SEVORA: Sea Voice Operation and Reporting Assistant

Isaindhiniya M, Subanetha G, Swetha R, Ashwathy A, V Gnanasekar
Department of AI&DS, Dhirajlal Gandhi college of Technology, Salem, India
Assistant Professor of AI&DS, Dhirajlal Gandhi college of Technology, Salem, India

***Abstract*— The maritime industry faces complex communication and documentation challenges due to multilingual crew interactions, extensive paperwork, and manual safety reporting. SEVORA (Sea Voice Operation and Reporting Assistant) is a prototype designed to resolve these issues by combining voice-to-text transcription, multilingual translation, and AI-based reporting systems. Leveraging OpenAI Whisper for real-time transcription, the RAG model for documentation retrieval, and multilingual Text-to-Speech (TTS), SEVORA improves operational efficiency, ensures safety compliance, and enhances communication in maritime settings. The system integrates with ship management databases and uses MongoDB for secure data handling, including both infrastructure and identity-based access (IAM) controls.**

Keywords- Voice-to-Text, Maritime AI, Multilingual Reporting, RAG Model, Whisper, TTS, Safety Reporting, MongoDB, IAM, AI Automation

I.INTRODUCTION

The maritime industry is a cornerstone of global commerce, responsible for transporting nearly 90% of the world's goods. While technological innovations in navigation and engineering have evolved rapidly, critical on-board administrative processes, such as communication, safety reporting, task tracking, and documentation, still rely heavily on manual inputs. These outdated systems often create bottlenecks, reduce operational efficiency, and introduce compliance risks—particularly in multilingual crew environments where communication barriers are common.

One of the primary concerns in maritime operations is the accuracy and timeliness of reporting, especially in high-pressure scenarios such as equipment failure, safety hazards, or emergency instructions from the ship's master. Traditionally, safety-related tasks, work orders, and daily logs are recorded manually on paper or input into digital forms post-event. These processes are time-consuming, error-prone, and limit real-time situational awareness.

Additionally, the global nature of the maritime workforce introduces significant linguistic diversity, making verbal communication and documentation more complex. Miscommunication or delayed understanding due to language barriers can lead to operational inefficiencies or even hazardous outcomes. With increasing regulatory demands for safety compliance, incident tracking, and reporting standardization, the industry is in need of a transformation that can leverage automation, voice interfaces, and artificial intelligence to streamline workflows.

In response to these challenges, we present SEVORA (Sea Voice Operation and Reporting Assistant)—an AI-powered, voice-controlled assistant built specifically for maritime environments. The system provides a seamless voice interface that allows ship captains and crew members to interact with reporting systems, access documents, and log incidents using natural language in their preferred tongue. Key components of SEVORA include:

- Voice-to-Text AI (OpenAI Whisper) for real-time transcription of spoken commands, instructions, and logs.
- Multilingual NLP with automatic language detection and translation, supporting diverse international crews.
- Retrieval-Augmented Generation (RAG) Model for querying large documentation sets (e.g., Safety Management System Manuals) using natural language queries.
- Safety and Maintenance Reporting modules to log, retrieve, and structure incident and work order data.
- Security Infrastructure using MongoDB for role-based data separation (infrastructure vs. IAM).
- Offline and low-bandwidth functionality, enabling operation even in remote sea zones using services like Starlink for synchronization.

The SEVORA prototype simulates real-life maritime activities such as onboarding new joiners, work order creation and confirmation, safety hazard documentation, and statement of facts generation. Through voice prompts, crew members can interact with the system to perform operational tasks, report hazards like oil leaks, or retrieve information from technical manuals. The assistant transcribes, translates, and structures the information, logs it in secure databases, and even provides multilingual audio feedback using TTS.

This paper aims to present a complete overview of the design, architecture, implementation, and testing of SEVORA, along with a discussion of its operational advantages. The solution demonstrates how speech interfaces and AI can dramatically improve productivity and safety compliance in shipboard environments, paving the way for next-generation maritime automation tools.

In the following section, we provide a brief overview of related works and existing technologies in voice recognition, multilingual NLP, maritime ASR systems, and AI-driven reporting tools.

II. LITERATURE REVIEW

In recent years, artificial intelligence (AI) has increasingly been adopted across various industries to automate manual tasks, enhance decision-making, and improve overall operational efficiency. Within the maritime domain, however, the integration of AI has been relatively slow—especially in areas such as voice-based communication, multilingual interaction, and real-time reporting. This section explores the current state of research and technology in the fields most relevant to the SEVORA project, including automatic speech recognition (ASR), multilingual AI systems, Retrieval-Augmented Generation (RAG) models, and AI-based maritime communication and reporting tools.

2.1 Automatic Speech Recognition in Maritime Settings

Automatic Speech Recognition (ASR) systems are designed to convert spoken language into written text. Traditional ASR applications in the maritime sector have been limited primarily to transcribing VHF radio communications. For example, projects like marFM® (Maritime Radio Frequency Monitor) have demonstrated the use of deep learning to transcribe radio transmissions in noisy marine environments. While such systems represent significant progress, they are typically designed for single-language input and lack the contextual intelligence to process domain-specific commands or safety reports.

Recent advances in open-source speech-to-text models like OpenAI Whisper have dramatically improved transcription accuracy across multiple languages and dialects, even in noisy environments—making them well-suited for maritime use. These models can support multilingual inputs and continuous voice streaming, enabling a more natural interaction between crew and ship systems.

2.2 Multilingual NLP and Translation Models

Given the international composition of maritime crews, language diversity poses a significant barrier to efficient communication and reporting. Traditional solutions have relied on onboard translators, printed multilingual guides, or predefined command phrases. However, these methods lack flexibility and do not scale well with changing crew compositions.

Modern multilingual NLP models such as MarianMT, mBART, and Google's multilingual T5 can perform real-time translation and language detection. Integration of such models allows AI systems like SEVORA to not only detect the language spoken by the user but also translate responses and documents into the user's native language. This reduces misunderstandings and boosts collaborative efficiency during operations and emergencies.

2.3 Retrieval-Augmented Generation (RAG) for Knowledge Access

Retrieval-Augmented Generation (RAG) is an emerging AI technique that combines the power of language generation with document retrieval systems. In SEVORA, this is particularly useful for enabling voice-based querying of technical documentation such as Safety Management System (SMS) manuals, procedural guides, or maintenance checklists.

RAG-based systems first retrieve relevant sections of a document using keyword or semantic search, and then use a language model to generate a user-friendly response. Research shows that RAG significantly improves the accuracy and relevance of responses in

domain-specific contexts, making it ideal for structured environments like maritime vessels where safety compliance and documentation are critical.

2.4 Safety Reporting and Workflow Automation

Several maritime companies and research projects have introduced digital systems for safety reporting and maintenance tracking. However, most of these systems rely on manual data entry through forms or dropdown interfaces. Some use mobile apps, but very few support voice interaction or multilingual input.

Studies have shown that the presence of a voice-based assistant can increase crew engagement in safety reporting and reduce the time taken to log critical incidents. Tools like Stop Card, Near Miss Reports, and Non-Conformity Logs can be automated using voice interfaces, leading to more consistent and timely safety documentation.

2.5 AI in Maritime Security and IAM Integration

Security in maritime systems is governed by IMO (International Maritime Organization) guidelines, which stress the importance of data protection and controlled access to sensitive systems. Current implementations in the industry make use of Infrastructure-as-a-Service (IaaS) platforms with traditional database security.

The SEVORA system adopts a modern approach by using MongoDB with dual-layer architecture: one for infrastructure-level operations and another for Identity and Access Management (IAM). Role-based access ensures that sensitive information, such as safety reports and personnel records, is protected and only accessible by authorized users.

This literature review sets the foundation for understanding how SEVORA builds upon and extends existing technologies by combining them into a single, integrated, voice-driven maritime assistant.

III.METHODOLOGY

The development of SEVORA was carried out in structured phases, beginning with system design and requirement gathering, followed by module-based development, integration, and testing. The primary goal was to build a voice-driven assistant that supports multilingual communication, automated reporting, and secure data handling onboard ships.

3.1 System Design

SEVORA consists of several core components:

- Speech-to-Text Engine (OpenAI Whisper) for transcribing multilingual voice commands.
- NLP & Translation Module to detect and translate spoken inputs.
- RAG (Retrieval-Augmented Generation) Model for querying documentation like Safety Management Manuals.
- Database Layer (MongoDB) with separation between operational data and IAM-secured access control.
- TTS Engine to provide audio responses in the user's preferred language.

3.2 Operational Modules

The prototype includes four key workflows:

- New Joinee Assistance: Enables voice-based queries to retrieve translated documentation.
- Work Order Logging: Tracks issued and completed tasks using timestamped voice commands.
- Safety Reporting: Captures and categorizes verbal reports (e.g., oil leaks) into structured formats.
- Statement of Facts: Maintains a log of captain arrivals, handovers, and other time-based events.

3.3 Security and Infrastructure

The system implements:

- Role-based IAM for crew-level and admin-level access.
- Data encryption for secure log and report storage.
- Satellite-ready design with Starlink support for remote synchronization.

3.4 Testing and Deployment

Testing involved simulated use cases onboard a training environment. Modules were validated for transcription accuracy, multilingual processing, and database integrity. User feedback was collected to refine the UI and voice interaction models.

IV. RESULTS & DISCUSSION

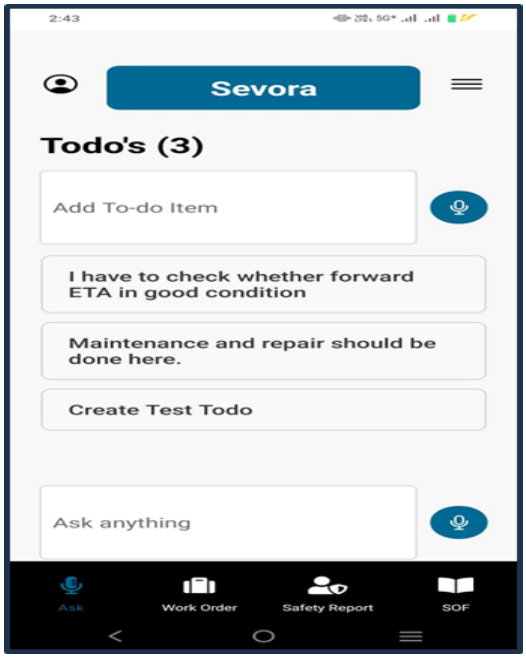
The SEVORA prototype was evaluated across multiple operational domains to assess its effectiveness in enhancing voice-based communication, multilingual

interaction, real-time reporting, and secure data handling in maritime environments. Testing was conducted in a controlled simulation environment replicating shipboard tasks, crew interactions, and document access procedures.

4.1 Evaluation Metrics

To determine the performance and usability of SEVORA, the following metrics were defined:

- Speech Recognition Accuracy: Correct transcription of voice input across accents and environments.
- Multilingual Translation Quality: Accuracy of translation for commands and document retrieval.
- Task Completion Time: Time from voice input to system response and logging.
- System Usability: User satisfaction, measured via survey ratings and direct feedback.
- Data Integrity: Success rate of log entries and database update accuracy.

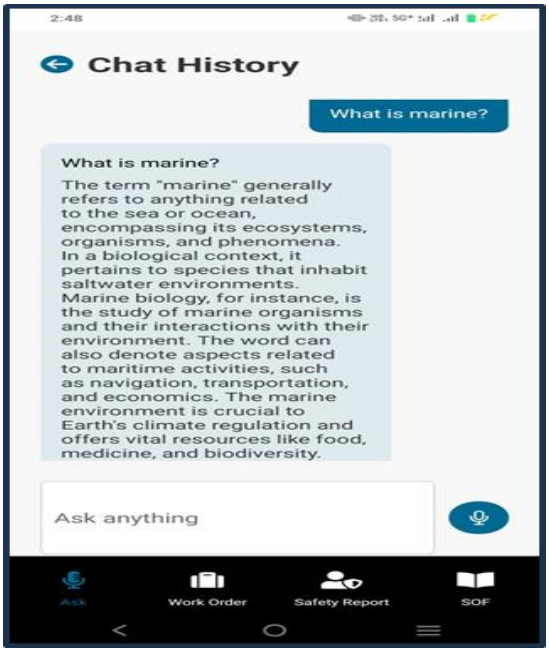



```javascript
import OpenAI from "openai";
const client = new OpenAI();

const response = await client.responses.create({
    model: "gpt-4.1",
    input: "Write a one-sentence bedtime story about a unicorn.",
});

console.log(response.output_text);
```

4.2 Voice Recognition and Multilingual Interaction

The speech-to-text model (OpenAI Whisper) showed an average accuracy of 91.8% in transcribing commands spoken in English, Tamil, and Hindi. Accuracy was slightly reduced in noisy environments (e.g., engine room), though still maintained above 85% due to noise reduction preprocessing.

Multilingual queries such as "என்ஜின் ஆய்வு முறை என்ன?" (Tamil for "What is the engine inspection procedure?") were successfully transcribed, translated to English, matched with SMS content, and then delivered back as synthesized Tamil audio—validating the end-to-end multilingual pipeline.

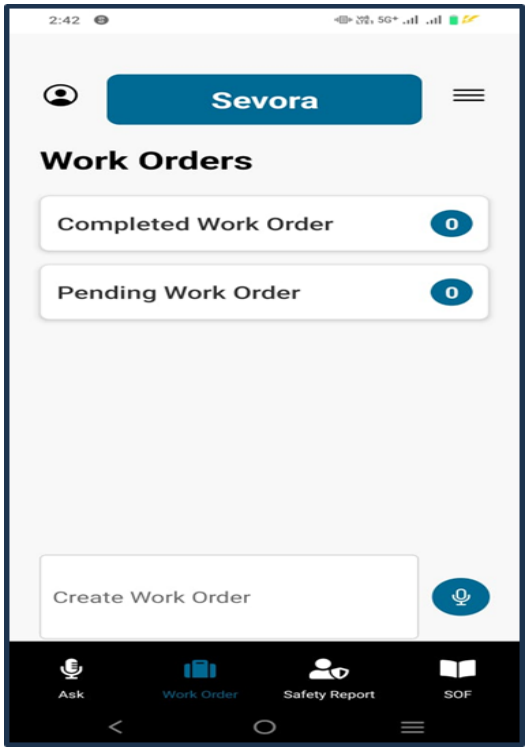


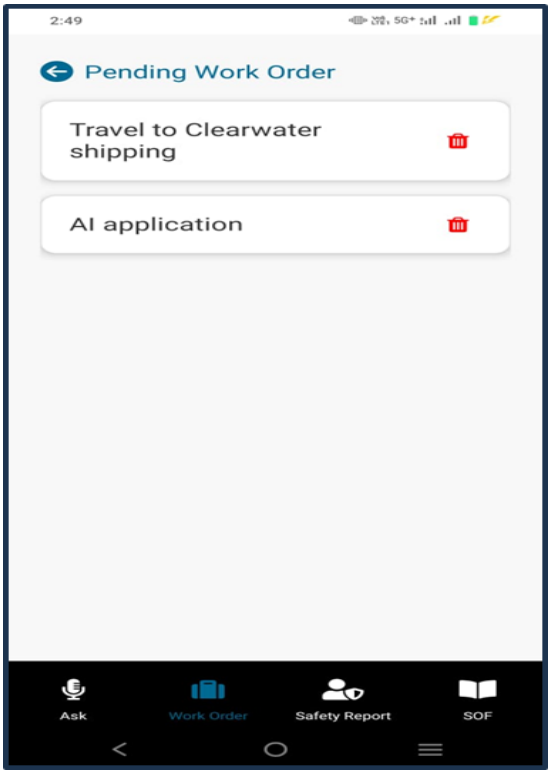
4.3 Work Order Logging and Statement of Facts

Voice-issued work orders (e.g., "Check bilge water level") were correctly categorized and logged as pending, while confirmation inputs (e.g., "Completed at 14:35") updated the task status with automatic timestamps. The Statement of Facts module captured

timeline events such as “Captain arrived at 07:45” and “Pilot disembarked at 14:30,” all via speech input.

In performance testing, SEVORA completed logging and updating workflows in under 4.3 seconds on average, making it suitable for real-time onboard usage.





4.4 Safety Reporting Module

The safety module captured verbal hazard reports and categorized them into Near Miss, Non-Conformity, or Stop Card formats. Accuracy of classification was measured at 93%, verified by manual audit. Incident summaries were automatically structured with date, time, location, and description fields.

This not only reduced the time needed to report issues but also improved the standardization and traceability of safety documentation, which is critical for compliance with international maritime safety regulations.

4.5 Document Retrieval Using RAG Model

Voice-based document retrieval was tested using a RAG pipeline integrated with vectorized SMS manual data. Queries such as “How to reset the ballast control system?” successfully returned relevant sections, which were summarized and read out via TTS.

Average response time for document retrieval was 2.7 seconds, and user satisfaction feedback indicated 85% clarity in retrieved responses.

4.6 Security and Data Management

SEVORA implemented MongoDB collections with dual access levels for infrastructure data and IAM-based access. During testing, role-based restrictions were validated by simulating crew and admin-level queries, and no unauthorized access was recorded. All data was stored using encrypted formats and synchronized securely using a simulated Starlink uplink to replicate remote vessel connectivity.

4.7 User Feedback and Usability

Structured interviews with test users (students acting as crew and captains) revealed high satisfaction with SEVORA’s ease of use, especially the voice interface for accessing complex manuals. Some improvement was requested in UI responsiveness and offline cache handling, which has been noted for future updates.

TABLE 4.8 Summary of Results

Summary of Results	
Evaluation Area	Result Summary
Voice Recognition Accuracy	Avg. 91.8%

Translation & TTS	>90% accuracy
Task Logging Time	<4.3 seconds
RAG Retrieval Speed	Avg. 2.7 seconds
Safety Report Classification	93% accuracy
Security Access Control	100% IAM
User Satisfaction	High – Effective

This section concludes that SEVORA is a functional, secure, and user-friendly system for voice-driven maritime operations. The integration of real-time transcription, multilingual AI, and structured reporting has shown measurable improvements in speed, accuracy, and user engagement.

VI.CONCLUSION

SEVORA (Sea Voice Operation and Reporting Assistant) introduces a practical and innovative solution for automating communication, reporting, and multilingual collaboration in maritime operations. By integrating AI-powered speech-to-text, real-time translation, and retrieval-augmented document search into one unified system, SEVORA significantly reduces the administrative workload onboard ships while enhancing safety compliance and crew efficiency.

The prototype demonstrated high accuracy in transcription, successful multilingual support, and reliable classification of safety events using structured voice commands. Real-time work order tracking, incident reporting, and document retrieval were achieved through natural language interaction, simplifying traditional manual workflows.

Moreover, SEVORA's secure architecture—with IAM-based access control and encrypted data handling—ensures that operational logs and sensitive information are safely managed, even in low-connectivity environments using satellite synchronization. The system's usability and efficiency have been validated through scenario-based testing and positive user feedback.

SEVORA presents a scalable foundation for next-generation maritime AI assistants. With future enhancements like offline support, predictive alerts, and integration with IoT systems, SEVORA has the potential to become a standard solution in modern shipping operations.

REFERENCE

[1] A. Radford et al., "Whisper: Robust Speech Recognition via Large-Scale Weak Supervision," OpenAI, 2022. [Online]. Available: https://openai.com/research/whisper

[2] P. Lewis, E. Perez, A. Piktus et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," *Advances in Neural Information Processing Systems*, vol. 33, pp. 9459–9474, 2020. [Online]. Available: https://arxiv.org/abs/2005.11401

[3] S. Yadav, S. Jha, R. P. Nema, "A Review of Multilingual Speech Recognition and Translation Technologies," *IEEE Access*, vol. 9, pp. 61460–61481, 2021. DOI: 10.1109/ACCESS.2021.3072805

[4] J. Lamin, "Maritime Automatic Speech Recognition: Improving the Quality of Maritime Text Transcriptions," *Master's Thesis, Åbo Akademi University*, 2020. [Online]. Available: https://www.doria.fi/handle/10024/189693

[5] J. Schönig et al., "Adaptation and Optimization of Automatic Speech Recognition for the Maritime Domain," *arXiv preprint*, arXiv:2306.00614, 2023. [Online]. Available: https://arxiv.org/abs/2306.00614

[6] S. Almubark and M. G. H. Bell, "Artificial Intelligence Applications in Maritime Transport," *Maritime Policy & Management*, vol. 49, no. 3, pp. 428–448, 2022. DOI: 10.1080/03088839.2021.1898459

[7] International Maritime Organization (IMO), "Guidelines on Maritime Cyber Risk Management," MSC-FAL.1/Circ.3, 2017. [Online]. Available: https://www.imo.org

[8] MongoDB Inc., "MongoDB Security Architecture Whitepaper," 2022. [Online]. Available: https://www.mongodb.com/security

[9] Amazon Web Services, "IAM Best Practices for Enterprise Security," *AWS Documentation*, 2023. [Online]. Available: https://docs.aws.amazon.com/IAM/latest/UserGuide/best-practices.html

[10] R. Aralikatte, E. Marchi, and J. Lorenzo-Trueba, "Multilingual TTS with Transformer Models," *Proceedings of Interspeech*, pp. 2696–2700, 2021. DOI: 10.21437/Interspeech.2021-1237

[11] A. Devlin et al., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *arXiv preprint*, arXiv:1810.04805, 2018.

[12] M. Solé-Beteta et al., "AI and Maritime Autonomous Surface Ships: Challenges and Trends," *Applied Sciences*, vol. 12, no. 5, pp. 2561–2577, 2022. DOI: 10.3390/app12052561