BIG DATA TECHNOLOGIES
Mini-Project

Marine Transportation and Logistics Platform
Guided by
Prof. Akshita Chanchalani

NAME	PRN	ROLL NO
Sanika Mulik	1032222540	01
Nikita Kedari	1032222562	04
Juee Shinde	1032222572	07

PROBLEM STATEMENT

To develop an international Marine Transportation and Logistics platform that integrates real-time tracking, route optimization, and environmental monitoring, providing comprehensive tools for global stakeholders to enhance maritime operations.

PURPOSE

1. The maritime industry faces numerous challenges, including piracy risks, complex weather patterns, and the growing importance of emerging ports.
2. Existing logistics platforms often lack the ability to address these specific issues comprehensively, focusing instead on generalised solutions.
3. There is a need for a marine logistics platform that integrates real-time data and analytics to enhance efficiency, safety, and sustainability across the oceans.

SCOPE

1. **Real-Time Vessel Tracking:** Utilising AIS and GPS data for continuous tracking of vessels across global shipping routes.[1]
2. **Route Optimization:** Implement algorithms to optimise shipping routes based on weather conditions and ocean currents.[1][8][9]
3. **Environmental Monitoring:** Integrate weather services to provide forecasts and monitor sea conditions in real-time.
4. **Port Integration:** Collect data from major ports worldwide for efficient docking and logistics coordination.[9][4]
5. **Stakeholder Support:** Provide tools for shipping companies, port authorities, and environmental agencies to manage logistics efficiently.
6. **Security and Compliance:** Ensure data protection and adherence to global maritime regulations.

OBJECTIVES

1. To implement a robust system for real-time vessel tracking using AIS and GPS technologies.
2. To optimise shipping routes based on real-time weather data and ocean conditions.
3. To integrate weather and environmental data for monitoring sea conditions and ensuring sustainable maritime operations.
4. To provide tools for efficient port operations, including docking schedules and logistics coordination.

5. To provide fuel consumption and emission monitoring features tailored to various shipping routes.

LITERATURE SURVEY

Paper Title	Paper Description	Gap Analysis
A Comprehensive Review on Big Data-Based Potential Applications in Marine Shipping Management	This paper assesses challenges and opportunities of big data in marine shipping, focusing on applications like vessel performance optimization and predictive analysis.	Limited exploration of specific case studies demonstrating successful big data applications in real-world shipping scenarios.
A Data Fusion Approach to Predict Shipping Efficiency for Bulk Carriers	This study presents a data fusion methodology to enhance shipping efficiency for bulk carriers, focusing on predictive analytics.	Lacks empirical validation of the proposed methodologies in real-world shipping contexts.
Port Logistic Issues and Challenges in the Industry 4.0 Era for Emerging Economies	This paper addresses the challenges of port logistics in the context of Industry 4.0, including the role of big data.	Does not explore specific big data solutions that could alleviate identified port logistics challenges.
Impact of COVID-19 on the Indian Seaport Transportation and Maritime Supply Chain	The study analyses the effects of COVID-19 on maritime supply chains and transportation in India, highlighting disruptions and adaptations.	Lacks a focus on how big data could have mitigated the impacts of the pandemic on maritime operations.
Identification and Ranking of Key Factors Impacting Efficiency of Indian Shipping Logistics Sector	This research identifies critical factors affecting the efficiency of shipping logistics in India, providing a ranking based on their impact.	Does not incorporate big data analytics as a tool for improving the identified efficiency factors.

In detail Survey Google Sheets
https://docs.google.com/spreadsheets/d/1Gloaaha8dzALFddr69DqF9PJsuORwKgo19wfcmdjwEQ/edit?usp=sharing

INTRODUCTION

The global maritime industry plays a pivotal role in international trade, responsible for transporting the majority of the world's goods across oceans. However, this vital sector faces numerous challenges, including unpredictable weather patterns, varying sea conditions, and stringent environmental regulations. Current logistics platforms often struggle to address these complexities due to their generalised approach and fragmented data sources. Our project aims to overcome these limitations by developing a specialised platform that integrates advanced tracking technologies and real-time data analytics to enhance maritime logistics.[1][9] This solution is designed to support shipping companies, port authorities, and environmental agencies in managing global marine operations with greater precision and efficiency.[4]

ABSTRACT

The maritime industry is hindered by existing logistics platforms that fail to provide integrated solutions for real-time vessel tracking, route optimization, and environmental compliance. Our innovative platform addresses these issues by utilising AIS and GPS technologies alongside real-time data analytics. It offers features such as weather and environmental monitoring, fuel consumption tracking, and emissions management. By integrating data from major ports and global shipping routes, the platform delivers actionable insights that enhance operational efficiency and reduce environmental impact. This approach not only addresses the current gaps in maritime logistics but also contributes to a more sustainable and connected global shipping network.

System Architecture Overview

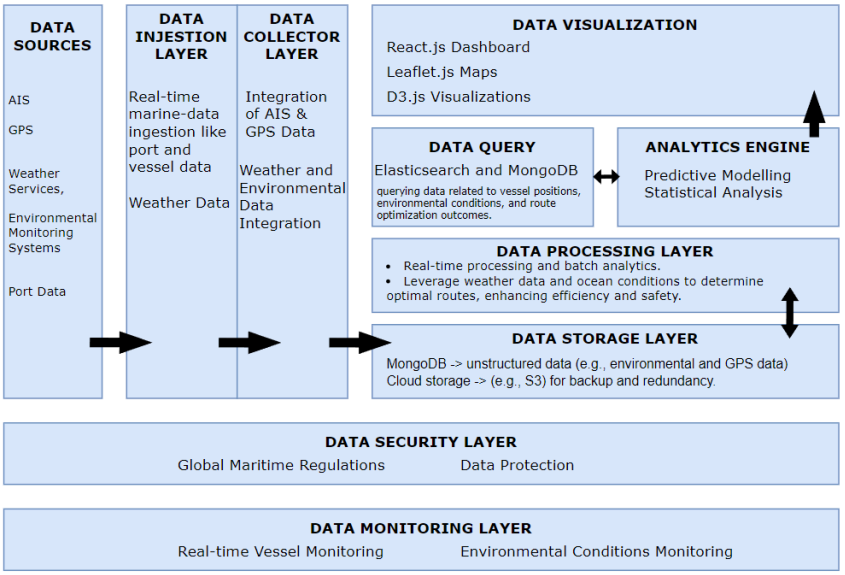

Fig:1 System Architecture and Big Data Cycle

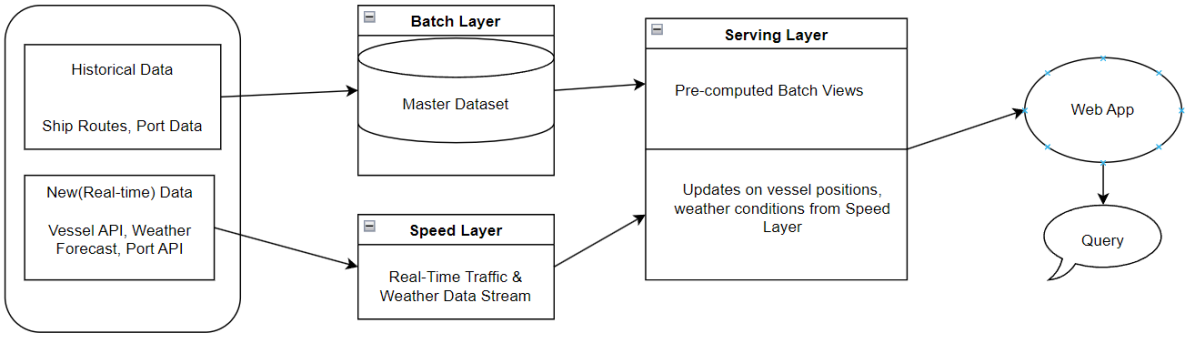

Fig:2 System Overview

---

1. **Data Sources**
    - AIS and GPS: Provide real-time and historical vessel positions for tracking.
    - Weather Services: Supply global forecasts and ocean conditions.
    - Environmental Monitoring: Collects data on ocean temperature and pollution levels.
    - Port Data: Includes locations, schedules, and traffic information.
2. **Data Ingestion and Collector Layers**
    - Apache Kafka: Manages real-time data ingestion from various maritime sources.

- Collector Layer: Integrates and preprocesses data for smooth analysis flow.

3. **Data Processing Layer**
    - Apache Kafka and Spark: Enable real-time and batch data processing.
    - Route Optimization: Algorithms use real-time weather and ocean conditions.
4. **Data Storage Layer**
    - MongoDB: Stores unstructured data like environmental and GPS information.
    - Elasticsearch: Allows efficient querying and analysis.
    - Cloud Storage: Provides backup and redundancy (e.g., S3).
5. **Data Query Layer**
    - Elasticsearch and MongoDB: Facilitate querying of vessel positions, environmental conditions, and optimization results.
6. **Analytics Engine**
    - Predictive Modelling: Optimizes routes, predicts fuel consumption, and assesses environmental impacts.
    - Statistical Analytics: Provides insights into operational efficiency.
7. **Data Visualization Layer**
    - React.js and D3.js: Create dynamic interfaces for data visualisation.
    - Leaflet.js: Offers interactive maps for real-time maritime monitoring.
8. **Security and Compliance Layer**
    - Standards and Regulations: Ensures adherence to maritime security and international compliance.
9. **Monitoring Layer**
    - Real-Time Monitoring: Continuously observes vessel operations and compliance checks.

DATASET INFORMATION

1. **AIS Data**: Provides real-time and historical vessel positions.
2. **GPS Data**: Offers precise real-time tracking of vessel locations.
3. **Weather Data**: Includes forecasts, ocean currents, and wave heights from sources like NOAA and CMEMS.
4. **Fuel Consumption Data**: Collected from vessel sensors and shipping companies.
5. **Port Data**: Information on port locations, docking schedules, and traffic.
6. **Environmental Data**: Covers oceanic conditions such as temperature and pollution levels.

SCHEMA DESIGN

Vessels Collection

Feature	Description	Data Type
_id	Unique identifier	String
name	Name of the vessel	String
mmsi	Maritime Mobile Service Identity number	String
imo	International Maritime Organization number	String
vessel_type	Type of vessel	String
owner	Name of the owning company	String
dimensions.length	Length of the vessel	Number
dimensions.width	Width of the vessel	Number
capacity	Cargo capacity in tons	Number

Positions Collection

Feature	Description	Data Type
_id	Unique identifier for each position entry	String
vessel_id	Reference to the vessel	String
timestamp	Date and time of the position record	ISODate (Excel equivalent: Date and Time)
latitude	Current latitude of the vessel	Number
longitude	Current longitude of the vessel	Number
sog	Speed over ground in knots	Number
cog	Course over ground in degrees	Number
destination	Destination port	String
eta	Estimated time of arrival	ISODate (Excel equivalent: Date and Time)

WeatherConditions Collection

Feature	Description	Data Type
_id	Unique identifier for each weather entry	String
timestamp	Date and time of the weather record	ISODate (Excel equivalent: Date and Time)
latitude	Latitude where the weather data is collected	Number
longitude	Longitude where the weather data is collected	Number
temperature	Temperature in degrees Celsius	Number
wind_speed	Wind speed in knots	Number
wind_direction	Wind direction in degrees	Number
wave_height	Wave height in meters	Number
visibility	Visibility in nautical miles	Number

Ports Collection

Feature	Description	Data Type
_id	Unique identifier for each port	String
port_name	Name of the port	String
country	Country where the port is located	String
location.latitude	Latitude of the port	Number
location.longitude	Longitude of the port	Number
capacity	Capacity of the port in terms of vessels it can handle	Number
facilities	Array of facilities available at the port	Array of Strings (Excel equivalent: Multiple columns)

YOUR NOVELTY(CONTRIBUTION)

Our platform revolutionises global marine transportation and logistics with:

- **Unified Data Integration**: Consolidates AIS, GPS, weather, and environmental data for better decision-making and predictive analytics.
- **Real-Time Route Optimization**: Uses machine learning to adjust routes based on current conditions, reducing travel time and fuel costs while enhancing safety.
- **Environmental Monitoring**: Provides real-time impact analysis and ensures compliance with regulations, promoting sustainability and improving market reputation.
- **Advanced Visualization**: Utilises React.js, D3.js, and Leaflet.js for intuitive, interactive data displays, enhancing transparency and management.
- **Scalable Architecture**: Built with Apache Kafka and Spark to handle large data volumes and adapt to industry needs, ensuring long-term effectiveness.

Market Impact: Our platform sets a new standard in maritime logistics, improving efficiency, safety, and sustainability, leading to cost reductions, regulatory compliance, and widespread industry adoption.

REFERENCES:

1. Port logistic issues and challenges in the Industry 4.0 era for emerging economies: an India perspective, Bishal Dey Sarkar, Ravi Shankar, Arpan Kumar Kar, Benchmarking: an International Journal, February 2022: https://www.emerald.com/insight/content/doi/10.1108/BIJ-08-2021-0499/full/html
2. A data fusion approach to predict shipping efficiency for bulk carriers, Dennis Sugrue, Peter Adriaens, Elsevier: Transport Research Part E: Logistics and Transportation Review, Vol 149, May 2021: https://doi.org/10.1016/j.tre.2021.102326

DATASETS

1. Vessel Tracking Data

- https://www.marinetraffic.com/en/ais-api-services

2. Weather and Ocean Conditions Data:

- https://data.marine.copernicus.eu/products?option=com_csw&task=results

3. Port Data

- https://msi.nga.mil/Publications/WPI