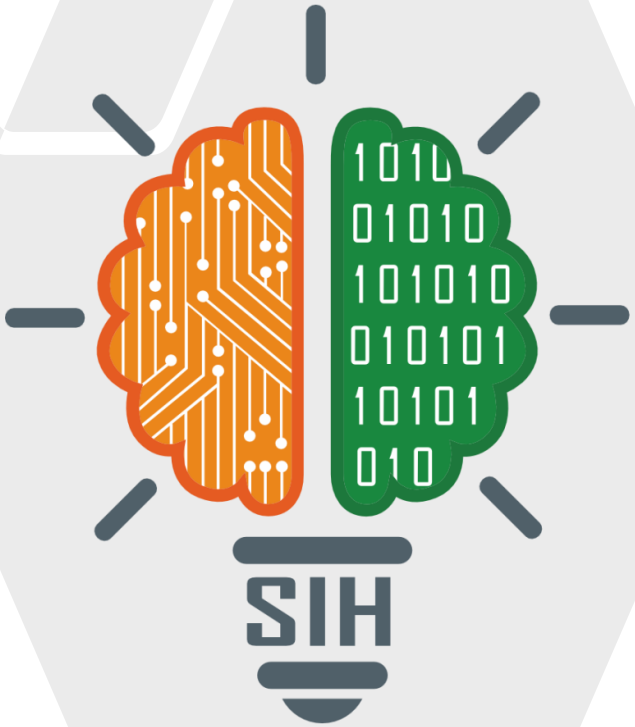
SMART INDIA HACKATHON 2025

- **Problem Statement ID – 25041**
- **Problem Statement Title - AI-Driven Unified Data Platform for Oceanographic, fisheries, and Molecular Biodiversity Insights**
- **Theme – Renewable / Sustainable Energy**
- **PS Category - Software**
- **Team ID – NA**
- **Team Name - Breaking Bugs**



Ministry of Earth Sciences
Government of India

Breaking Bugs

OceanusAI

SMART INDIA HACKATHON 2025

Problem Statement :

- Marine data in India is **siloed across formats** (oceanography, taxonomy, morphology, molecular/eDNA).
- Scientists, conservationists & policymakers **cannot cross-link datasets in real time**, slowing discovery & decision-making.
- India's **Blue Economy and food security** depend on sustainable fisheries & biodiversity conservation.
- Today, it takes **weeks to months** to manually correlate ocean parameters with biodiversity — too late for timely policy action.

Proposed Solution :

- User asks queries in natural language → routed to specific **domain RAG indexes** (oceanography, taxonomy, otolith morphology, eDNA).
- Left pane: **ChatGPT-style textual answers with evidence**.
- Right pane: **Cinematic interactive map** that visualizes zones, anomalies, and biodiversity patterns in real-time.
- IoT devices expand the knowledge base (temperature, salinity, camera/sonar) → live feeds integrated into platform.

Target Audience

Research Organisations

Government & Policy Makers

Investors /Startups

Environmentalists & Conservationists

Fisheries & Blue Economy stakeholders

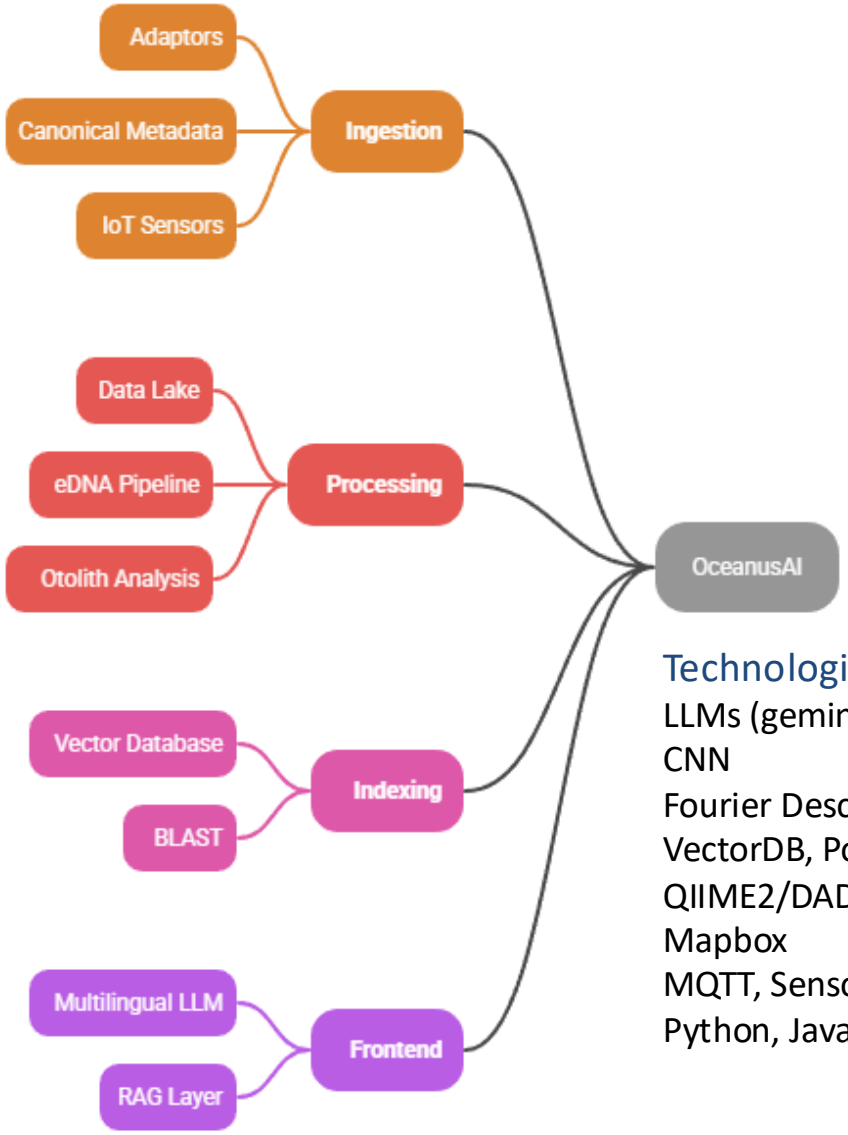
Our Model Promises :

- **Provides** unified access to oceanographic, taxonomy, otolith, and molecular datasets in one portal.
- **Ensures** real-time visualization + explainable AI answers for scientists & policymakers.
- **Supports** multilingual queries + cross-domain correlation for faster research.
- **Improves** decision-making for fisheries & biodiversity by **70% faster insight generation**.

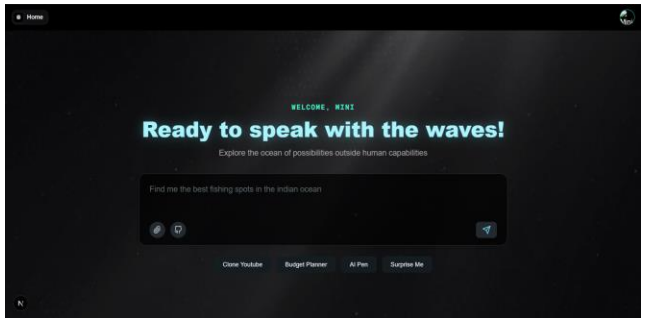
Breaking Bugs

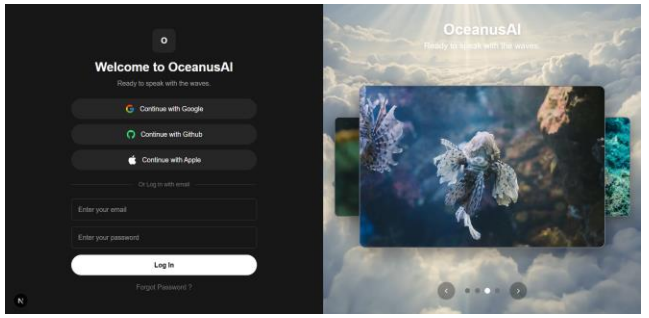
TECHNICAL APPROACH

SMART INDIA HACKATHON 2025

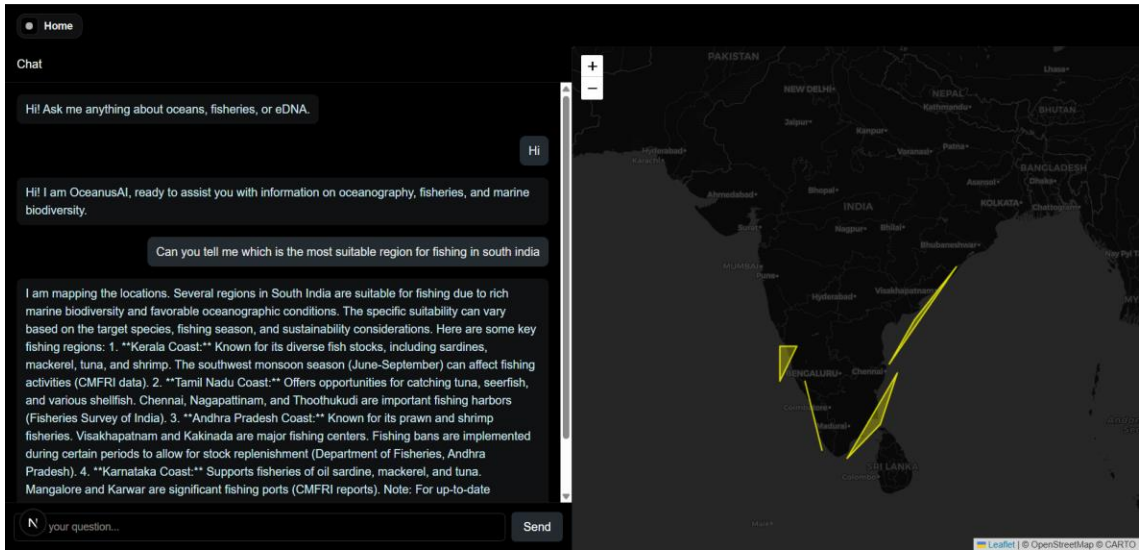


Input





Output



Technologies Used

LLMs (gemini/openAI
CNN
Fourier Descriptors
VectorDB, PostGIS
QIIME2/DADA2, BLAST
Mapbox
MQTT, Sensors (IOT)
Python, Javascript

Breaking Bugs

FEASIBILITY AND VIABILITY

SMART INDIA HACKATHON 2025

- Analysis of the Feasibility of the Idea

- **Technical Feasibility**: The solution uses proven tools (LLMs, RAG, vector DBs, QIIME2/DADA2, CNNs, Mapbox) that are already open-source and widely adopted. Integration is the key innovation, not building from scratch — making it technically achievable.
- **Economic Feasibility**: Cloud-native microservices and modular pipelines reduce infrastructure costs. Open-source libraries (Milvus, QIIME2, Deck.gl) keep expenses low, ensuring scalability without huge budgets.
- **Operational Feasibility**: Designed for CMLRE, research labs, and government users with simple Chat+Map UI. Multilingual support increases adoption across India. Role-based access ensures secure use by scientists, policymakers, and conservation groups.

- Potential Challenges :
  - Complex, heterogeneous datasets.
  - Accuracy issues in species detection or otolith classification.
  - IoT connectivity problems in deep-sea conditions.
  - Scaling costs for large datasets.
  - Trust issues with AI answers (hallucinations).

- Strategies to Overcome :
  - Standardize data with Darwin Core & NetCDF/CF formats.
  - Use curated databases + expert validation for biological data.
  - Edge computing & buffered IoT sync for reliability.
  - Cloud-native modular scaling with cost monitoring.
  - Explainable AI with citations & provenance for trust.

Breaking Bugs

IMPACT AND BENEFITS

SMART INDIA HACKATHON 2025

Impact :

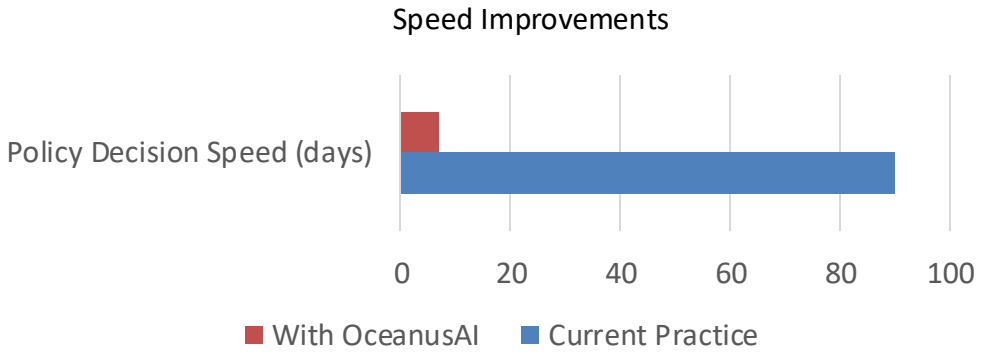
OceanusAI will act as India's national marine data backbone, enabling scientists, policymakers, and conservationists to access unified insights instantly. It reduces weeks of manual data processing into real-time, evidence-backed answers for sustainable ocean management.

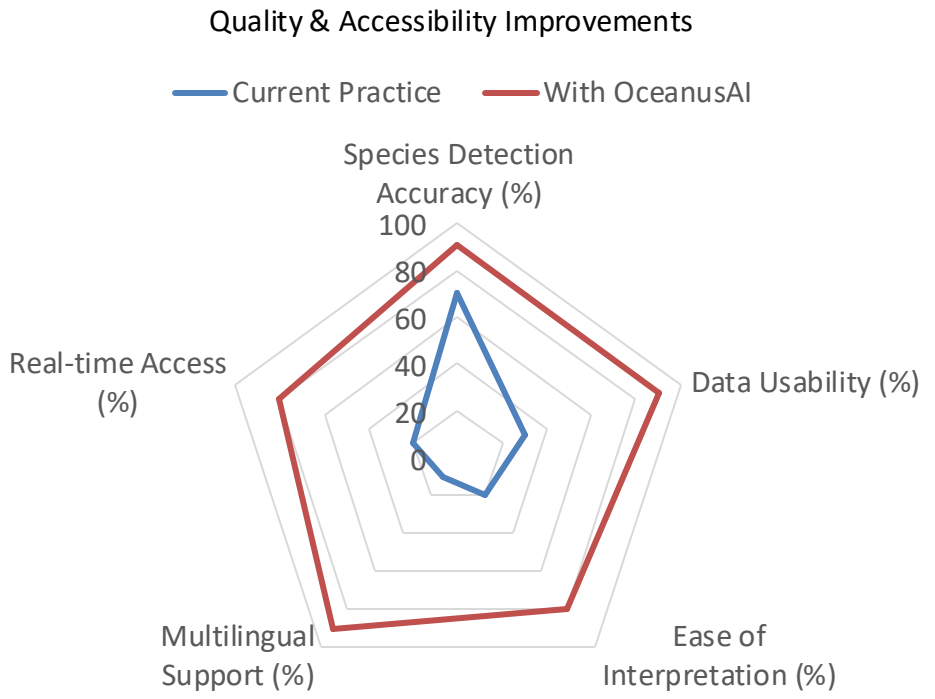
Analogy :

Just like Google Maps revolutionized how we navigate cities, OceanusAI will revolutionize how we navigate marine biodiversity and fisheries data — turning scattered information into one smart, interactive map.

Benefits :

- Faster scientific discovery & decision-making.
- Better fisheries & biodiversity management.
- Stronger blue economy & food security.
- Easy adoption with multilingual natural-language interface.





Breaking Bugs

RESEARCH AND REFERENCES

SMART INDIA HACKATHON 2025

- **Key References**

- **CMLRE (MoES)** — Annual Reports & Marine Living Resources Programme details (basis of the PS).
- **Darwin Core (TDWG)** — International biodiversity metadata standard.
- **CF Conventions (NetCDF/CF)** — Oceanographic data standard for interoperability.
- **QIIME2 / DADA2** — Leading bioinformatics pipelines for eDNA metabarcoding.
- **BLAST / VSEARCH** — Sequence matching & species identification tools.
- **Otolith Fourier Descriptors + CNN studies** — Proven approach for fish species & age classification.
- **Pinecone / Milvus (Vector DBs)** — Industry-standard solutions for scalable RAG.
- **LangChain / RAG frameworks** — Best practices for retrieval-augmented generation.
- **Mapbox / Deck.gl / CesiumJS** — Advanced visualization frameworks for geospatial analytics.
- **OGC SensorThings API / MQTT** — Standards for IoT data ingestion & real-time telemetry.
- **FAO Fisheries & UN Blue Economy Reports** — Global benchmarks for sustainable fisheries and marine ecosystems.

OceanusAI builds on globally accepted scientific standards and cutting-edge AI tools, ensuring credibility, scalability, and real-world impact.