frequency, route overlap with transport networks.

4. ML Model Processing

- Mode detection (Random Forest/XGBoost).
- Trip purpose classification (semi-supervised learning).

5. Data Storage

- Anonymized, structured records stored in **PostgreSQL (trip tables)**.
- Raw logs stored in **MongoDB** for validation.

6. Dashboard & Visualization

- Origin-Destination matrices, heatmaps, congestion maps.
- Accessible through **React + D3.js dashboard**

System Events Example

Trip Detection Event

1. App detects movement start via accelerometer.
2. GPS records trajectory & velocity.
3. ML model predicts mode = bus.
4. When trip ends, system logs origin, destination, mode, start time, end time.
5. Data is anonymized & synced to NATPAC server.

Dashboard Example (Planner View)

- **Real-time OD matrices** (Home → Work, Work → Market, etc.).
- **Mode Share Charts** (e.g., Car 40%, Bus 25%, Bike 20%).
- **Trip Chains** (e.g., Home → School → Office → Shop → Home).
- **Hotspot Maps** (visualizing congestion & travel density).

4. Feasibility & Impact

Feasibility

The proposed solution is highly feasible as it leverages mature and cost-effective technologies. The mobile app will be built using Flutter, enabling a single codebase for both Android and iOS. For the backend, FastAPI/Django with a PostgreSQL + MongoDB database will handle structured and unstructured data efficiently.

Our AI/ML models for mode detection and trip classification can be trained using publicly available mobility datasets (e.g., GPS trajectory datasets, Google Transit feeds) and further refined with regional travel survey data from NATPAC. The architecture avoids costly cloud dependencies by relying on open-source frameworks and lightweight deployment options (DigitalOcean/Railway VPS), making it both technically and financially viable. **Scalability &**

Scalability & Sustainability

Project DRISHTI is designed for **nationwide scalability**.

- The **modular architecture** allows expansion from a pilot in the Northeast region to other states, without major redesigns.
- Trip detection models can continuously improve as **more real-world data** is collected, ensuring adaptability over time.

- Sustainability is ensured by integration with **government agencies, transport authorities, and urban planning bodies**, making it a valuable long-term data collection tool.
- The system reduces the cost and inefficiency of manual household surveys, freeing resources for policymakers while providing richer datasets.

Potential Impact

The successful implementation of Project DRISHTI will have a **transformative effect on transportation planning**:

- **Data-Driven Policy Making:** Enables planners to create accurate Origin-Destination (OD) matrices, mode share reports, and congestion analyses.
- **Improved Urban Mobility:** Facilitates better route planning, infrastructure development, and identification of underserved areas.
- **Time & Cost Savings:** Eliminates dependence on small-scale manual surveys, capturing large-scale data automatically.
- **Enhanced Public Engagement:** Provides citizens with a transparent way to contribute to planning, increasing trust in government initiatives.

User Experience (UX)

The mobile app is designed to be **minimalist, intuitive, and non-intrusive**.

- Data collection runs **silently in the background**, minimizing user effort.
- Occasional, **simple feedback prompts** (e.g., verifying detected mode of travel) make participation easy without survey fatigue.
- A **gamified experience** (badges, contribution scores) can motivate users to stay engaged.
- The planner dashboard will feature **interactive visualizations** (OD heatmaps, charts, congestion hotspots) with **role-based access control** for government officials and researchers.

5. References & Research links

Core Concepts & Research Areas

- **Mobility Data Analysis & Trip Detection:** Research on using smartphone sensor data to identify travel modes and patterns. (Link to IEEE Xplore)
- **Origin-Destination (OD) Matrix:** The concept of using mobile data to estimate travel flows, replacing traditional surveys. (Link on ScienceDirect)
- **Privacy-Preserving Data Collection:** Techniques like data anonymization to protect user privacy in location-based services. (Link to a survey paper)

Technologies & Frameworks

- **Flutter:** A cross-platform framework for building the iOS and Android mobile app. (https://flutter.dev/)
- **React JS:** A JavaScript library for building the interactive NATPAC scientist dashboard. (https://react.dev/)
- **Node.js (Express.js):** A backend JavaScript runtime for building server-side applications. (https://nodejs.org/)

- **FastAPI:** A high-performance Python web framework for building APIs. (https://fastapi.tiangolo.com/)
- **TensorFlow:** An open-source platform for machine learning. (https://www.tensorflow.org/)
- **Scikit-learn:** A Python library for traditional machine learning algorithms like Random Forest. (https://scikit-learn.org/)
- **PyTorch:** An open-source machine learning framework known for its flexibility. (https://pytorch.org/)
- **Hugging Face:** A platform for NLP models, useful for the optional AI chatbot. (https://huggingface.co/)
- **MongoDB:** A NoSQL database for storing unstructured data like raw sensor logs. (https://www.mongodb.com/)
- **PostgreSQL:** An object-relational database system for structured trip data. (https://www.postgresql.org/)
- **Redis:** An in-memory data store used for caching and fast data retrieval. (https://redis.io/)

Relevant APIs & Datasets

- **Google Maps SDK:** For integrating maps and location services into the app. (https://developers.google.com/maps)
- **Google Activity Recognition API:** To detect user activities like walking, cycling, or driving. (https://developers.google.com/location-context/activity-recognition)
- **General Transit Feed Specification (GTFS):** The standard format for public transit schedules and geographic data. (https://gtfs.org/)
- **GeoLife GPS Trajectories:** A public dataset from Microsoft Research for training initial ML models. (Link to dataset)

Organizations & Context

- **NATPAC (National Transportation Planning and Research Centre):** The client organization for this project. (https://natpac.kerala.gov.in/)
- **Government of Kerala:** The overarching government body associated with the project. (https://kerala.gov.in/)
- **Smart India Hackathon (SIH):** The innovation competition context for the problem statement. (https://www.sih.gov.in/)

Deployment & Security

- **Docker:** A platform for containerizing and deploying the application. (https://www.docker.com/)
- **NGINX:** A high-performance web server and reverse proxy. (https://www.nginx.com/)
- **JWT (JSON Web Tokens):** A standard for creating access tokens for authentication. (https://jwt.io/)
- **OAuth 2.0:** The industry-standard protocol for authorization. (https://oauth.net/2/)