Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 12, No. 3, 2024

Research on Ship Data Analysis Platform Development and Its Application in Teaching

Juan Zhang, Enze Wu, Taizhi Lv

College of Information Engineering, Jiangsu Maritime Institute, Nanjing, 211170, China

Abstract: This paper delves into the significant importance of developing a ship data analysis platform, which aims to enhance the operational efficiency and safety of the shipping industry by integrating big data processing technologies. The development of the platform centers on crucial aspects such as data collection, preprocessing, analysis, and visualization, enabling real-time monitoring and intelligent prediction of ship operational status. This platform not only assists in optimizing route planning and reducing energy consumption costs but also provides robust support for navigational safety. In project-based teaching, this platform serves as a bridge between theory and practice, cultivating students' abilities in big data processing and analysis through case studies, practical exercises, and other forms, thereby promoting the cultivation of high-quality talents and technological innovation in the field of shipping.

Keywords: Ship data analysis, big data technology, multi-level data storage architecture, data acquisition, project-based teaching, ideological and political education in courses.

1. Introduction

In the era of big data, the importance of ship data analysis has become increasingly prominent, serving as a crucial force driving the transformation and upgrading of the shipping industry. With the rapid development of information technology, the volume of data generated during ship operations has exploded, covering various dimensions such as navigation status, cargo information, energy consumption, and equipment maintenance, forming the cornerstone of shipping big data [1]. The application of ship data analysis primarily lies in enhancing operational efficiency. Through deep mining of navigation data, it is possible to optimize route planning, reduce unnecessary sailing time and energy consumption, thereby lowering costs [2]. At the same time, meticulous management of cargo loading and unloading processes can significantly improve port operation efficiency, reducing waiting times and waste of resources. Secondly, data analysis plays an irreplaceable role in ensuring navigation safety. By analyzing historical accident data, potential safety hazards can be identified, allowing for timely preventive measures to be taken. Furthermore, real-time monitoring of ship status enables the prompt detection and handling of abnormalities, ensuring the safety and control of the navigation process. Additionally, ship data analysis contributes to enhancing the scientific nature of decision-making. Prediction models based on big data can accurately forecast market trends and cargo demand, providing strong support for the strategic planning of shipping companies. Meanwhile, detailed analysis of operational costs can also help companies identify breakthrough points for cost reduction and efficiency improvement, thereby enhancing market competitiveness. Ship data analysis holds a pivotal position in the era of big data. It not only improves the operational efficiency and safety of the shipping industry but also provides a scientific basis for corporate strategic decisions, driving the transformation and upgrading of the shipping industry towards intelligence and greenness. Therefore, strengthening the construction of ship data analysis capabilities is both an inevitable choice for the shipping industry to address current challenges and an important way to seize future development opportunities. As big data technology continues to advance, ship data analysis will play a more significant role in the shipping industry, leading it towards a new stage of development [3-4].

Integrating ship data analysis platforms into professional teaching in the field of big data technology not only enhances students' professional skills and industry awareness but also stimulates their innovative potential, providing strong support for cultivating high-quality big data talents that meet future market demands. Firstly, through practical operation of ship data analysis platforms, students can closely integrate the theoretical knowledge learned in the classroom with practical applications, gaining a deep understanding of the entire process of big data processing, analysis, and visualization. This "learning for application" teaching mode greatly enhances students' interest and motivation in learning, facilitating the transformation of knowledge into skills. Secondly, ship data analysis platforms involve multiple links such as data collection, cleaning, storage, analysis, and visualization, requiring students to master a series of big data technology tools and methods. This not only enhances students' data processing capabilities but also familiarizes them with industry-specific analysis needs and solutions, laying a solid foundation for their future work in related fields. Lastly, by analyzing ship operation data, students can gain a deep understanding of the operation mode, market trends, and challenges of the shipping industry, cultivating sensitivity and insight into specific industries. This interdisciplinary perspective is crucial for cultivating compound big data talents and helps them better serve various industries in their future careers.

2. Platform Functionality and Architecture

2.1. Functional modules

As shown in Figure 1, the ship data analysis platform comprises ten systems, including the Ship Registration Data Analysis System, Ship AIS (Automatic Identification System)

Data Analysis System, Ship Water Pollutant Emission Data Analysis System, Ship Shore Power Data Analysis System, Ship Entry and Exit Report Data Analysis System, Ship Manning Data Analysis System, Ship Electronic Fence Data Analysis System, Ship Risk Data Analysis System, Ship Hazardous Chemical Management Data Analysis System, and Ship Epidemic Prevention Data Analysis System. The platform processes data through a big data processing platform and achieves data visualization using front-end and back-end technologies.

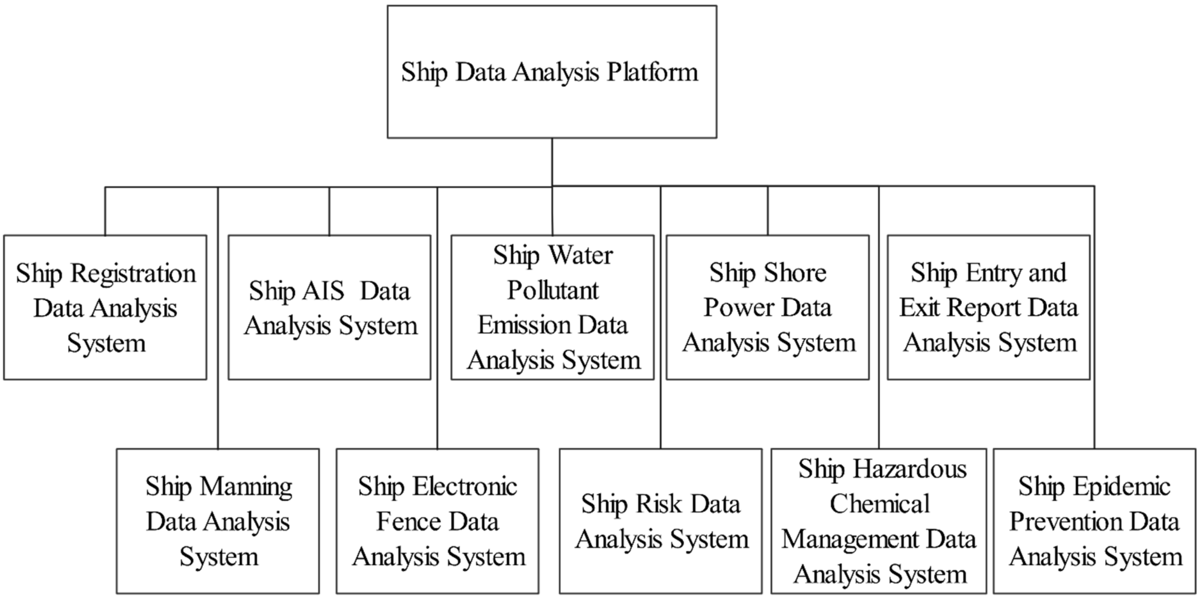

Figure 1. The Functional Module Diagram of the Platform

Ship registration data is one of the most crucial pieces of information in maritime management and forms an integral part of shipping big data. Its accuracy and reliability directly impact the precision and efficiency of maritime operations. The ship registration data analysis function primarily includes querying and retrieving basic ship information, statistically analyzing ship dynamic data, and recording and analyzing ship violations. The intelligent ship monitoring system utilizes a highly efficient and stable Kafka message queue to transmit AIS data, combined with various data processing technologies such as Spark and Flink, as well as a multi-level storage architecture, to construct an efficient Lambda-based data processing framework. This framework realizes functions such as risk ship alerts, ship traffic statistics, and ship accident rescue, thereby enhancing maritime safety and shipping efficiency. The ship water pollution data analysis module facilitates the identification of factors influencing pollutant quantities through statistics on the number of manifests and pollutants. This enables the control of pollutant quantities and the setting of risk threshold values. The ship shore power data analysis module provides statistics on the electricity consumption of shore power equipment when ships are supplied with power from shore-based sources during their stay at ports. This analysis aids in the research and study of shore power technical data, bringing significant social and economic benefits to port cities. It not only achieves energy conservation and emission reduction but also provides support and guidance for the layout of shore power for ships in ports. The ship entry and exit report module analyzes the movement of ships between different provinces and examines historical cargo data to create initial ship profiles. This analysis helps to study the degree of dependence on cargo by different provinces and their positioning within the industrial chain. Insufficient ship manning poses a significant safety hazard and may lead to serious maritime accidents. The implementation of this system enables regulatory authorities to have real-time knowledge of ship manning situations, effectively enhancing the intelligence level of ship manning management. In water transportation, the core of virtual fence technology lies in real-time tracking of ship positions and effective monitoring of their entry and exit behaviors, with the AIS system playing a crucial role [5]. Based on the AIS application background, the platform has constructed an electronic data fence analysis system for ships under a virtual fence, realizing the map-based display of virtual fences and ship positions, and improving the level of information visualization. Ship transportation risk analysis plays a vital role in ensuring navigation safety, enhancing transport efficiency, reducing environmental impact, and ensuring compliance. The ship risk data analysis system deeply analyzes various factors in the ship transportation process, not only revealing potential problems and areas for improvement but also significantly enhancing the effectiveness and level of maritime regulation. With the development of water transportation, the volume of hazardous chemical shipments and the risk of accidents have increased, making it one of the key management focuses for maritime regulatory authorities. The platform has designed and implemented an AIS-based real-time analysis system for risk data of hazardous chemical ships, improving the intelligence level of hazardous chemical ship management.

2.2. Data processing architecture

As shown in Figure 2, the platform adopts Lambda as its data processing architecture. Lambda Architecture is a framework for processing big data, which combines batch processing and stream processing to handle data, supporting fast real-time queries and complex analytical operations [6]. The platform is divided into five layers: the source data layer, the data acquisition layer, the data storage layer, the data computation layer, and the data service layer. The source data layer consists of business data warehouses, Internet of Ships (IoS) data collection tools, and so on. The data acquisition layer pushes data from the source data layer to the data storage layer through data acquisition tools such as Kafka message queues or Sqoop. The data storage layer stores massive amounts of ship data through multi-level distributed storage. The data computation layer is divided into real-time computation and offline computation. Offline computation is responsible for storing and processing offline historical data sets, generating specific query views through pre-computation to ensure data accuracy and completeness. The real-time layer processes incrementally generated real-time data streams, producing real-time data views with low latency, as it deals with incremental data. The service layer combines the result sets generated by the batch processing layer and the acceleration layer, and responds to user query requests.

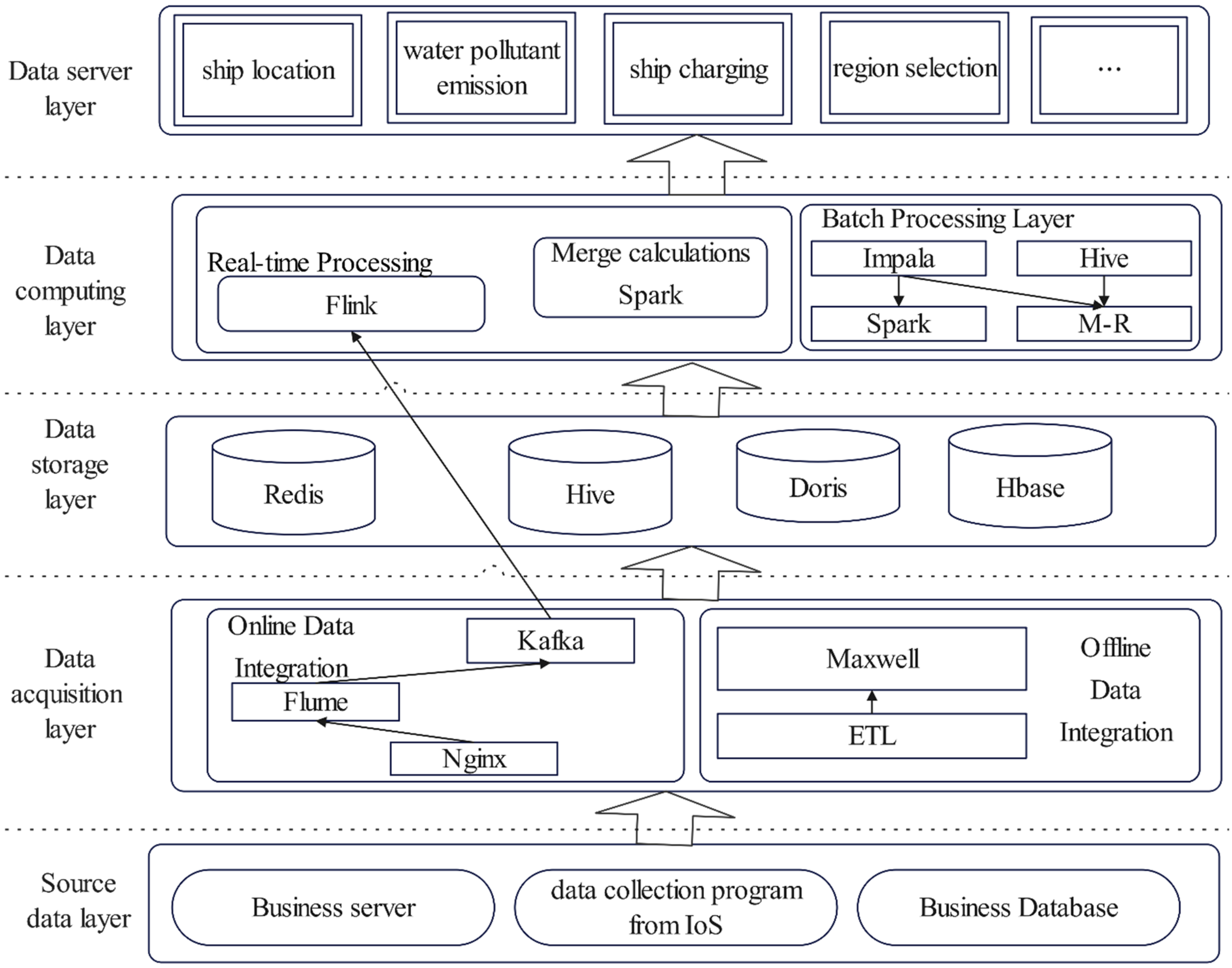

Figure 2. The Functional Module Diagram of the Platform

3. Platform Implementation

3.1. Data acquisition

The data originates from various sources. Specifically, the ship entry and exit data comes from a secondary lower-level database, which is based on Oracle. Ship AIS (Automatic Identification System) data is collected in real-time from various base stations. Water pollution data from ships is sourced from the ship water pollution business system, which utilizes a MySQL database. Basic ship information data is derived from the integration of multiple databases, including the ship registration system, ship radio management system, and static AIS data of ships. Ship shore power data comes from the ship shore power system, also based on a MySQL database. As shown in Figure 3, to collect data from different sources into Doris, the platform uses incremental scripts to push the data into the Kafka message queue. Subsequently, Kafka Connect is utilized to import the data into the Doris data warehouse.

For the data acquired from IoS, the platform employs a message queue to facilitate real-time data acquisition and transmission. It utilizes the Producer API to receive data, which is then dispatched to a Flink cluster or stored in the data storage layer. As for the log data from the business database, a Flume client program is used to collect it in real-time from the Nginx server. This data is subsequently transmitted to the data storage layer via the Kafka message queue.

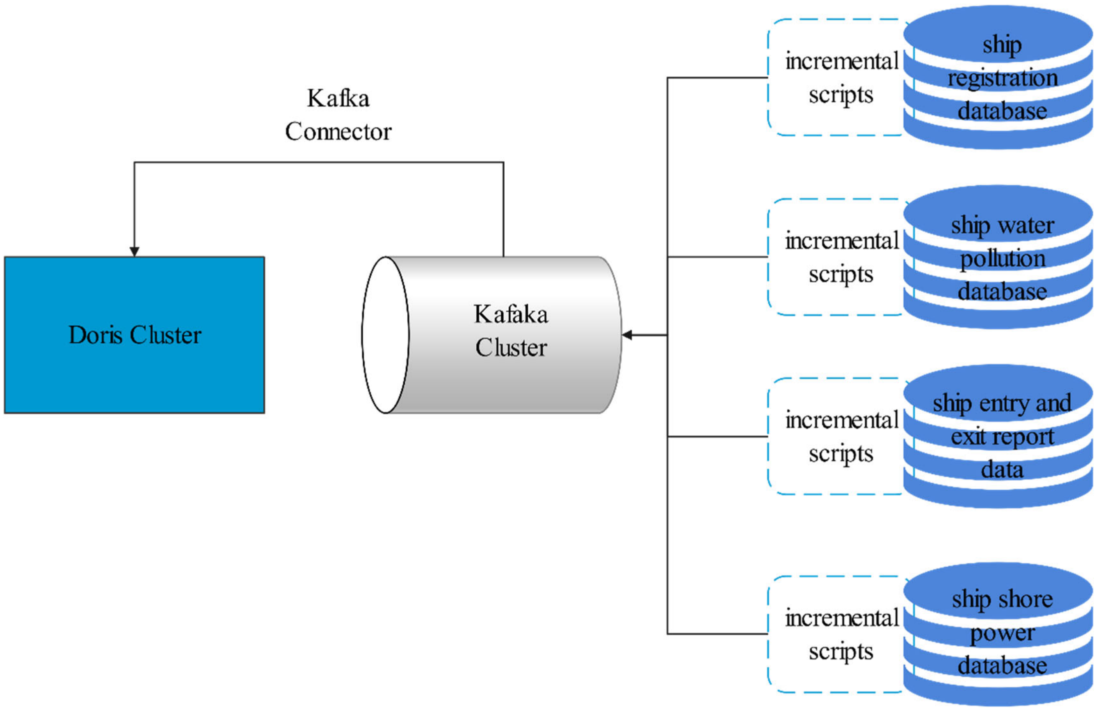

Figure 3. The data collection process from the business database

3.2. Data computing

In the early stages, the system adopted a traditional relational data warehouse architecture, but this failed to meet the growing demands of data processing. It was unable to handle large volumes of data queries, particularly those related to ship trajectory information. To enhance data processing reliability and real-time performance, the platform transitioned to a Lambda-based data processing architecture. This architecture design boasts strong fault tolerance and high query flexibility, with its batch processing layer, speed layer, and service layer all being easily scalable. The data processing architecture model combines batch processing and real-time processing methods, aiming to provide a high-performance, low-latency, and fault-tolerant data processing solution. The computing layer employs Flink as the streaming data processing engine, responsible for reading shipping data streams in real-time from Kafka. The tasks of this layer involve real-time processing and storage of received ship data to support the system's real-time monitoring and analysis functions. For batch processing, Spark is used to analyze and mine data within specific time ranges.

3.3. Data server layer

The data service layer is implemented using a front-end and back-end separated architecture model. The back-end primarily consists of the View layer, Controller layer, Service layer, and Mapper layer. The front-end mainly follows the development method of responsive systems, designing and implementing a responsive ship data visualization system based on Vue technology. The base map plays a crucial role in the display of the entire electronic chart, and AIS-based ship display requires accurate and reliable electronic base maps. Utilizing third-party Mapbox to implement the base map ensures the accuracy of basic roads, rivers, and regions. To efficiently and uniformly load map data, the front-end employs the high-performance Openlayers interactive map library to display maps. The Mapbox map service supports HTTP and HTTPS protocols, adhering to the Web Map Tile Service (WMTS) standard of the Open Geospatial Consortium (OGC), and provides services such as vector base maps, vector annotations, image base maps, and image annotations.

4. Application of The Platform in Education

4.1. Project-based teaching

Project-based learning is a student-centered teaching method that facilitates student learning and growth through hands-on project experiences; the selection of project cases is crucial to teaching effectiveness, as appropriate cases can stimulate students' interest and enhance their practical abilities and problem-solving skills [7]. The curriculum system for the big data direction revolves around real-world smart shipping projects, with all courses covering the full range of knowledge and technologies required for project development. The big data application module for the ship's shore power system is a project that runs through the core courses of the professional group, serving as a common teaching carrier to organically link the courses within the professional group. The basic skills for big data applications in the ship's shore power system include system platforms, programming fundamentals, and data visualization basics, with the technical content used in the project being broken down into corresponding courses. Centered around the big data processing flow for the ship's shore power system, from acquisition, storage, analysis to visualization, this flow is integrated into the curriculum system. In the process of completing the project, students not only need to comprehensively apply professional knowledge and hone their operational skills, but also possess the abilities of autonomous learning, innovative thinking, teamwork, stress resilience, and responsibility. The project is disassembled into several sub-modules (courses), integrated with unified technical standards, and teachers collaborate and cooperate to implement project-based, modular teaching. A "dual-mentor system" is implemented between schools and enterprises, adopting a "mentor team + project" co-creation model between teachers and students, transforming the traditional teacher-centered "indoctrination" and "presentation" teaching methods into a student-centered, project-based, integrated learning and exploratory collaborative learning approach, stimulating students' intrinsic motivation to learn, and cultivating their innovative thinking, technological abilities, as well as their ability to integrate theory with practice, discover and solve problems.

4.2. Mining ideological and political elements in projects, developing curriculum-based ideological and political education with maritime-themed

The significance of curricular ideological and political education lies in organically integrating ideological and political education into various course teachings, reflecting the school's educational characteristics and educational philosophy [8]. Through curricular ideological and political education, not only can students' ideological and political qualities be enhanced, but the school's cultural heritage and educational characteristics can also be highlighted. The major delve deeply into the educational resources of ideological and political theory related to the strengthening of shipping, guiding and educating students to fully understand the status of industry and social development, continuously enhance their professional honor and industry pride, consciously serve the construction of a strong shipping nation and national strategies, integrate their individual "small selves" into the "greater self" of the motherland and the "greater self" of the people, and actively contribute to the development of the country. The major will make full use of distinctive educational resources such as projects from the Maritime Administration to ensure that ideological and political courses are easily conducted, grounded, and effective. For example, by introducing the display of ship positions based on Beidou signals, teachers can encourage students to strive for excellence through comparative analysis of the domestic and international technological status of navigation technology, cultivating students' meticulous craftsmanship spirit as a great nation, and inspiring their mission to serve the country through technology and build a powerful transportation and maritime nation.

5. Conclusion

Ship data analysis demonstrates significant value in the era of big data. It not only optimizes shipping efficiency and safety but also provides scientific support for corporate decision-making, driving the intelligent transformation of the shipping industry. In this context, integrating the ship data analysis system into big data technology teaching achieves a deep fusion of theory and practice, enhancing students' professional skills and industry insights. Through practical operations, students grasp the core skills of big data processing and analysis while cultivating innovative thinking and problem-solving abilities, significantly enhancing their competitiveness in the job market and supplying high-quality big data talent for the shipping industry and broader fields.

With the continuous development of big data technology, ship data analysis will increasingly focus on the application of advanced data analysis and mining technologies. For example, leveraging machine learning, deep learning, and other algorithms to conduct deep processing of massive ship data, uncovering potential relationships between data, predicting changes in ship performance, and optimizing navigation strategies. By constructing intelligent models, ship operation data can be automatically analyzed and processed, providing intelligent decision-making suggestions for ship managers, thereby improving ship operation efficiency and safety. Ship data analysis will no longer be limited to a single data source but will integrate multi-source data such as ship data, channel data, and meteorological data for comprehensive and accurate information systems, providing a more scientific basis for ship operation decisions.

Acknowledgements

This work was financially supported by the funding of the Research Project on Education and Teaching Reform of Higher Education in Jiangsu Province (2023JSJG647), the Research Program of Jiangsu Maritime Institute (2022ZKyb02), and the Excellent Teaching Team for QingLan Project of the Jiangsu Higher Education Institutions of China (Big Data Technology Teaching Team with Shipping Characteristic).

References

[1] Munim, Ziaul Haque, et al. "Big data and artificial intelligence in the maritime industry: a bibliometric review and future research directions." Maritime Policy & Management 47.5 (2020): 577-597.

[2] Aslam, Sheraz, Michalis P. Michaelides, and Herodotos Herodotou. "Internet of ships: A survey on architectures, emerging applications, and challenges." IEEE Internet of Things journal 7.10 (2020): 9714-9727.

[3] Aiello, Giuseppe, Antonio Giallanza, and Giuseppe Mascarella. "Towards Shipping 4.0. A preliminary gap analysis." Procedia Manufacturing 42 (2020): 24-29.

[4] Sepehri, Arash, et al. "The impact of shipping 4.0 on controlling shipping accidents: A systematic literature review." Ocean engineering 243 (2022): 110162.

[5] Lv, Taizhi, et al. "Research on Ship AIS Data Analysis Based on Stream Computing and Virtual Fence." 2023 4th International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE). IEEE, 2023.

[6] Müller, Ingo, Renato Marroquín, and Gustavo Alonso. "Lambada: Interactive data analytics on cold data using serverless cloud infrastructure." Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data. 2020: 115-130.

[7] Zhang, Lu, and Yan Ma. "A study of the impact of project-based learning on student learning effects: A meta-analysis study." Frontiers in psychology 14 (2023): 1202728.

[8] Tian, Feng. "Case Study and Practice of Ideological and Political Education in the Course." Advances in Educational Technology and Psychology 7.13 (2023): 64-71.