Big Data Management in Maritime Transport

Jović, Marija; Tijan, Edvard; Marx, Rebecca; Gebhard, Berit

Source / Izvornik: **Pomorski zbornik, 2019, 57, 123 - 141**

Journal article, Published version
Rad u časopisu, Objavljena verzija rada (izdavačev PDF)

https://doi.org/10.18048/2019.57.09.

ISSN 0554-6397
UDK: 004.65
656.61.025
Review article
Received: 10.10.2019.

Marija Jović
E-mail: jovic@pfri.hr
Edvard Tijan
E-mail: etijan@pfri.hr
University of Rijeka, Faculty of Maritime Studies, Studentska 2, 51000 Rijeka, Croatia
Rebecca Marx
E-mail: rebecca.marx@outlook.de
Hochschule Wismar, Fakultät für Wirtschaftswissenschaften, Philipp-Müller-Straße 14, 23966 Wismar, Germany
Berit Gebhard
E-mail: berit.gebhard@student.jade-hs.de
Jade Hochschule Wilhelmshaven/Oldenburg/ Elsfleth, Fachbereich Seefahrt und Logistik, Weserstr. 52, 26931 Elsfleth, Germany

Big Data Management in Maritime Transport

Abstract

As maritime transport produces a large amount of data from various sources and in different formats, authors have analysed current applications of Big Data by researching global applications and experiences and by studying journal and conference articles. Big Data innovations in maritime transport (both cargo and passenger) are demonstrated, mainly in the fields of seaport operations, weather routing, monitoring/tracking and security. After the analysis, the authors have concluded that Big Data analyses can provide deep understanding of causalities and correlations in maritime transport, thus improving decision making. However, there exist major challenges of an efficient data collection and processing in maritime transport, such as technology challenges, challenges due to competitive conditions etc. Finally, the authors provide a future perspective of Big Data usage in maritime transport.

Keywords: Big Data Management, Big Data Analysis, Maritime Transport, Shipping Safety, Energy Efficiency

1. Introduction

Mobility and maritime transport have always played an important role in economic, ecological and social development [23]. With the development of international trade and globalization, traffic and cargo volumes keep increasing. Therefore, numerous stakeholders (such as maritime logistics companies, forwarders, agents etc.) are compelled to accept changes in the maritime transport sector and turn to more effective practices by introducing technologies that can gather and process massive amounts of information (in a cost-effective way) [35].

A large amount of data is generated from different sources and in different formats in maritime transport on a daily basis. This includes traffic data, cargo data, weather data and machinery data [65]. Due to the size of the maritime transport network that includes the aforementioned stakeholders, there exist large scale planning problems at the strategic, tactical and operational level [10]. Nowadays, Big Data analytics are applied to many industries (among which the maritime transport industry) to promote better quality of decision-making processes [6].

According to Acharjya and Ahmed [1], "Big Data means more than just dealing with a large quantity of data. In general, it refers to the collection of large and complex datasets which are difficult to process and analyse using traditional database management tools or data processing applications". When handling Big Data in large quantities, advanced data-processing techniques and tools are required in order to effectively analyse and utilize these data [6].

Various stakeholders involved in cargo transports (from origin to destination) are constrained to rely on the information provided by other parties involved. The more exact and detailed such information, the smoother running of shipping operations. The level of cooperation and coordination that is essential for the success of these businesses and competitiveness is based on information systems and information technology [61]. Decisions based on the information extracted from raw data can lead to several advantages: increased safety and resource utilization, as well as a higher degree of efficiency, sustainability and environmental protection [17].

The authors will analyse recent Big Data innovations, solutions and applications in maritime transport, and the advantages of such applications. The goal of the research is to accentuate the importance and possibilities of Big Data management in maritime transport. The research problem stems from increased costs and insufficient prediction accuracy (for example, predictions of vessel arrival times and the needed speed adjustment) due to inefficient Big Data management and analysis. This paper presents a review of research papers dealing with this topic. The research results contribute to a better understanding of Big Data management in maritime transport (considering different possible applications of Big Data in maritime transport). The paper has been drafted in connection with Activity 3.2. of the DigLogs project "Impact analyses of Big Data management for freight and passengers' mobility". The DigLogs project involves cooperation between two countries - Italy and Croatia - and is being implemented with the aim of creating the necessary concepts, technological solutions, models and planning to establish the most advanced digitized logistic processes for the multimodal freight traffic and passenger services in the program area.

2. Big Data in General

The concept of Big Data comes with a set of related components that enable organizations to put the data to practical use and solve a number of business problems. These include the IT infrastructure needed to support Big Data, the analytics applied to the data, technologies needed for Big Data projects, related skill sets, and the actual use cases that make sense for Big Data [63]. According to Nita and Mihailescu, the sources of data could be communication, transport, social media, climate, search engines, GPS signals, online shopping, mobile devices, and many others [41]. The sources of data can be the Internet of Things (IoT) as well, creating large data sets on a daily basis. IoT consists of various devices, which are provided with digital sensors and connected for interacting. The number of these devices, such as mobile phones or Radio Frequency Identification tags (RFID) is rising rapidly [36], [15].

Even though there is no official definition of Big Data, the concept of Big Data goes beyond the literal meaning of the data great volume. Taylor-Sakry [57] defines Big Data as "large sets of complex data, both structured and unstructured which traditional processing techniques and/ or algorithms are unable to operate on." This definition also points out the importance of specific technology or methods, such as algorithms, to handle Big Data.

Big Data is also described by its characteristics of the "3 V's" that are represented by Volume, Variety and Velocity [52], [15]. Those characteristics were first identified by Gartner analyst Doug Laney in 2001 [14]. Sometimes this concept is extended into four or five V's by adding the characteristics of Veracity or Value [44], [57], [36]. Fundamentally, characteristics of Big Data can be described as stated below:

- Volume: The quantity of data, that needs to be stored and analysed, is large and consistently growing. Hammad et al. [15] expect the volume of data in the world to exceed 40 zettabytes by 2020. In the traditional database system, the relationship between data items can be explored easily as the traditional database system does not contain massive or voluminous data. However, Big Data contains massive or voluminous data that increase the level of difficulty in figuring out the relationship between data items [2].
- Variety: As mentioned in the definition, Big Data contains structured and unstructured data. This means that their analyses need to include different structures such as texts, audio, video and images at the same time [15]. Also, the number of resources causes varieties in data sets [44].
- Velocity: Velocity describes the high production rate of upcoming new data that consequently requires the need of faster processing [44].
- Veracity: Since the quality and reliability of analysis outcomes depend on the input data, data resources have to be reliable and secure. In addition, data need to be protected from unauthorized access [44].
- Value: While the value of data is low in relation to its volume just by collection, high (economical) value will be created after the Big Data Process, which

provides extensive insights. Consequently, Big Data can impact companies and the society [36].

The latter presumes a transformation of value within Big Data Processes. Those processes can be divided into two main aspects. The first aspect is represented by data management that consists of acquiring and preparing the data, which means e.g. cleaning and aggregating them. The second aspect is the part of analytics, in which data are to be firstly analysed, and the results to be afterwards interpreted [13]. These processes are shown in Figure 1.

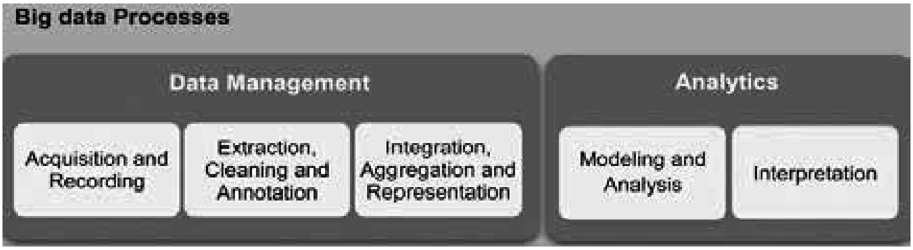

Figure 1: Big Data Processes
Source: [13]

As stated above, Big Data Analyses (BDA) require specific technology and techniques that go beyond traditional (statistical) methods. Thus, wider insights are generated and affect the quality of interpretation. Consequently, De Mauro et al. [36] emphasize that Big Data changes the way the well informed decision making shifts “from the logical, causality-based reasoning to the acknowledgment of correlation links between events”. The so-called data-based or knowledge-based decision making is already present in numerous successful applications in industry, such as the link prediction and personalized commercials in e-Commerce and Social Media, as well as in the medical sector and much more [44].

3. Big Data in Maritime Transport

Maritime transport is characterized by a large number of stakeholders such as seaport operators, shipbuilders and ship owners, agents, brokers, shipping and insurance companies or classification societies, etc. A large variety of stakeholders means a large variety of business procedures and interests, among which there are various interests in data types. Thus, a clear definition of the term “Maritime Big Data” does not exist. In dependence on the target, Maritime Big Data includes details of ships’ performance, freight rates, weather data, labour costs, oil or even metal prices [24]. In addition, the number of digital sensors in the maritime transport is also increasing. Consequently, the generation of Big Data in the maritime context is increasing, adding to the quantity of data, and the data are provided by numerous different sources [46]. Beside weather forecasts and historical data, most important data resources in the case of voyage data are provided by the bridge equipment to be recorded by the Voyage Data Recorder (VDR), and by external monitoring such as the Automatic Identification System (AIS) [24], [38].

The main purpose of VDR is to provide data for analyses in case of accidents. Therefore, VDR stores all recorded data of a voyage [38]. As the International Maritime Organization (IMO) determines (Regulation 20, Chapter 5 of the International Convention for the Safety of Life at Sea – SOLAS, 1974), a VDR needs to be installed on all passenger ships and other ships of 3000 gross tonnage or more on international voyages [55]. Information collected by VDR are: date and time, ship's position, speed, heading, bridge audio, communication audio, wind speed and direction, main alarms on the bridge and rudder or engine orders and responses and more [21]. According to Resolution MSC.333(90) (adopted on 22 May 2012), the time for which all stored data items are retained should be at least 30 days/720 hours on the long-term recording medium and at least 48 hours on fixed and float-free recording media. Data items older than this may be overwritten with new data. A standard international voyage can last longer than 30 days and data collected during this journey will be overwritten and not extracted and saved on an external device. On short voyages, the data will also be deleted after the voyage or during the upcoming voyage, making them inaccessible for a long term analysis [24]. VDR collects a massive volume of data, providing deep insights into the voyage. As the VDR data are overwritten after a specific period and replaced with new data, the data should be sent ashore or be saved and exported manually or automatically. Mirović et al. [38] claim that such information is still analysed only in cases of accidents, and is otherwise deleted without consideration.

The AIS tracking system was originally developed as the collision avoidance tool that enables commercial vessels to 'see' each other more clearly in any conditions and to improve helmsman's information about the surrounding environment. The AIS does this by continuously transmitting vessel's position, identity, speed and course, along with other relevant information, to all other AIS equipped vessels within the range [3]. In this respect, the AIS enables safer and more efficient navigation by tracking all ships within the range (enabling the exchange of ship data among ships and the shore) [29], [38]. AIS regulations are defined in the SOLAS as well, and require AIS transponders to be installed on all passenger ships and all other ships of 300 gross tonnage and more [54]. The AIS provides other ships and coastal authorities with static data (IMO number, ship's length and beam etc.), voyage related data (destination and ETA, etc.), dynamic data (position, course, speed etc.) and VTS data (short content information related to various safety warnings and information on areas with warnings about navigational and other dangers) [62], [38], [9]. Thus, it helps collision avoidance and decision making on board in real time [29]. In this respect, the AIS provides data as a basis for BDA in the maritime transport. The frequency of AIS data refreshing rate depends on the vessel movement and navigational status, meaning that the data quantity will increase as the vessel moves faster and/or alters the course.

However, the growing trend of the AIS has been such that in some most congested waters the system is already overloaded as of today. Given the danger that this overload can represent for the main mission of the AIS, that is collision avoidance, the International Association of Marine Aids to Navigation and Lighthouse Authorities (IALA) and a number of national maritime authorities have started the work on the VHF Data Exchange System (VDES). Rather than an evolution of AIS, VDES is a communications system encompassing different subsystems of communications, one of them being the AIS [27], [19]. New techniques providing higher data rates than those used for the AIS are a core element of VDES. The VDES supports e-Navigation [27] and could have a significant positive impact on the provision of maritime information services, such as maritime safety information, general data communications at high data rates, locating, vessel traffic management, satellite communications, etc. [59].

Maritime transport offers favourable conditions for Big Data Processes, by being provided with a wide range of data from various sources and by being obliged to record such data. As mentioned before, Big Data Analyses provide deeper insight into processes and events and allow for a well informed decision making. In the same manner as in other industries, maritime industry might benefit Big Data Analyses. In the maritime transport (as part of the maritime industry), innovations in BDA will bring advantages to efficient routing, operational planning and improvements in safety, as shown in the following paragraphs.

There exist several current applications of Big Data in the maritime transport and one of them is “operations”. In this respect, ship owners can determine the optimum speed for fuel consumption, taking into consideration factors such as bunker cost, freight rates and schedules. Speed reduction will result in the fuel consumption optimization and will also significantly contribute to decreasing the emission of harmful gases, thus getting in consonance with the development of ecological legislation [67]. Other applications of Big Data are: voyage operations (vessels can be tracked using dashboards instead of relying on notes, emails or phone calls), tracking etc. [58]. For example, Freight Metrics provides the location visibility of a ship or vessels located around a particular port. Freight Metrics has plotted locations of major shipping ports around Australia on Google Maps. It enables a simplified connection to the port website for shipping schedules [51].

In Singapore and Malaysia, ports utilize Big Data techniques to create advanced inspection systems to assess the history and cargo type of importers. The Hamburg Port in Germany uses a cloud-based analytics tool called SmartPortLogistics. The tool pulls different types of data, such as vessel positions, height and breadth of bridges, and planned sailing routes. These data can be viewed in real-time on mobile applications [58].

3.1. Big Data and e-Navigation

With the development of navigational systems, sensors and tracking systems following recent advances in technology, the maritime industry is opening up to the benefits of the digital era [38]. In this respect, the International Maritime Organization (IMO) has developed an e-Navigation concept in order to improve shipping through better organization of data on ships and on shore, and better data exchange and communication between ships and between the ship and the shore [33]. The e-navigation Strategy Implementation Plan (SIP) introduces a vision of e-navigation which is embedded in general expectations for onboard, onshore and communications elements. The main objective of the present SIP is to implement the five prioritized e-navigation solutions [20]:

1. Improved, harmonized and user-friendly bridge design;
2. Means for standardized and automated reporting;
3. Improved reliability, resilience and integrity of bridge equipment and navigation information;
4. Integration and presentation of available information in graphical displays received via communication equipment; and
5. Improved Communication of Vessel Traffic Service (VTS) Service Portfolio (not limited to VTS stations).

Consideration should be given to the IMO Formal Safety Assessment (FSA) from which a number of required tasks have been identified. These tasks, when completed in the period 2015–2019, should provide the industry with harmonized information, in order to start designing products and services to meet the e-navigation solutions [20].

The Maritime Connectivity Platform (MCP), a cloud-based technology and an open source digital maritime domain, enables common Internet standards to maritime navigation and transport systems [16]. It is a communication framework enabling efficient, secure, reliable and seamless electronic information exchange between all authorized maritime stakeholders across available communication systems [34]. As navigation systems become more advanced, the amount of ship performance and navigation data thus generated is becoming ever more significant. Big Data analytics tools make it possible for these large quantities of data to be analyzed in order to gain the insight that supports the decision-making [38].

3.2. Big Data Innovations in the field of Seaport Operations

Similarly to previous analyses, the BDA can also be used for predictions of arrival times and calculations of the required speed. In general, shipping companies are under the pressure of meeting the agreed arrival times at the seaport, while being faced with the problem of the named uncertainties of weather conditions. Consequently, managers often decide to deviate from the efficient low steaming in order to prevent delays

[28]. On the other hand, seaport operations are sometimes delayed due to different reasons, such as arrival times of prior ships, loading or unloading processes, etc. Both circumstances will lead to waiting times on available slots for the arriving ship, being anchored outside of the port.

To prevent this situation, the performance and weather data can be analysed and considered with tracking data, as provided by the AIS. The planned (optimum) route matched with real time data on outside circumstances, like the weather, and the tracked position, can thus be used to forecast arrival times [38]. However, it is necessary to divide the standard AIS transceiver from AIS Application Specific Messages (referred here). An AIS transceiver transmits and receives static data (Maritime Mobile Service Identity, ship name and callsign), as well as dynamic data (position data, speed and course over ground) via VHF on two channels specially reserved for AIS [56]. AIS Application Specific Messages (ASMs) are a quick way to get important information through the VDL in small, predefined packages. A ship, for example, can use the ASMs to transmit important details such as: dangerous cargo information, number of persons on board, extended ship static and voyage-related data and route information [18].

Data Science Campus (DSC) [8] has explored the likelihood that a ship would be delayed and would arrive at its intended destination some time after its estimated time of arrival. Ship delays were predicted using the supervised machine learning and these predictions can be used to predict port loading at a point in time and can support any subsequent operational port planning. Both the Automatic Identification System (AIS) and the Consolidated European Reporting System (CERS) datasets have been used and a means by which they could be joined has been developed. The common identifier between the AIS and CERS data is the Maritime Mobile Service Identity (MMSI, the unique identifier of each ship) [8]. To merge the two datasets together they were first inner joined on the MMSI, the datasets were then filtered by restricting each timestamp within the AIS to its closest estimated time of arrival (ETA) or estimated time of departure (ETD) (both ETA and ETD were considered since the AIS data contained both the inbound and outbound portions of the journey). Records were retained where the ETA or ETD were falling within 24 hours of the timestamp. As it is possible for more than one ETA or ETD to meet these criteria, the ETA or ETD closest to the timestamp was selected. After the merging, the complete dataset included 727 voyages relating to 235 unique ships. An in-depth explanation of the research can be found in the report [8].

Big Data enables optimized usage of resources and infrastructures as well. For example, a typical crane operator works only one-quarter of the time, remaining idle for three-quarters of the time, waiting to get a container ready to load or for an empty truck to load a container. Increasing the number of trucks may not be a viable solution owing to the congestion it would cause. Rather, Big Data analysis could synchronize movements, so that the crane operator can work more time. For instance, signals related to the crane position, status, and GPS position signals could sync the movement of trucks and containers, to reduce idling time [30].

The Bahri maritime company has rolled out a series of initiatives aimed at harnessing the power of "Data Innovation" to enhance productivity, unlock opportunities for growth and transform the existing operations model in the shipping industry [47]. In this respect, the company has created BahriData, a Big Data platform, to improve the operating performance and to unlock growth opportunities. The company has developed various data models in its BahriData platform to cover various key business areas such as chartering, voyage management, fleet operations, maintenance and customer services [58].

3.3. Big Data Innovations in the field of Security

The security of ship operations can be affected by several different reasons. In order to improve the general security of ship operations and to eliminate causes such as acts of piracy, various data about ship details and positions need to be monitored. The use of Big Data can not only prevent accidents caused by weather conditions, but also those caused by other participants in the maritime traffic. In some seas, acts of piracy still pose a great risk to shipping. Information regarding most recent pirate activities can help to divert other ships that are currently on the same route.

One of the systems in the field of security is the Long Range Tracking and Identification (LRIT). It is an international tracking and identification system incorporated by the IMO under the SOLAS convention to ensure a thorough tracking system for ships across the world [37]. The system aims to enhance security for government authorities. LRIT provides ship identity and current location information in sufficient time for a government to evaluate the security risk posed by a ship off its coast and to respond to reduce the risk if necessary. The system complements existing modes of tracking ships using coast-based AIS stations. Furthermore, the combined use of LRIT and Satellite-AIS data can increase the tracking quality and coverage of ships registered in the EU CDC [11].

3.4. Big Data Innovations in the field of Weather Routing

Weather conditions may cause several issues in the maritime transport. Due to the direction or strength of winds and waves, the engine power needs to be increased to maintain the speed. By BDA of historical weather data, eventual harsher weather conditions on certain routes can be detected and data could be used to adapt the future strategic route planning [28]. Furthermore, data about the ship performance and navigation information, such as the location, average draught, main engine power, speed and fuel consumption combined with wind speed and direction, can support navigation strategies, as well as intelligent decisions, in real time as well [45]. With the provision of real-time weather data (wind speed and direction, wave heights, storm forecasts) and ship data (regarding the trim and rolling), dangerous situations can be identified immediately. Furthermore, alternative (less hazardous) routes outside critical areas can be calculated faster. The data regarding the ships that are already within a hazardous area can help other ships to avoid the same difficulties, thus improving safety of the crew and other ships. The risk of sinking or cargo loss due to severe weather conditions may be minimized. For example, the port of Singapore uses the platform that enables information on ship positions and weather data to be collected. The platform helps avoid accidents by inferring the most likely path ships would take in a given situation [58].

FastSeas [12], as another example, is a cloud-based weather routing and passage planning tool that enables calculating the fastest route (from point A to point B) according to the current NOAA GFS weather forecast (a weather forecast model produced by the National Centers for Environmental Prediction), current oceanic currents, vessel performance, and “comfort criteria”. FastSeas has to be “taught” the vessel performance. The performance settings must be set only once. FastSeas will save these settings and use them next time. After the calculation is complete, the route will be displayed on the map.

In order to achieve an efficient maritime transport service, shipping companies can also implement specialized software for performance and route analyses [26]. Weather-routing software can help users determine the optimal departure time and routing based on user-specified parameters. These parameters can include maximum acceptable wind speeds or wave heights, as well as the percentage of time that the vessel will be on a particular point of sail [50]. Weather-routing can also enable fuel savings, which will not only bring the benefit such as the cost reduction, but will also support the achievement of emission standards and support the Ship Energy Efficiency Management Plan (SEEMP) [38], [45]. There exist several examples of such software such as the Transas Navi-Planner, Octopus and BonVoyage. Transas Navi-Planner utilises one of the world’s largest navigational databases, as well as the artificial intelligence to auto-create a route that is safe to sail. It calculates weather optimisation, supports hazard identification, creates a voyage plan, and provides at all times the latest charts and data automatically [40]. The OCTOPUS software assists ship officers and engineers in making real time decisions enabling them to be proactive in safety and efficiency actions, resulting in more immediate benefits than just a traditional post voyage analysis [32].

The BonVoyage System is a combination of software and data. BonVoyage is a user-friendly icon-driven graphical marine weather briefing system. By using real-time weather data to generate color-coded graphics, the software lets the user see into the heart of severe weather systems. This provides a much more complete picture of the storm than traditional text-based weather bulletins or radio facsimiles, because it captures the detailed shape of each storm system and visually displays dangerous wave generating areas [7].

3.5. Big Data Innovations in the field of Monitoring/Tracking

To avoid collisions of ships in seaport areas, forecasts of vessel arrival times with real-time information about their current speed are made to illustrate the estimated traffic [58]. Through this technology, collision issues are identified and vessels are notified before accidents could happen. The encounter of two vessels in narrow areas can be avoided. For example, the Vessel Traffic Service is designed to improve maritime safety, to support security activities in the maritime sector, to promote smooth and efficient maritime traffic, and to prevent accidents and the potential environmental harm they may cause. The system displays a graphical environment with movements of vessels in the approach areas, putting each of these overlapping vessels to a digital nautical chart, in its real geodesic position and informs the identification of each vessel [60].

Risky shore areas (such as bunkering facilities for fuel or liquid cargo) require a special control in order, for example, to detect unauthorized ship movements as early as possible by monitoring the information about all vessel and vehicle movements [53]. Collisions outside the seaport areas can be avoided by collecting data about the standard routes of vessels on a long-term basis. These data reveal if a vessel deviates from the declared path [38]. This process is shown in Figure 2.

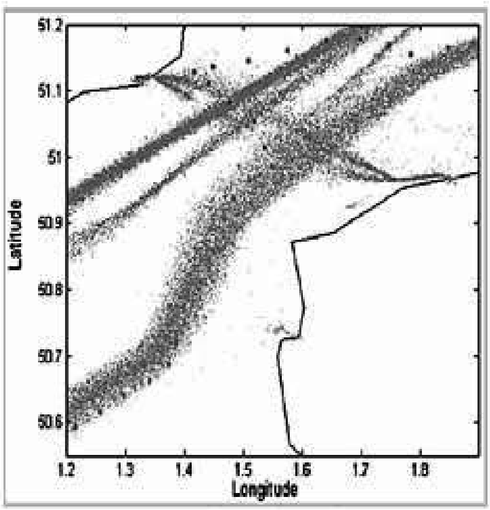

Figure 2: Deviations from the declared path
Source: [4]

The plots in Fig. 2 depict the maritime traffic in the Dover Strait; two-weeks of raw AIS data are represented [4]. The vessel deviates from the known route (red dots), approaching a port without stopping and finally returning to its declared route.

Monitoring of technical equipment enables higher safety for the crew and environment and increased longevity of ship components. During the use of the equipment and engines, data are collected and sent to manufacturers, who can detect irregularities and increase their service quality [24].

The Satellite AIS (S-AIS) is a vessel identification system that tracks the location of vessels in the most remote areas of the world, especially over open oceans and beyond the reach of terrestrial-only AIS systems. S-AIS is used for collision avoidance, identification and location information as well as for maritime domain awareness, search and rescue, environmental monitoring and maritime intelligence applications [43]. Analyses of oceangoing ship routes are also important in the function of environmental protection. The analyses of ship routes in the mentioned oceanic areas covered by S-AIS could be very important to estimate air pollution emissions from ships, seeking to minimize their negative impact [49].

4. Challenges of Efficient Data Collection and Processing

Big data challenges include data capturing, data storage, data analysis, search, sharing, transfer, visualization, querying, updating, information privacy and data source [25]. During the early days of Big Data use in the maritime sector, three main types of challenges to the efficient data collection and processing were detected: competitive conditions [24], [58], technology [46], [64] and human resources [31].

4.1. Challenges due to Competitive Conditions

By storing many datasets in the same infrastructure it becomes difficult to overview different owners and to avoid the monopolization of property rights because a strong demand for only one data provider leads to monopolistic deals [24]. Mutually accepted agreements and legal framework with clear instructions and rights need to be implemented by international and local policies to ensure competitive conditions. For instance, ship owners use Big Data to examine the efficiency of their schedules to fit in with supply lines, to consider which cargo to transport, when and how, and with what fuel consumption to keep costs down etc. [5]. Although the ship owner (on whose ship the data are collected) is considered the data owner, the issue arises when third-party services are included in order to analyse the data. The absence of data ownership regulations and the inability to ensure access for all those involved in the process of collecting and deriving such data present an important issue [46]. To ensure access for all those involved, a centralized unique database is required. Currently, the collected data cannot be properly shared and some pieces of information are still unavailable to some parties in the maritime transport. Therefore, any comprehensive decision making is impossible.

Past shipping crises, the falling freight rates, and the oversupply and margin pressure have slowed down further development of Big Data solutions. In such an unstable environment, investors are uncertain about implementing new technologies [58].

4.2. Technology Challenges

The majority of Big Data challenges can be seen in technological fields. As mentioned before, a standard international voyage can take more than 30 days and the data collected during this journey will be overwritten and not extracted and saved on an external device. On short voyages, the data will also be deleted after the voyage or during the upcoming voyage, thus becoming inaccessible for a long term analysis [24]. The same applies to the rapid pace the generation of new data follows. Processing of incoming data, that should provide proper real-time information in support of the ongoing transportation operations, becomes impossible due to the lack of suitable processing and storage devices. The traditional toolsets are no longer enough, as well as the offered bandwidth for the high seas communication [38].

The quality of collected data can be affected by various factors (accidental or deliberate). Data that needs to be inserted manually into report forms, computer systems or AIS transceivers, present an important source of errors, as there is no way to validate it [46]. Unauthorized individuals can wilfully insert, delete or update information in order to reach economic advantages or to disguise illegal activities. “Ships are increasingly transmitting false or stolen identifying marks, taking advantage of the AIS ‘honours system’, as ships are required to transmit their information […]” [64]. One percent of all vessels are using wrong IMO-Numbers, more than 50% do not report their next port of call, and 25% are turning off their AIS devices regularly [64]. A short-term solution to this information gap is the extensive data validation [24].

Some of the challenges of the S-AIS technology include satellite revisit times, message collision, and ship detection probability. The data processing latency and lacking the continuous real-time coverage have made it less reliable for end users in certain aspects of monitoring and data analysis. Recent developments and improvements by leading S-AIS service providers have reduced latency issues [49].

Another obstacle concerns cyber threats. Wireless exchange of data always involves the risk of interception or penetration from cyber criminals and terrorists. The implementation of Big Data solutions makes the maritime transport more vulnerable to hostile attacks than before. Definitions regarding illegal activities, a strategy to render them harmless, and moreover, guidelines for secure handling of personal data are missing. More than 90% of major container operators are vulnerable to cyber-attacks, according to a report by the Maritime Cybersecurity company CyberKeel [58].

The IMO has published guidelines for the maritime cyber risk management and they recommend that stakeholders need to take the necessary steps to safeguard shipping from current and emerging threats and vulnerabilities relating to the digitisation, integration and automation of processes and systems in shipping. The IMO is not prescriptive in how these recommendations should be implemented, but refers to best practices from NIST, industry organisations and ISO/IEC 27001 as sources of additional guidance and standards [39]. The ISO/IEC 27001 is the best-known standard in the family providing requirements for an information security management system (ISMS) [22].

4.3. Human Resources Challenges

Developing new devices for data processing and validating the collected data is labour intensive and requires a high degree of technological knowledge in analytics, statistics and software modelling, as well as experts such as computer scientists, mathematicians and data scientists [17]. Maritime transport may suffer from a lack of skilled workforce, a problem that is expected to increase in the future because the development of a Big Data solution requires new skill sets, technological knowledge in analytics, statistics, software modelling etc. Enough quantity and quality of human resources needs to be ensured through cooperation between universities and the private sector for further development and implementation of Big Data solutions in maritime transport [24].

5. Future perspectives of Big Data usage in maritime transport

Since the use of Big Data is increasingly present in the maritime transport, a number of policy makers and organizations have become increasingly aware of the importance of Big Data and BDA. According to Voss et al. [61], "Regarding the future of logistics, we might concede that "the technology" is there, it just has to be applied".

According to Zgurovsky and Zaychenko [66], "Big data and Internet of Things work in conjunction. Data extracted from IoT devices provide mapping of device interconnectivity. IoT is also increasingly adopted as a means of gathering sensory data, and this sensory data has also been used in transportation contexts". The evolution of IoT and the use of Big Data create for logistics the prospect of becoming a data-centric industry, where information takes precedence in propositions of the value for logistics services over the actual ability to move cargo [42].

Despite numerous benefits, increased costs related to Big Data storage and manipulation may represent a large issue. When handling Big Data, advanced data-processing techniques and tools are required in order to effectively analyse and utilize the data. Specific technologies or methods, such as algorithms to handle Big Data are needed. However, although the maritime transport industry is suffering from recent crises [48], [58], a rising number of companies are offering high-technology solutions to optimize ship operations regarding the optimum speed, fuel consumption etc. This number is expected to rise continuously in the future. Internal platforms are being developed to provide more efficient use of collected data. Future prospects anticipated through the more efficient use of collected data is a higher level of cooperation between actors in a transport operation or supply-chain, which provides real-time information, thus enabling cost savings. Due to the analysis of weather data, a more efficient routing can decrease fuel costs [58]. A long-term analysis of waiting times for a berth in a specific seaport can enable cost savings through reducing the average speed or changing the route of the ship in order to avoid waiting times. Collecting ship engine data or machinery data can provide increased safety, thus lowering the operational risk. Errors can be detected and even dealt with immediately. Damage can be avoided, and fixed maintenance schedules will become less frequent or even unnecessary. In a long run, this can reduce high docking cost for larger damages caused by the ongoing use of defective equipment and needless service, because maintenance can now be scheduled when needed.

Ship owners and operators are exploring the use of Big Data analytics to reduce their bunker costs. Low bunker costs can offset the record low freight rates in the market. Maritime software allows companies to achieve fuel savings through energy efficiency retrofits, using big data collection and analysis. For example, the ClassNK-NAPA GREEN software, collectively developed by ClassNK and NAPA, offers a real-time Big Data analysis performance monitoring and optimization solution for fuel efficiency purposes [58].

6. Conclusion

Maritime transport involves a large number of stakeholders such as ship owners, insurance companies, agents, etc., and numerous business procedures and interests. In this respect, maritime transport produces a large amount of data from various sources and in different formats. Several applications of Big Data in the maritime transport can be found, such as: operations (enabling ship owners to determine the optimum speed, thus affecting fuel consumption), voyage operations, tracking or monitoring etc.

Big Data and its analyses provide deep understanding of causalities and correlations in maritime transport, improving decision making. For example, BDA can be used for predictions of arrival times and calculations of the needed speed. It has been also shown that BDA enables ship owners to determine the optimum speed for fuel consumption, taking into consideration several factors such as bunker cost, freight rates and schedules. Furthermore, BDA ensures higher safety at sea by preventing collisions and machinery failures. Besides, the analysis of Big Data can bring advantages to the maritime transport sector regarding efficient routing, operation optimization and safety improvements. Big Data innovates the maritime transport by improving shipping efficiency and operations as well as by increasing safety at sea.

However, Big Data usage in the maritime transport has yet to overcome challenges to fully exploit these advantages due to competitive conditions, lack of suitable processing and storage devices and shortcomings in human resources. Issues such as the lack of regulations for the market environment, cyber-crime and lack of skilled experts have slowed down the BDA development in the maritime transport sector. For example, extensive data sets collected by VDR are deleted without being considered for analyses.

As the Big Data usage is increasing in the maritime transport sector, policy makers and organizations have become increasingly aware of the advantages of Big Data and BDA. Furthermore, a rising number of companies are offering high-technology solutions in order to achieve optimized ship operations etc. This number is expected to rise in the future. The IoT is one of the examples of the aforementioned high-technology solutions, closely connected to the Big Data usage. In this respect, the IoT technology and the use of Big Data provide logistics with the prospect of becoming (including maritime transport) a data-centric industry. Besides, the collection and analysis of Big Data regarding, for example, the volume of traffic to avoid collisions in narrow seaport areas is expected to be the first step towards autonomous shipping.

Acknowledgments

This work has been supported by the “DigLogs - Digitalising Logistics Processes” (Interreg V-A Italy-Croatia 2014-2020) project.

References

1. Acharjya, D. P. and Ahmed, K. (2016) ‘A Survey on Big Data Analytics: Challenges, Open Research Issues and Tools’, *International Journal of Advanced Computer Science and Applications*, 7(2). doi: 10.14569/ijacsa.2016.070267.
2. Aggarwal, D. (2016) *Difference between traditional data and big data*. Available at: https://www.projectguru.in/publications/difference-traditional-data-big-data/ (Accessed: 6 August 2019).
3. AISHub (2019) *AIS data sharing and vessel tracking by AISHub*. Available at: https://www.aishub.net/ (Accessed: 6 August 2019).
4. Alessandrini, A. *et al.* (2017) ‘Mining Vessel Tracking Data for Maritime Domain Applications’, in *IEEE International Conference on Data Mining Workshops, ICDMW*. IEEE, pp. 361–367. doi: 10.1109/ICDMW.2016.0058.
5. *Big data and their hidden threats to shipping industry - SAFETY4SEA* (2019). Available at: https://safety4sea.com/big-data-and-their-hidden-threats-to-shipping-industry/ (Accessed: 19 August 2019).
6. *Big Data in the Shipping Industry* (2018). Available at: https://www.morethanshipping.com/big-data-in-the-shipping-industry/ (Accessed: 16 August 2019).
7. *Bon Voyage System* (2019). Available at: http://www.ships-for-sale.com/bon_voyage_weather_forcasting.htm (Accessed: 9 December 2019).
8. Bonham, C. *et al.* (2018) *Analysing port and shipping operations using big data*. Available at: https://datasciencecampus.ons.gov.uk/wp-content/uploads/sites/10/2018/06/here.pdf (Accessed: 6 August 2019).
9. Bošnjak, R., Šimunović, L. and Kavran, Z. (2012) *Automatic Identification System in Maritime Traffic and Error Analysis*, *Transactions on Maritime Science*. Available at: https://hrcak.srce.hr/file/132144 (Accessed: 19 August 2019).
10. Brouer, B. D., Karsten, C. V. and Pisinger, D. (2016) ‘Big Data Optimization: Recent Developments and Challenges (Big Data Optimization in Maritime Logistics)’, in. Springer International Publishing.
11. EMSA (2019) *VESSEL TRACKING GLOBALLY - Understanding LRIT*. Available at: http://www.emsa.europa.eu/lrit-home/download/452/256/23.html (Accessed: 6 August 2019).
12. FastSeas (2019) *What is FastSeas?* Available at: https://fastseas.com/#what (Accessed: 6 August 2019).
13. Gandomi, A. and Haider, M. (2015) ‘Beyond the hype: Big data concepts, methods, and analytics’, *International Journal of Information Management*. Elsevier Ltd, 35(2), pp. 137–144. doi: 10.1016/j.ijinfomgt.2014.10.007.
14. *Gartner's Big Data Definition Consists of Three Parts, Not to Be Confused with Three Vs* (2013).

Available at: https://www.forbes.com/sites/gartnergroup/2013/03/27/gartners-big-data-definition-consists-of-three-parts-not-to-be-confused-with-three-vs/#35419c1942f6 (Accessed: 6 August 2019).
15. Hammad, K. A. I. *et al.* (2015) 'Big Data Analysis and Storage', in *International Conference on Operations Excellence and Service Engineering*. Orlando, Florida, USA, pp. 648–659.
16. HARBOURS REVIEW (2019) *Launch of the Maritime Connectivity Platform Consortium (MCC)*. Available at: http://harboursreview.com/launch-of-the-maritime-connectivity-platform-consortium-(mcc).html (Accessed: 6 August 2019).
17. Heilig, L., Schwarze, S. and Voss, S. (2017) 'An Analysis of Digital Transformation in the History and Future of Modern Ports', in *50th Hawaii International Conference on System Sciences*. Hawaii, pp. 1341–1350. doi: 10.24251/HICSS.2017.160.
18. IALA (2013) *IALA Guideline No. 1095 On Harmonised implementation of Application-Specific Messages (ASMs) - Edition 1*. Available at: https://www.e-navigation.nl/sites/default/files/1095 Implementation of Application-Specific Messages.pdf (Accessed: 19 August 2019).
19. IALA (2019) *VDES FAQ*. Available at: https://www.iala-aism.org/technical/connectivity/vdes-vhf-data-exchange-system/ (Accessed: 6 August 2019).
20. International Maritime Organisation (IMO) (2014) *NCSR 1/28: E-Navigation Strategy Implementation Plan. London: IMO*. Available at: http://www.imo.org/en/OurWork/Safety/Navigation/Documents/enavigation/SIP.pdf (Accessed: 6 August 2019).
21. International Maritime Organization (IMO) (2012) *ADOPTION OF REVISED PERFORMANCE STANDARDS FOR SHIPBORNE VOYAGE DATA RECORDERS (VDRs), MSC 90/28/Add.1, Annex 21*. Available at: http://www.imo.org/en/KnowledgeCentre/IndexofIMOResolutions/Documents/MSC - Maritime Safety/333(90).pdf (Accessed: 6 August 2019).
22. ISO (2019) *ISO/IEC 27001 INFORMATION SECURITY MANAGEMENT*. Available at: https://www.iso.org/isoiec-27001-information-security.html (Accessed: 9 December 2019).
23. Jović, M. *et al.* (2019) 'The Transition of Croatian Seaports into Smart Ports', in *42nd international convention on information and communication technology, electronics and microelectronics*. Opatija, Croatia, pp. 1618–1622. Available at: https://www.bib.irb.hr/986081 (Accessed: 13 May 2019).
24. Koga, S. (2015) *Major challenges and solutions for utilizing big data in the maritime industry*. Available at: http://commons.wmu.se/all_dissertationshttp://commons.wmu.se/all_dissertations/490 (Accessed: 6 August 2019).
25. Kotlarsky, J., Oshri, I. and Willcocks, L. (2018) *Digital Services and Platforms: 12th Global Sourcing Workshop 2018*. La Thuile, Italy.
26. Lambrou, M. (2016) 'Innovation Capability, Knowledge Management and Big Data Technology: A Maritime Business Case', *International Journal of Advanced Corporate Learning (iJAC)*, 9(2), p. 40. doi: 10.3991/ijac.v9i2.6010.
27. Lázaro, F. *et al.* (2019) 'VHF Data Exchange System (VDES): an enabling technology for maritime communications', *CEAS Space Journal*, 11(1), pp. 55–63. Available at: https://link.springer.com/article/10.1007/s12567-018-0214-8.
28. Lee, H. *et al.* (2018) 'A decision support system for vessel speed decision in maritime logistics using weather archive big data', *Computers and Operations Research*. Elsevier Ltd, 98, pp. 330–342. doi: 10.1016/j.cor.2017.06.005.
29. Lensu, M. and Goerlandt, F. (2019) 'Big maritime data for the Baltic Sea with a focus on the winter navigation system', *Marine Policy*. Elsevier Ltd, 104(January), pp. 53–65. doi: 10.1016/j.marpol.2019.02.038.
30. Mahapatra, N. (2019) *The Use of Big Data in Ports and Terminals*. Available at: https://www.academia.edu/27232641/The_Use_of_Big_Data_in_Ports_and_Terminals (Accessed: 6 August 2019).
31. Mammadova, M. and Jabrayilova, Z. (2017) 'Opportunities and Challenges of Big Data Utilization in the Resolution of Human Resource Management', *Problems of Information Technology*, 07(1), pp. 33–40. doi: 10.25045/jpit.v07.i1.05.
32. *Marine Advisory System - OCTOPUS* (2019). Available at: https://new.abb.com/marine/systems-and-solutions/digital/advisory (Accessed: 19 August 2019).
33. MARINET (2019) *e-Navigation testbed and development of equipment for e-Navigation*. Available at: https://marinet.org/e-navigation-testbed-and-development-of-equipment-for-e-navigation/

(Accessed: 6 August 2019).
34. *Maritime Connectivity Platform (MCP)* (2019). Available at: https://maritimeconnectivity.net/ (Accessed: 6 August 2019).
35. Marshall University (2019) *Big Data, Big Data Analytics*. Available at: https://www.marshall.edu/bigdata/ (Accessed: 23 August 2019).
36. Mauro, A. De *et al.* (2014) 'What is Big Data? A Consensual Definition and a Review of Key Research Topics', in *AIP conference proceedings*, pp. 97–104. doi: 10.13140/2.1.2341.5048.
37. MI News Network (2019) *The Long-Range Tracking and Identification (LRIT) System: Tracking and Monitoring Ships*. Available at: https://www.marineinsight.com/maritime-law/the-long-range-tracking-and-identification-lrit-system-tracking-and-monitoring-ships/ (Accessed: 6 August 2019).
38. Mirović, M., Miličević, M. and Obradović, I. (2018) 'Big data in the maritime industry', *Nase More*, 65(1), pp. 56–62. doi: 10.17818/NM/2018/1.8.
39. Nettitude (2019) *The IMO Approach to Maritime Cyber Risk Management*. Available at: https://blog.nettitude.com/the-new-imo-maritime-cyber-security-guidelines-nettitude (Accessed: 9 December 2019).
40. Network Mi News (2019) *Wärtsilä Navi-Planner Lifts Voyage Planning and Optimisation To Unprecedented Levels*. Available at: https://www.marineinsight.com/shipping-news/wartsila-navi-planner-lifts-voyage-planning-and-optimisation-to-unprecedented-levels/ (Accessed: 19 August 2019).
41. Nita, S. L. and Mihailescu, M. I. (2017) 'Importance of Big Data in Maritime Transport', *Scientific Bulletin of Naval Academy*, XX (1), pp. 2–6. doi: 10.21279/1454-864x-17-i1-079.
42. Notteboom, T. (2017) *The Future of Port Logistics: Meeting the Challenges of Supply Chain Integration*. Available at: www.ing.be (Accessed: 6 August 2019).
43. ORBCOMM (2019) *Networks: Satellite AIS*. Available at: https://www.orbcomm.com/en/networks/satellite-ais (Accessed: 6 August 2019).
44. Özköse, H., Ari, E. S. and Gencer, C. (2015) 'Yesterday, Today and Tomorrow of Big Data', in *World Conference on Technology, Innovation and Entrepreneurship: Procedia - Social and Behavioral Sciences 195*, pp. 1042–1050. doi: 10.1016/j.sbspro.2015.06.147.
45. Perera, L. P. and Mo, B. (2016) 'Machine intelligence for energy efficient ships: A big data solution', in *Proceedings of the 3rd International Conference on Maritime Technology and Engineering (MARTECH 2016)*. Lisbon, Portugal, pp. 143–150. doi: 10.1201/b21890-21.
46. Rødseth, Ø. J., Perera, L. P. and Mo, B. (2016) 'Big Data in Shipping - Challenges and Opportunities', in *Proceedings of the 15th International Conference on Computer Applications and Information Technology in the Maritime Industries (COMPIT 2016)*. Lecce, Italy, pp. 361–373.
47. safety4sea (2016) *Bahri relying on Big Data to enhance productivity*. Available at: https://safety4sea.com/bahri-relying-on-big-data-to-enhance-productivity/ (Accessed: 6 August 2019).
48. SAFETY4SEA (2018) *Crisis management explained: The shipping industry case*. Available at: https://safety4sea.com/cm-crisis-management-explained-the-shipping-industry-case/ (Accessed: 13 August 2019).
49. Šakan, D. *et al.* (2018) 'Near real-time S-AIS: Recent developments and implementation possibilities for global maritime stakeholders', *Pomorstvo*, 32(2), pp. 211–218. doi: 10.31217/p.32.2.6.
50. Schmidt, D. (2019) *Weather Routing Software for Sailors*. Available at: https://www.cruisingworld.com/weather-routing-software-for-sailors/ (Accessed: 6 August 2019).
51. *Ship Tracking* (2019). Available at: http://www.freightmetrics.com.au/Marine/tabid/402/Default.aspx (Accessed: 6 August 2019).
52. Sivan, A. P., Johns, J. and Venugopal, J. (2014) 'Big Data Intelligence in Logistics Based on Hadoop And Map Reduce', in *International Journal of Innovative Research in Science, Engineering and Technology*. Madurai, Tamil Nadu, India, pp. 2634–2640.
53. Society for Gas as a Marine Fuel (SGMF) (2017) *Gas as a marine fuel - Safty Guidelines Bunkering FP07-01. Version 2.0*. Available at: https://www.sgmf.info/shop/preview/9 (Accessed: 6 August 2019).
54. *SOLAS Guidance on Chapter V – Safety of Navigation - Annex 17* (2002). Available at: http://solasv.mcga.gov.uk/Annexes/Annex17.htm (Accessed: 6 August 2019).
55. *SOLAS Guidance on Chapter V – Safety of Navigation - Regulation 20* (2002). Available at: http://solasv.mcga.gov.uk/regulations/regulation20.htm (Accessed: 6 August 2019).
56. SVB (2019) *AIS Transceivers*. Available at: https://www.svb24.com/en/category/electronics-gps-

a-i-s-devices-ais-transceivers.html (Accessed: 19 August 2019).
57. Taylor-Sakyi, K. (2016) *Big Data: Understanding Big Data*. Available at: http://arxiv.org/abs/1601.04602 (Accessed: 6 August 2019).
58. Trelleborg Marine Systems (2018) *Use of big data in the maritime industry*. Available at: https://www.patersonsimons.com/wp-content/uploads/2018/06/TMS_SmartPort_InsightBee_Report-to-GUIDE_01.02.18.pdf (Accessed: 6 August 2019).
59. Valčić, S., Žuškin, S. and Brčić, D. (2018) 'VHF Maritime Mobile Band - A New System to Declutter AIS Channels', *Sea technology*, 59(7), pp. 24–27.
60. *Vessel Traffic Service (VTS)* (2019). Available at: https://www.sheltermar.com.br/en/vts/.
61. Voss, S., Sebastian, H.-J. and Pahl, J. (2017) 'Introduction to Intelligent Decision Support and Big Data for Logistics and Supply Chain Management Minitrack', in *Proceedings of the 50th Hawaii International Conference on System Sciences*, pp. 1516–1517. doi: 10.24251/hicss.2017.159.
62. Watson, R., Lind, M. and Haraldson, S. (2017) 'Physical and Digital Innovation in Shipping: Seeding, Standardizing, and Sequencing', in *Proceedings of the 50th Hawaii International Conference on System Sciences (2017)*. doi: 10.24251/hicss.2017.579.
63. *What is big data analytics? Everything you need to know | InfoWorld* (2018). Available at: https://www.infoworld.com/article/3220044/what-is-big-data-analytics-everything-you-need-to-know.html (Accessed: 19 August 2019).
64. Windward (2014) *AIS Data on the High Seas: An Analysis of the Magnitude and Implications of Growing Data Manipulation at Sea*. Available at: http://www.arbitrage-maritime.org/fr/Gazette/G36complement/Windward.pdf (Accessed: 6 August 2019).
65. Zaman, I. *et al.* (2017) 'Challenges and opportunities of big data analytics for upcoming regulations and future transformation of the shipping industry', *Procedia Engineering*. Elsevier B.V., 194, pp. 537–544. doi: 10.1016/j.proeng.2017.08.182.
66. Zghurovsky, M. and Zaychenko, Y. P. (2019) *Big Data: Conceptual Analysis and Applications*. Available at: https://www.springer.com/gp/book/9783030142971.
67. Žuškin, S., Grubišić, N. and Sumner, M. (2015) 'Shipowner management in accordance with mutual agreement', *Pomorstvo: scientific journal of maritime research*, 29(1), pp. 69–74.