NAV.AI

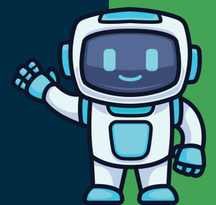
Overview About Us Services

Contact Us

NAV.AI

NAV.AI : The Future of Delivery, Spoken

Theme: Economic empowerment through AI (from Grab’s vision + AI)



UMHACKATHON 2025

PREPARED BY KERISBYTE

NAV.AI

OBJECTIVE

Safe Driver Interaction

Enables hands-free, voice-activated interactions

Speech Recognition Accuracy

Improves clarity in noisy environments

Resilience in Diverse Environments

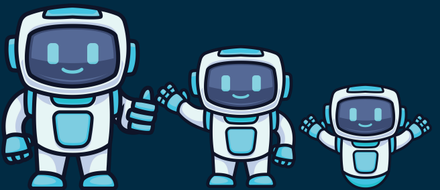
Maintains performance in various conditions

Adaptability to Speech Patterns

Handles diverse accents and languages

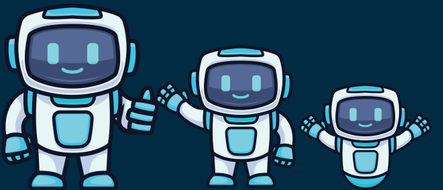
Reliability with Partial Clarity

Functions effectively with imperfect audio



NAV.AI

SOLUTION OF ARCHITECTURE



PREPARED BY KERISBYTE

NAV.AI

KEY CHALLANGES (1)

Traffic And Road Noise

Constant urban noise masking speech and reducing accuracy.

Vehicle Engine Sounds

Intrusive engine noise overpowering driver's voice.

Weather Conditions

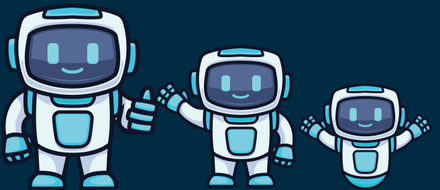
Adverse weather disrupting speech clarity significantly.

Urban Ambient Noise

Complex urban sounds challenging voice system isolation.

Audio System Feedback

Feedback loop interfering with voice command precision.



NAV.AI

KEY CHALLANGES (2)

Speech Pattern Complexities

Regional Accents and Dialects

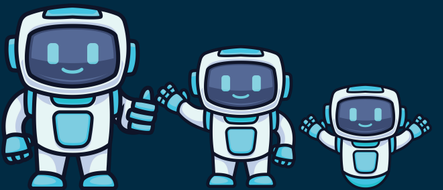
Accents and dialects vary pronunciation and intonation.

Variations in Speech Speed

Speech speed varies with comfort and urgency.

Use of Colloquial Expressions

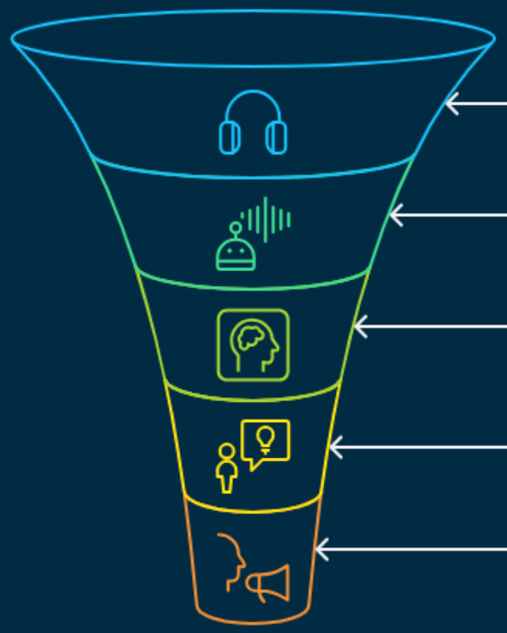
Colloquial expressions include slang and idioms.



NAV.AI

PROCESS

Human Speech Rhythm



Noise Filtering

Speech-to-Text Conversion

Text Interpretation

Response Generation

Text-to-Speech Conversion

User Interaction

1 Human Speech Rhythm

The process begins with human speech input.

2 Server Processing

The speech is processed and cleared on a server.

3 Speech to Text

Whisper is used to transcribe speech into text.

4 Text to LLM

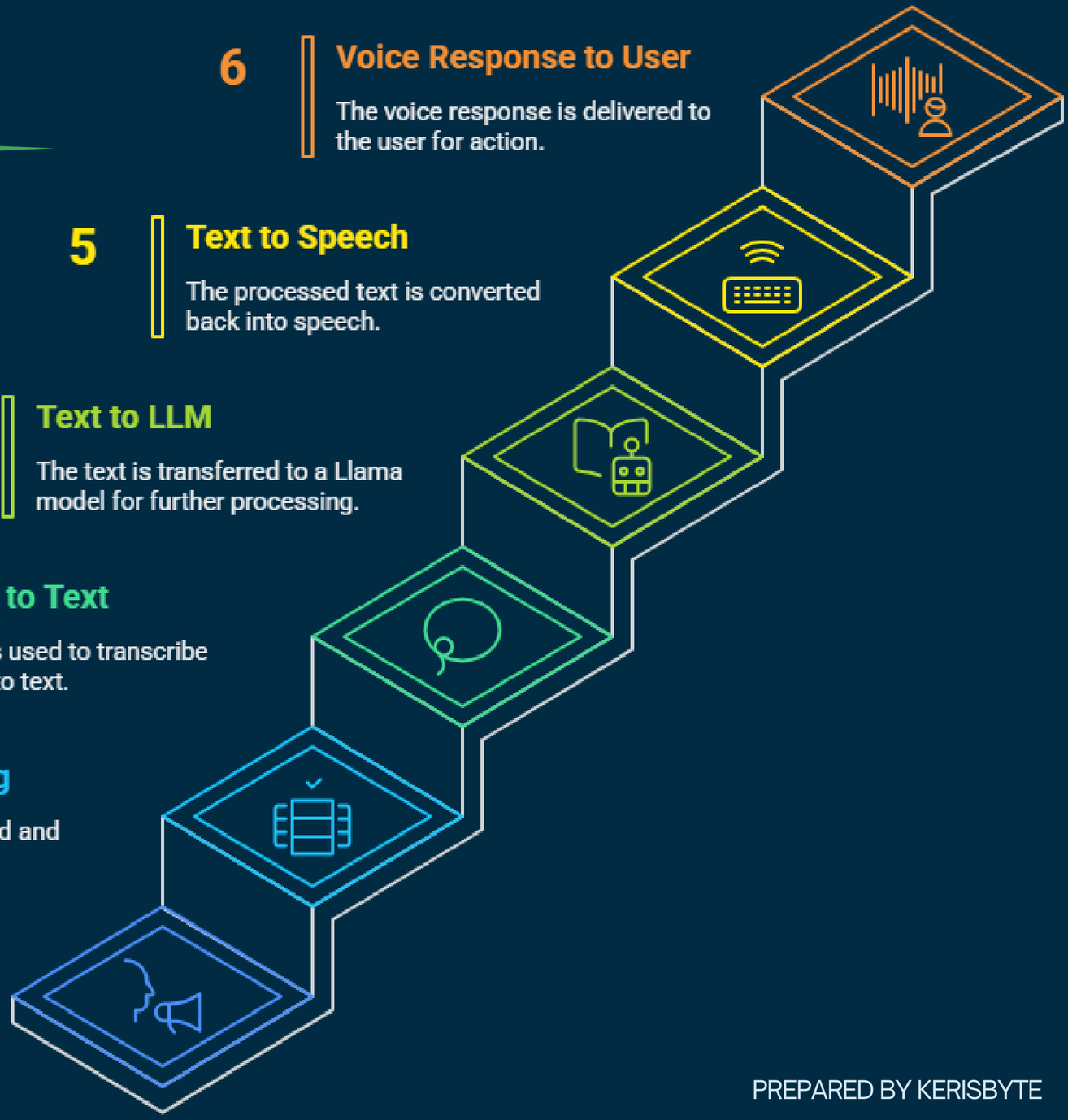
The text is transferred to a Llama model for further processing.

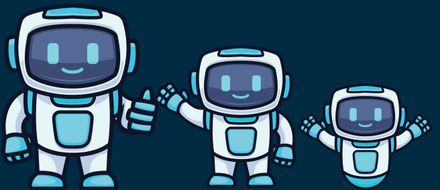
5 Text to Speech

The processed text is converted back into speech.

6 Voice Response to User

The voice response is delivered to the user for action.





PREPARED BY KERISBYTE

NAV.AI

LANGUAGE UNDERSTANDING PROCESS

Relevant Response Generation

The assistant generates context-aware responses.

Contextual Interpretation

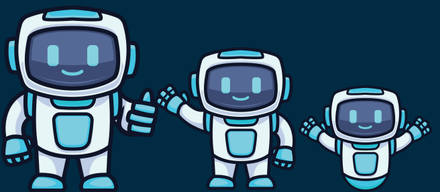
LLaMA 3.2 interprets the driver's intent.

Analysis with LLaMA 3.2

The text is analyzed using LLaMA 3.2.

Transcribed Text Input

The process begins with transcribed text input.



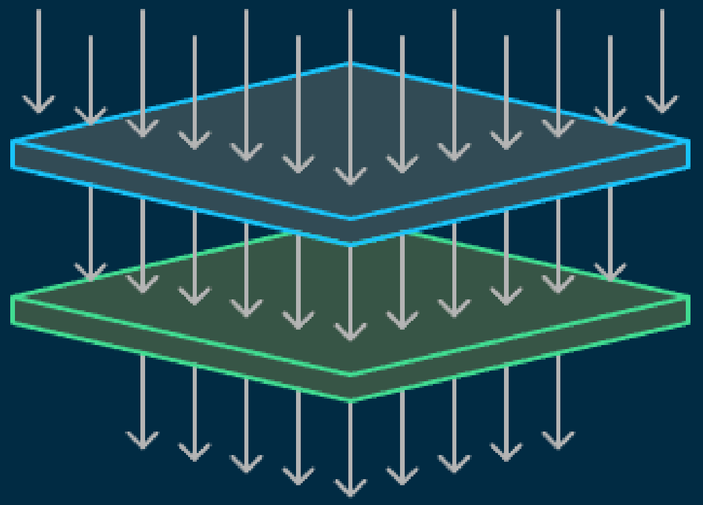
NAV.AI

KEYWORD DETECTION (LLAMA)

INCASE AUDIO OVER COMPROMISED
(Degraded Audio Quality)

Keyword Detection

Command Interpretation



Reliable Driver Commands

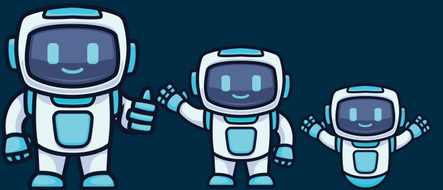
Reliable Command Interpretation

Robust in Noisy Environments

Precise Command Execution

Requires Clear Audio Input

Balancing Robustness and Precision in Speech Recognition



NAV.AI

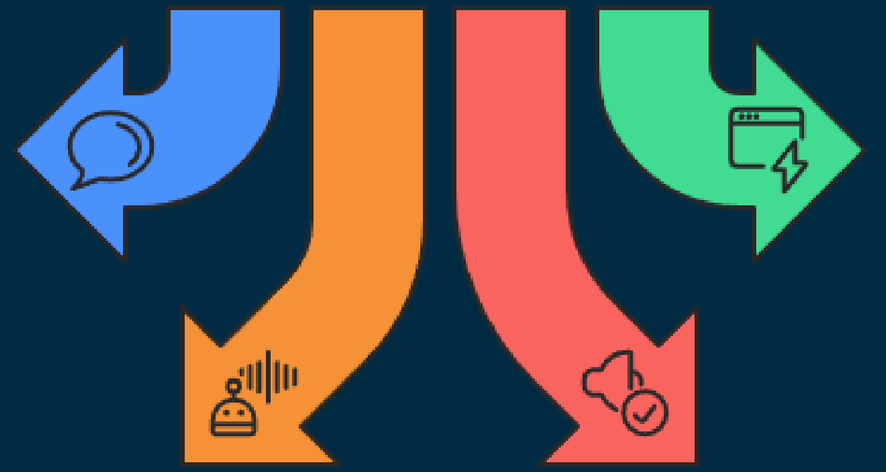
UI / UX

Voice-First Interaction

Focus on conversational commands and multilingual support for diverse users.

Minimal Visual UI

Use big fonts and animations for quick glances and feedback.

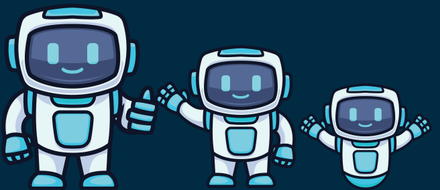


Smart Feedback System

Provide short, informative voice responses and interrupt options.

Fail-Safe Design

Ensure audible confirmation and fallback to visual interface.



NAV.AI

FEATURES

Display Maps

Displaying Google Maps for user interaction.

Real-time Tracking

Tracking user's location in real time.

Speech Input

Integrating speech-to-text for voice commands.

Speech to Search

Converting speech into location-based searches.

Draw Routes

Drawing polylines (routes) on the map.

State Management

Managing application state using Provider.

Voice Interaction

Real-time voice interaction using VOSK and PyAudio.

Audio Denoising

Audio denoising achieved through DeepFilterNet integration.

Multilingual Support

Supports multiple languages including Malay, Chinese, and English.

Smart Navigation

Smart navigation implemented with natural confirmation flow.

Model Integration

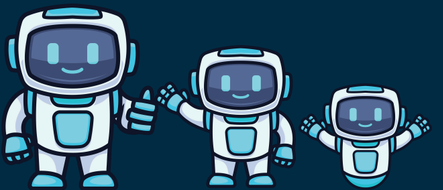
Integrates local LLaMA and OpenAI GPT-4 models.

JSON Support

Supports JSON command and function calling.

Adaptive Parsing

Adaptive intent parsing and conversation flow implemented.



NAV.AI

TECHNOLOGY APPLICATION

Voice Clarity Enhancement

DeepFilterNet ensures clear voice capture in noisy environments.

Natural Language Processing

LLaMa 3.2 interprets and responds to user queries intelligently.

Mobile Interface Development

React Native builds responsive and cross-platform mobile interfaces.

Speech Recognition

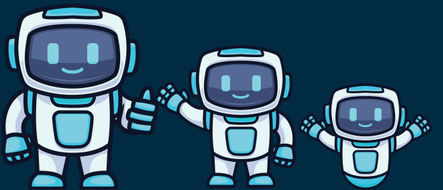
Whisper accurately converts speech to text across languages.

Rapid Prototyping

BOLT.New accelerates feature development and testing.

Backend Hosting

Google Cloud provides scalable and reliable hosting for AI services.



NAV.AI

SUMMARY OF TECHNOLOGY USED

DeepFilterNet	Whisper	LLaMa 3.2	BOLT.New	React Native	Google Cloud Console
Noise Suppression	Speech-to-Text	Language Understanding	Prototyping Framework	Mobile Interface	Backend Hosting
Enhances voice clarity	Converts speech to text	Interprets user queries	Rapidly prototypes features	Builds mobile interface	Hosts backend infrastructure