Safe-Karachi

Kaavish Project Proposal

By

Jibran Sheikh	(js08312@st.habib.edu.pk)
Wania Hussain	(sh08450@st.habib.edu.pk)
Ghulam Mustafa	(gm08291@st.habib.edu.pk)
Daniyal Farooqui	(mf08320@st.habib.edu.pk)
Ikhlas Ahmed	(ia08308@st.habib.edu.pk)

September 28, 2025

In partial fulfillment of the requirement for

Bachelor of Science

Computer Science

Dhanani School of Science and Engineering

Habib University

Fall 2023

1 Abstract

Karachi, Pakistan's economic hub and most productive city, has long been ranked among the least livable urban centers in the world. Despite its role as the country's financial backbone, residents continue to struggle with daily challenges that compromise both safety and mobility. These include rampant street crime, chronic traffic congestion, broken or flood-prone roads, poor street lighting, civic disruptions, and the lack of timely and reliable information about conditions on the ground. Existing navigation apps are designed to optimize routes based only on distance and time. While useful, they overlook the equally critical dimension of safety. This leaves commuters vulnerable when a faster route passes through unsafe or disrupted areas. The absence of a system that prioritizes both speed and security creates a gap in urban navigation services for the people of Karachi. SafeKarachi addresses this gap by offering a mobile-first platform that enables safer and more informed travel across the city. Instead of focusing solely on traffic, it brings together information about crime, road conditions, and civic disruptions to guide residents toward routes that are not just quicker, but also safer and more reliable. The app gathers information from three primary sources: reports submitted directly by users, updates shared on social media, and credible news outlets. This information is then organized and simplified into clear updates, displayed on an interactive map of Karachi. To ensure trust, an internal system evaluates the credibility of each report and classifies it as strong, weak, or uncertain. These updates are shared with users as headline-style alerts within the app, helping them remain informed about both local and city-wide developments. Community engagement is another key element of SafeKarachi. Users can join area-based groups—such as Gulshan, Johar, or Saddar—to interact, share updates, and engage in discussions. These groups function like neighborhood forums but are directly linked to issues of safety and mobility. This creates a collaborative environment where people contribute to the flow of real-time information while also benefiting from the reports of others. The navigation system goes beyond conventional shortest-path routing. It factors in distance, live traffic, road conditions, and even historical patterns of incidents to provide multiple route options. By combining real-time updates with predictive insights, SafeKarachi helps users avoid high-risk areas and choose safer alternatives without significantly increasing travel time. Throughout their journey, users remain continuously updated through in-app notifications that deliver timely, context-aware alerts. Whether it is a traffic disruption, civic protest, flooding, or a safety incident nearby, the system ensures that people receive relevant updates tailored to their location and travel patterns. This ongoing flow of information allows them to make quick, confident decisions on the move. The initial version of SafeKarachi will include user onboarding, incident reporting, community engagement, live maps, route recommendations, predictive insights, and integrated notifications.

2 Problem definition

Karachi, despite being Pakistan's economic hub, consistently ranks among the least livable cities due to persistent challenges in safety, infrastructure, and mobility [1]. Residents face unpredictable commutes, broken or flooded roads, and constant personal safety risks. Street crime—phone and vehicle theft, extortion, and harassment—creates widespread fear, with the CPLC estimating that up to 80% of incidents go unreported [3]. This leaves citizens dependent on unreliable updates from informal networks such as WhatsApp, Facebook groups. The transport system further compounds these issues, with congested corridors, inadequate drainage, and poor lighting turning routine trips into hazards. Evaluations under the Karachi Mobility Project emphasize that accessibility, reliabil-

ity, and safety must be addressed together, yet current navigation apps only optimize for time and distance [5]. Citizens therefore lack a credible, real-time tool to stay informed and to make safer, more reliable travel decisions.

3 Social relevance

Karachi's citizens face daily risks from rampant street crime, frequent traffic accidents, and poor infrastructure. CPLC statistics show thousands of incidents each month, including widespread mobile and motorcycle snatching, while media reports highlight the psychological toll of residents constantly fearing victimization [4, 3]. Traffic accidents add hundreds of fatalities annually, worsened by broken roads, flooding, and long-delayed civic projects that create hazards and blockages. Current navigation apps like Google Maps focus only on time and distance, offering no warnings about crime-prone spots, accident-heavy intersections, or closed roads. SafeKarachi addresses this gap by merging crime, accident, and civic disruption updates into a single, reliable platform that provides safer route options and real-time alerts. The project also benefits from external collaboration with Mr. Irtiza Ali (IBA graduate, Acumen), whose direct awareness of Karachi's street-crime and civic challenges ensures the solution remains grounded in real community needs.

4 Originality/Novelty

Addressing Karachi's safety and mobility gap provides clear, measurable value for its residents: fewer exposures to street crime, more predictable commutes, and greater confidence in everyday travel. Karachi repeatedly ranks among the world's least livable cities, reflecting chronic shortfalls in security, infrastructure, and services that residents feel on every trip. Tackling these issues directly improves quality of life and also unlocks broader economic participation at scale (Tribune2025).

Key Contributions

- **Personal safety:** Transforms citizen, social, and news reports into structured intelligence, with AI-based classification of credible vs. weak signals. Prior research shows safety-aware routing increases distance from crime hotspots while keeping routes efficient [7].
- **Reliability:** Goes beyond time/distance by factoring in flooding, closures, lighting, and other risks—aligned with Karachi Mobility Project priorities on safety and accessibility [5].
- **Information quality:** Uses NLP pipelines to clean and structure noisy chatter, adapting methods proven in platforms like GDELT and CrisisNLP [8, 9].
- **Trust:** Responds to chronic underreporting of crime (up to 80% unrecorded) by cross-checking and elevating credible signals [3].

Comparison with Existing Solutions

While navigation and safety apps exist, SafeKarachi is novel in its integration of safety-weighted routing, real-time classification, and community engagement:

- **Google Maps/Waze:** Focus only on time and distance, ignoring crime and infrastructure risks [10, 11].
- **Citizen, CrimeMapping, SpotCrime, Ring Neighbors, Zuluhood:** Provide alerts or community feeds but lack integration with routing, hotspot prediction, or Karachi-specific coverage [14, 15, 17, 16, 18].

5 CS contribution

The SafeKarachi project applies advanced Computer Science concepts to address real-world challenges of safety and mobility in Karachi.

1. Natural Language Processing (NLP)

We use multilingual NLP techniques such as classification and named entity recognition to process noisy reports in English, Urdu, and Roman Urdu. This enables extraction of structured intelligence from social media, news, and user inputs.

2. Machine Learning and Prediction

Spatio-temporal models are applied to forecast crime hotspots and disruptions. Core ML and Deep Learning concepts guide risk scoring, model calibration, and evaluation, ensuring reliable predictive insights.

3. Databases and Geospatial Systems

Using PostgreSQL/PostGIS and MongoDB, we design schemas and spatial queries to support city-scale data storage and retrieval for real-time routing and incident reporting.

4. Algorithms and Routing

Building on graph algorithms (e.g., Dijkstra, A*), routing is customized with safety weights such as crime density and road conditions, enabling safer path recommendations.

5. Software Engineering and Systems Design

Principles from Software Engineering and Systems Design are applied in modular APIs, microservices, and CI/CD pipelines to ensure scalability, maintainability, and resilience.

6 Scope and Deliverables

Scope

The project will develop a mobile-first platform that integrates community reporting, real-time safety updates, and risk-aware navigation. Key components include:

- React Native app with map visualization and community interaction.

- Node.js/Express backend with authentication, APIs, and data storage.
- Integration of user reports, social media feeds, and news sources.
- Multilingual NLP pipeline for classification and entity extraction.
- Machine learning models for hotspot prediction and safety scoring.
- Safety-aware routing engine (OSRM with custom weights).
- Trust & safety features such as credibility scoring and moderation.

Out of scope are automated image/video analysis, city-wide deployment, and direct government API integrations, which are reserved for future extensions.

Deliverables

The final deliverables will include:

1. **Mobile App:** Registration, incident reporting, community groups, live maps, safe-route recommendations, and push alerts.
2. **Backend Services:** REST/GraphQL APIs, PostgreSQL/PostGIS for geospatial queries, and MongoDB for unstructured reports.
3. **AI/NLP Modules:** Multilingual classification and NER, hotspot prediction models, and a credibility scoring system.
4. **Documentation & Demo:** A final report and live demonstration of the end-to-end system.

7 Feasibility

SafeKarachi is technically feasible owing to the availability of diverse data sources, stable APIs, and mature open-source libraries that enable end-to-end development within the project timeline.

Data Sources

Incident data will be aggregated from **user reports**, **credible news websites** (Dawn, GeoNews, ARY), and **social media streams** (Twitter/X, Facebook groups). Supplementary datasets, such as CPLC crime statistics, will provide additional ground truth.

Tools and Libraries

Routing will combine **Google Maps API** (routing, geocoding) with **OpenStreetMap + OSRM** for open-source, customizable pathfinding. Data ingestion will use **Facebook Graph API**, **NewsAPI/World News API**, and web scraping with **Scrapy** and **Selenium**. NLP pipelines will rely on **spaCy**, **NLTK**, and regex (`re`) for entity extraction and classification, while PostGIS will support spatial queries, hotspot visualization, and heatmaps.

Challenges and Mitigation

- **Imperfect Data:** Addressed via a reliability scoring model combining source credibility, corroboration, and report volume.
- **Multilingual Text:** Managed with multilingual NLP models for English, Urdu, and Roman Urdu.
- **API Costs:** Mitigated by fallback to open-source OSRM/OSM for routing and scraping-based ingestion.

Resources and Hardware

Development requires only standard laptops and a modest cloud VM for backend hosting and databases. NLP tasks will run efficiently on CPUs, with GPUs (Colab/Kaggle) used during training of heavier multilingual models. The mobile client will run on standard Android/iOS devices without additional hardware dependencies.

Performance

The pipeline is designed for near real-time use: news/API queries respond in seconds, NLP classification in sub-seconds, and PostGIS spatial queries in milliseconds. OSRM enables sub-second route calculations, ensuring reporting, map updates, and safe-route suggestions remain responsive for daily commuters.

8 Team Dynamics and Suitability

Member & Role	Responsibilities	Strengths	Weakness	Mitigation
Ghulam Mustafa Backend & APIs (MongoDB)	Service design; MongoDB REST; Postgres/PostGIS schema; auth/rate-limits; ingestion jobs; observability	Spring Boot backend internship; DB/SE/DSA coursework; GIS/Earth Engine exposure	Limited city-scale throughput/DevOps	Queue-based ingestion; caching for read-heavy endpoints; early metrics/alerts. *Evidence:* Spring Boot REST work; DB & SE courses.
Jibran Sheikh Mobile (React Native [RN]) & UX	RN app; map/alerts UX; FCM push; offline states; performance & testing	Android app with push/auth; React/Node projects; web-GIS apps; HCI/UX exposure	Less native-module/geofencing depth	Use stable SDKs (Mapbox/Expo Maps); isolate native bridges; E2E tests on report/alert *Evidence:* Android internship; RN projects.
Daniyal Farooqui NLP Pipeline (Classification & NER)	Scraping/ETL; taxonomy; classifier & NER; geo/time extraction; human-in-the-loop labeling	NLP coursework; routing/graph project; GIS RA; Python ETL	Limited production NLP deployment	Iterate: rule-augmented classical models → small transformer baseline; active-learning with curator tool. *Evidence:* NLP course projects; graph/routing work.

Member & Role	Responsibilities	Strengths	Weakness	Mitigation
Ikhlas Ahmed				
ML Prediction (Hotspots & Risk)	Spatio-temporal modeling; risk scoring; calibration; model evaluation; MLOps handoff	ML/DL projects (e.g., transformers); strong eval/metrics; stats background	Lighter GIS tooling	Pair with backend for PostGIS features; documented feature store; calibration (Brier/PR). *Evidence:* DL coursework/projects.
Syeda Wania				
Data Science & SDP Policy	Data collection, cleaning, preprocessing; data QA/metrics; SDP: privacy/consent, T&S policies, moderation playbooks; stakeholder liaison	InfoSec internship (security/compliance); survey & mixed-methods research; Python/pandas; GEE suitability analysis	Less mobile/UI implementation	Focus on DS/SDP track; reviewer tooling & SOPs; in-app policy surfaces & user education. *Evidence:* InfoSec experience; research methods courses; DS projects.

9 Tech stack

SafeKarachi will be built on a modern, scalable stack that integrates mobile, backend, geospatial, and AI components:

- **Mobile:** React Native (TypeScript) with Mapbox for maps and Firebase for push notifications.
- **Backend:** Node.js/Express APIs with authentication and integration services.
- **Databases:** PostgreSQL/PostGIS for spatial queries, MongoDB for unstructured data, and Redis for caching.
- **Routing:** OSRM with OpenStreetMap, extended with safety weights (crime, flooding, road quality).

- **AI/NLP:** Python (FastAPI) services using spaCy, HuggingFace, scikit-learn, and PyTorch for classification, entity extraction, and hotspot prediction.
- **Data Ingestion:** APIs and scraping pipelines (Celery/Kafka, Scrapy, Selenium) for social media and news.
- **Security:** Firebase Auth with role-based access and data privacy controls.
- **DevOps:** Docker for deployment and GitHub Actions for CI/CD.

10 References

References

[1] The Express Tribune. "Karachi fourth worst city in global livability index 2025." 2025. Available: https://tribune.com.pk/story/2551326/karachi-fourth-worst-city-in-global-livability-index-2025

[2] The Express Tribune. "Street crime epidemic." 2025. Available: https://tribune.com.pk/story/2538103/street-crime-epidemic

[3] Ali, Imtiaz. "Frightening statistics of crimes in city only 'tip of the iceberg'." *Dawn*, January 4, 2024. Available: https://www.dawn.com/news/1803089

[4] Citizens-Police Liaison Committee (CPLC). "Karachi Crime Statistics – August 2025." Available: https://web.cplc.app/Statistic/2025/AUGUST_2025.pdf

[5] World Bank. *Pakistan – Karachi Mobility Project (P166732)*. 2019. Available: https://documents.worldbank.org/en/publication/documents-reports/documentdetail/327981561946644468/pakistan-karachi-mobility-project

[6] World Bank. "Karachi Mobility Project (P166732)." Washington, D.C.: World Bank, 2025. Available: https://documents1.worldbank.org/curated/en/099530005162260146/pdf/P16673206235d80a408be3035d2857d7bf7.pdf

[7] Levy, Daniel, et al. "SafeRoute: Learning to Navigate Streets Safely in an Urban Environment." Proceedings of NIPS, 2018. Available: https://ebelding.cs.ucsb.edu/sites/default/files/publications/saferoute.pdf

[8] GDELT Project. "Global Database of Events, Language, and Tone." 2021. Available: https://www.gdeltproject.org/

[9] CrisisNLP. "Crisis Data Analysis Tools." 2020. Available: https://crisisnlp.qcri.org/

[10] Google. "Google Maps Help." 2023. Available: https://support.google.com/maps

[11] Waze. "Waze Help Center." 2023. Available: https://support.google.com/waze

[12] Google. "Android Auto Help." 2023. Available: https://support.google.com/androidauto/answer/6348166?hl=en

[13] Citizens-Police Liaison Committee (CPLC). "CPLC Emergency+ App." 2023. Available: https://web.cplc.app/simple_powerful/

[14] Citizen. "Citizen App." 2023. Available: https://citizen.com

[15] CrimeMapping. "CrimeMapping.com." 2023. Available: https://www.crimemapping.com/

[16] Ring. "Ring Neighbors." 2023. Available: https://ring.com/neighbors

[17] SpotCrime. "SpotCrime." 2023. Available: https://spotcrime.com/

[18] Zuluhood. "Zuluhood: Safer and Connected Communities." 2023. Available: https://www.zuluhood.com/blog/safer-and-connected-communities

[19] Dawn News. "Road crashes killed nearly 500, injured 4,879 in Karachi in 2024." Feb 2025. Available: https://www.dawn.com/news/1891767

[20] Dawn News. "Karachi residents face traffic jams due to repair work on tattered roads in flood aftermath." Aug 2025. Available: https://www.dawn.com/news/1937093

[21] The Express Tribune. "University Road repair." Mar 2025. Available: https://tribune.com.pk/letter/2256652/university-road-repair

[22] Business Recorder. "We are barely surviving: delay in Karachi's Karimabad underpass project takes toll on nearby businesses." 2025.
Available:https://www.brecorder.com/news/40379696/we-are-barely-surviving-delay-in-karachis-karim

Undertaking of Kaavish advisement as an External Supervisor

I hereby affirm that I have read the project details as described on the preceding pages and agree to undertake advisement of this Kaavish project as an External Supervisor. I understand that this role entails the following.

Meeting Meeting the project team regularly, at least once every two weeks, for the entire duration of the Kaavish. The meetings may be held remotely if required.

Advisement Providing supervision and advice to the team in order to ensure steady progress of the project toward its goals.

Liaison Liaising with the Internal Supervisor as required, e.g. to provide feedback or engage in grading.

Other Any other task, depending on availability and suitability, relevant to the Kaavish as communicated by the Internal Supervisor or Kaavish Working Group.

Name: _____

Email: _____

Phone: _____

Designation: _____

Affiliation: _____

Signature: _____