

Zaman I, Pazouki K, Norman R, Younessi S, Coleman S.

[Challenges and Opportunities of Big Data Analytics for Upcoming Regulations and Future Transformation of the Shipping Industry.](#)

*Procedia Engineering* 2017, 194, 537-544.

**Copyright:**

© 2017 Published by Elsevier Ltd. by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>). Peer-review under responsibility of the organizing committee of the 10th International Conference on Marine Technology.

**DOI link to article:**

<https://doi.org/10.1016/j.proeng.2017.08.182>

**Date deposited:**

07/11/2017



This work is licensed under a

[Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence](http://creativecommons.org/licenses/by-nc-nd/4.0/)



10th International Conference on Marine Technology, MARTEC 2016

## Challenges and Opportunities of Big Data Analytics for Upcoming Regulations and Future Transformation of the Shipping Industry

Ibna Zaman<sup>a,c,\*</sup>, Kayvan Pazouki<sup>a</sup>, Rose Norman<sup>a</sup>, Shervin Younessi<sup>c</sup>, Shirley Coleman<sup>b</sup><sup>a</sup>*School of Marine Science and Technology, Newcastle University, Newcastle upon Tyne NE1 7RU, United Kingdom*<sup>b</sup>*Industrial Statistics Research Unit, Newcastle University, Newcastle upon Tyne NE1 7RU, United Kingdom*<sup>c</sup>*Royston Ltd, Unit-3 Walker Riverside, Newcastle upon Tyne NE6 3PF, United Kingdom*

### Abstract

Shipping is a heavily regulated industry and responsible for around 3% of global carbon emissions. Global trade is highly dependent on shipping which covers around 90% of commercial demand. Now the industry is expected to navigate through many twists and turns of different situations like upcoming regulations, climate change, energy shortages and technological revolutions. Technological development is apparent across all marine sectors due to the rapid development of sensor technology, IT, automation and robotics. The industry must continue to develop at a rapid pace over the next decade in order to be able to adapt to upcoming regulations and market pressure. Ship intelligence will be the driving force shaping the future of the industry. Ships generate a large volume of data from different sources and in different formats. So big data has become the talk of the industry nowadays. Big data analysis discovers correlations between different measurable or unmeasurable parameters to determine hidden patterns and trends. This analysis will have a significant impact on vessel performance monitoring and provide performance prediction, real-time transparency, and decision-making support to the ship operator. Big data will also bring new opportunities and challenges for the maritime industry. It will increase the capability of performance monitoring, remove human error and increase interdependencies of components. However, the industry will have to face many challenges such as data processing, reliability, and data security. Many regulations rely on ship data including the new EU MRV (Monitoring, Reporting and Verification) regulation to quantify the CO<sub>2</sub> emissions for ships above 5000 gross tonnage. As a result, ship operators will have to monitor and report the verified amount of CO<sub>2</sub> emitted by their vessels on voyages to, from and between EU ports and will also be required to provide information on energy efficiency parameters. The MRV is a data-oriented regulation requiring ship operators to capture and monitor the ship emissions and other related data and although it is a regional regulation at the moment there is scope for the International Maritime Organization (IMO) to implement it globally in the near future.

© 2017 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license

[\(http://creativecommons.org/licenses/by-nc-nd/4.0/\)](http://creativecommons.org/licenses/by-nc-nd/4.0/).

Peer-review under responsibility of the organizing committee of the 10th International Conference on Marine Technology.

**Keywords:** Carbon emission; data-oriented;MRV; big data

### 1. Introduction

The shipping industry is key to global trade and currently operates under a complex set of different international and national regulations. Shipping is a volatile industry and is now in a turbulent condition due to the current poor economic situation, energy price fluctuations, technological immaturity and upcoming increases in regulations. Due to the pace of technological development in the 21<sup>st</sup> century, many emerging technologies have created a significant

\* Corresponding author. Tel.: +44-191-280-6718 ; fax: +44-191-280-5491.

E-mail address: [ibna.zaman@newcastle.ac.uk](mailto:ibna.zaman@newcastle.ac.uk)

impact on the maritime sector. New technology has the power to deliver great benefits to the industry but also introduces direct and indirect threats. Marine applications require the integration of multiple technologies and two technology areas in particular will shape the shipping industry in 2030 and have a significant impact on vessel system design and ship operation [1]. The first technology area originates from within the industry and includes ship building, propulsion systems and the so-called “smart” ship [1]. The second technology area comes from other sectors and it includes sensors, big data analytics, advanced materials and communications [1]. Those two technology groupings are inter-connected and will encourage technological sophistication, operational efficiency and commercial performance in the shipping industry [1]. The use of sensor technologies is rapidly expanding in the shipping industry and will allow real-time monitoring and control of systems and processes. A ship can create a huge amount of data from multiple sensor system and this data is very large in volume and complex to process; big data has now become a burning issue for the industry. Big data analytics can facilitate operations such as the monitoring of emissions and predictive analysis of the vessel performance. Shipping is a highly carbon-efficient mode of transportation but currently has no regulatory system to limit the future growth of greenhouse gas emission [2]. Shipping is an essential sector for EU economy as it provides 2.3 million jobs and contributes € 145 billion to EU GDP [6]. Approximately 4% of EU greenhouse emissions are generated by shipping and this is expected to increase in the future [3]. If left unchecked, shipping-related emissions are expected to increase by 51% in the EU and 200% globally by 2050 [2,3]. The European Commission has therefore proposed the introduction of a monitoring, reporting and verification (MRV) system for CO<sub>2</sub> emissions from vessels greater than or equal to 5000 gross tonnage calling at any EU ports from 2018 [5]. CO<sub>2</sub> emissions and additional data will be collected on a per voyage basis and must be verified by independent parties. This paper discusses features and risks associated with big data, how big data analytics will be turned into added value for the future maritime industry as well as the upcoming data oriented EU MRV regulations.

### Nomenclature

EU	European Union
MRV	Monitoring, reporting and verification
IT	Information technology
CO <sub>2</sub>	Carbon dioxide
GPS	Global Positioning System
KPI	Key performance indicator
GDP	Gross domestic product
Mt	Million Tonnes

## 2. What is Big Data

The term Big Data is used to describe large and complex data sets that are difficult to process and analyse using traditional data processing techniques and applications [1]. It refers to the collection and subsequent analysis of any significant large collection of unstructured data that may contain hidden insights [7]. According to O’ Reilly, “Big data is data that exceeds the processing capacity of conventional database systems. The data is too big, moves too fast, or does not fit the structures of existing database architectures. To gain value from this data, there must be an alternative way to process it” [7]. The growth of big data increases daily and there is no limit in sight. The annual data generation is predicted to increase by 4300% by 2020. [1] Big data analytics is not just about gathering data but also covers the process of analysis to discover hidden insights, trends and correlations. That information has the potential to bring a competitive advantage to any industry.

### 2.1. Big Data Characteristics

Big data covers information from various sources i.e. sensors. There are many difficulties associated with capturing, sorting, analysing and managing data. Fig.1 illustrates the features of big data according to IBM. Big data has four main characteristics and referred to as the 4V’s (Volume, Variety, Velocity and Veracity) [8].

- Volume refers to the massive quantity of the data. Nowadays, sensors produce a massive amount of data in terabytes, petabytes and beyond.
- Variety refers to the form of the data. In big data, the datasets are stored across in multiple formats. Data variation differentiates big data from traditional data.
- Velocity defines the speed of the data creation and movement. The data is created at different rates and must be stored for processing. Generally, a huge amount of data is created in real-time and data flow rates are themselves increasing rapidly.
- Veracity refers to the data accuracy and trustworthiness. Datasets from different sources may use different scales to measure the same variable and this raises issues of how to maintain data quality. Veracity needs to be addressed and maintained throughout the data lifecycle.

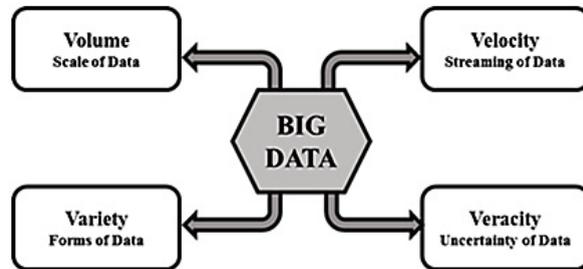


Fig. 1: Four Vs of big data [10].

### 3. Big Challenges of Big Data

Despite the much lauded potential, using big data has brought huge challenges in terms of data acquisition, management, process, storage and analysis. Big data is unlike traditional data in its characteristics of high-volume, high-velocity, high-variety of sources and the requirement to integrate all of it for analysis. Traditional data management and analytical systems are based on relational and structured database systems [7]. Those systems are not designed for the huge volume and heterogeneity of big data. Table 1 shows a comparison of big data with traditional data.

Table 1: Comparison of big data with traditional data[7].

Parameters	Traditional Data	Big Data
Type of Data	Structured	Unstructured
Volume of Data	Terabyte	Petabytes and Exabytes
Architecture	Centralised	Distributed
Relationship between data	Known	Complex

The management and analysis of big data is becoming increasingly important and having a greater impact on a wide range of industries. However, there are many uncertainties and issues associated with big data. Many of these issues are common for all industries but some are associated only with particular sectors. Fig.2 shows the key issues for using big data.

- Data Quality is extremely important. Having a huge volume of data does not necessarily imply a better output. Typically, most data scientists spend 75%-80% of their time cleaning up data [7]. Analysing poor quality data result in misleading information.
- Consistency refers to how the values will be presented and the formatting of the data. It is also related to the form in which the specific sample of data comes from the source.



Fig. 2: The main issues for using big data applications[8,9].

- Data Reliability depends on collection methods and definition of measurements [7]. Data reliability depends on how the raw data was collected. The strength of analysis is related to data quality and reliability.
- Data Availability relates to on-time accessibility. Data needs to be available at any time for analysis. If sensors provide data for a measurement at a particular sample rate then the raw data would be accessible for a specific period of time. Archive or historical data should be available for collection and analysis.
- Linking Data refers to how to connect the data together. Usually, the data is collected for a specific task or purpose and might require different types of data from various datasets. If a single user accesses multiple datasets then a primary key is needed to connect the data together.
- Data Presentation is related to how the data may be viewed to facilitate meaningful interpretation. Poor data presentation could obstruct data analysis and reduce the value of the original data. Data presentation for different datasets has become important for efficient operations.
- Dataset Scalability defines the capability of a system to support any type of dataset. All datasets have their own structure, type, semantics and accessibility. The analytical algorithms applied to big data must have the capability to support increasingly expanding and complex datasets [9].
- Data Compression refers to the filtering of data to reduce the volume. The data generated from sensor networks generally contains a high level of redundancy. That data must be filtered and compressed at orders of magnitude to reduce redundancy.
- Data Life Cycle Management relates to how frequently and which data will be stored or discarded. Generally, the value of big data depends on data freshness [9]. Sensor network systems can generate data at unprecedented rates and scales, so processing and storage become key issues for such massive datasets.
- Data Confidentiality refers to safety and protection of data. Data may be analysed or shared with third parties increasing the potential safety risk. The data needs to be protected from any risks.

#### 4. Challenges for Shipping Industry of Big Data Analytics

The shipping industry generates a huge amount of data from different sources and in different formats, this includes traffic data, cargo data, weather data and machinery data. The volume and variety of data continues to increase day by day due to the application of sensor technology in the industry. The data is generally collected and processed remotely with high transmission rates. So the industry is now getting familiarized with many challenges of big data. Big data analytics are new to the shipping industry and address many issues such as adaptability and integration. Many challenges are associated with big data implementation in the shipping industry:

- *Data Transfer:* Ships typically have a very large number of sensors onboard. A major cause of uncertainty comes from data transfer from those sensors. Every sensor requires a specific communication bandwidth, so it is important to have appropriate data communication for the individual sensor to transmit the information to the database. The data transfer speed may be accelerated with the help of high-tech communication systems.
- *Cybersecurity:* This is a burning issue for any IT system. The safety and security of the data network and data management will become vital for future shipping [1]. This will need to be protected from external interventions such as piracy, viruses or terrorist attacks [1]. Cybersecurity will be the key issue for any naval system to prevent

disruption in maritime security. A cyber-attack on the sensor network would interrupt the overall system and could be responsible for significant losses in the business.

- **Data Quality:** Low-quality data would potentially lead to errors in interpretation. The database will not be able to keep track of all new entries. So ideally the data should be error free. Data quality will be a big concern for the industry.
- **Data Integration:** The current data collection systems in the marine industry are inconsistent and often unreliable [1]. Data from different sources will need to be integrated for analysis. For example, fuel consumption, GPS data and engine data would need to be integrated to monitor the vessel performance.
- **Data Ownership:** Ownership allows access to the data to read, create, update and delete entries in databases as well as allow traceability through the data lifecycle [11]. The shipping industry is based on a complex supply chain; many stakeholders are associated with it including ship owners, operators, customers, port authorities, and Classification Societies. Ship operators will have access to the full set of machine data and Classification Societies will get access to data for safety or classification purposes. Port state authorities will require access to cargo and personnel information. Ownership of data is crucial to the shipping industry and it will become more challenging for ship operators to distribute the data ownership and the level of authority in the future. Fig. 3 illustrates the owners of ship data.



Fig. 3: Ship data ownerships[1].

- **Data Protection:** Data will move between individual parties because of different interests. Sensitive data will probably need to be shared externally making security and privacy priorities for data protection and to maintain the data quality.
- **Adoption and Standard Management:** The industry has to look forward, adopting big data analytics to understand the hidden features and benefits of using the data it has available. The shipping industry will need to create an environment and awareness across the stakeholders to adopt new technologies, tools and processes and also to regulate standards.
- **Human factors and Practice:** It will become more important to increase the connectivity between the crew and shore staff in shipping companies. In the future, the data transfer between a ship and shore and from shore to ship will increase to drive towards optimal operational efficiency and safety. The ship and shore personnel will be required to undertake additional training to provide support for this.
- **Business Model:** The shipping industry is moving towards significant technological change. This will lead to a change in the business model of the industry; a model which will enable the development of a transparent industry associated with the transfer of knowledge and data-driven systems.

#### 4.1. Data-oriented applications for the future shipping industry

Many data-driven concepts have already been introduced in the shipping industry such as the smart ship, the connected ship, autonomous surface and underwater vehicle. Those concepts are based on the implementation of digital technology and automation. Ship intelligence will be the driving force that will direct the future of the industry [12]. Big data has become the buzzword in the maritime sector. The shipping industry has produced a large amount of data which leads to a move to the big data where the analysis and management of this data will become increasingly essential and is expected to have a great impact in the marine industry [13]. The data transmission rate is getting faster meaning that more data collect in the shortest possible time. Different types of sensors and by turning the data into value, there is significant potential to improve ship onboard operations and maintenance. Fig.4 shows different data sources for the maritime industry and the benefits that may arise from analysis of the data. The ship data needs to be integrated for analysis. Data analytics will have the potential to increase the vessel optimisation, asset utilisation and

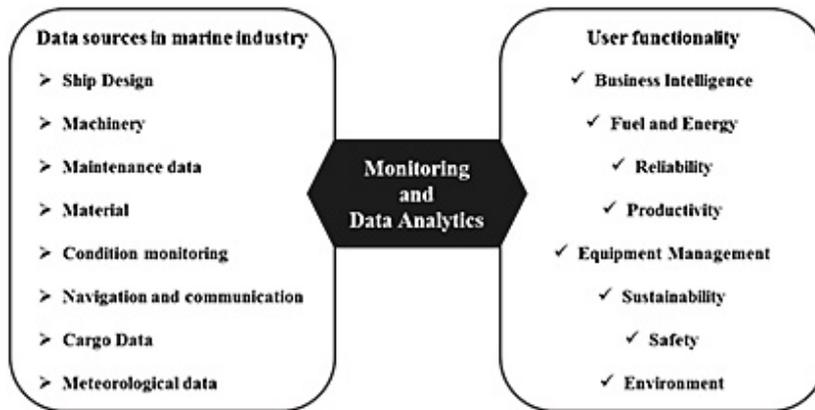


Fig. 4: Ship data sources and user functionality after analysis.

performance. Maintenance, navigation and communications managed by onboard data analytics connected to onboard and onshore decision support systems will enhance the effectiveness of operational scheduling. A future data-driven maritime might include:

- *Remote Sensing:* Ships will be monitored continuously from remote locations. Data will be collected autonomously by using remote sensor networks. A robust wireless network with high transmission capabilities would be required for the shipping industry. The real-time sensor data will come to the database and be distributed to the interested parties giving them up-to-date information on what is happening onboard.
- *Voyage Planning:* Ship operators or charters will be able to implement voyage planning after analysing the route, vessel performance and meteorological data. The voyage planning will be based on the vessel performance on the same and different routes. A reliable forecast of wind and ocean current data will be required for voyage planning. Data analytics will help to identify the most efficient route for the journey, accurately estimated arrival time and alternative routes can be planned to avoid any delay or disturbance.
- *Intelligent Traffic Management:* Port authorities will have access to the ship data for safety. Intelligent traffic management systems will be introduced as data-oriented applications in the shipping industry. Ship current position, cargo and personnel information will be transferred to the port so that the port authorities would be able to monitor the congestion and improve cargo handling performance.
- *Operational Predictability:* Vessel operational performance can be monitored in real-time by analysing ship data. Ship operators will gain the capability to predict the vessel performance based on current operational conditions. This predictability will assist in making decisions on maintenance.
- *Energy Management:* Shipping is moving towards flexible and alternative energy systems. Energy production, storage and re-use will become part of the energy management system; in fact, battery operated vessels have already been introduced in the industry. The ship energy management system will be run on the basis of real-time data of load requirements and power availability from all sources. The system will distribute and balance ship and shore power.
- *Environmental legislation monitoring:* Most emissions are related to fuel type. Ship operators need to comply with environmental legislations including the requirement to switch fuel in emission controlled areas i.e. the Sulphur content of the fuel should not be more than 0.1% on and after 1 January 2015[4]. New data-oriented regulations are coming in to force in the EU to monitor CO<sub>2</sub> emissions and the system will give an indication of fuel switching as well as monitoring the current emissions. The ship operators will also be able to set the KPI for the vessel after analysing the emissions data.
- *Performance Monitoring and Optimisation:* Automation expands the capability of the control of machinery and vessel optimisation. A range of data measurements are required for monitoring vessel performance and optimising performance. Access to historical data is essential for optimising and forecasting. Vessel optimisation and

efficiency will be measured by combined analysis of onboard data and historical data in the current operating condition.

- *Vessel Safety and Security*: The use of wireless sensors and extensive satellite communication systems will increase vessel safety. Sensor data analytics will provide information on vessel maneuvering to avoid collision. The inbuilt sensors on machinery will provide real-time information about their current condition; information that will be useful for crew safety. Vessel safety and security will be increased by the adoption of innovative technology.
- *Condition Monitoring*: This will be improved by analysing asset data. This will be applicable to machinery such as engines, pumps, boilers and compressors. Sensors will help to monitor the machinery and give early warning of the need for maintenance. It will determine the condition of machinery and record data during operation.
- *Predictive Maintenance System*: This will detect the need for maintenance to avoid potential failure. The potential failures would be detectable and measurable. The system will record all data and indicate risk of failure by synchronising related data such as engine data, fuel consumption, running hours etc. It will reduce the cost of asset failures and minimise unscheduled downtime. The system relies on machine health monitoring to dictate when and what maintenance would be required before failure occurs so that the crew would not be required to spend additional time on maintenance and scheduling activities.
- *Automatic Mode Detection System*: This system detects the vessel's operational mode automatically based on sensor data (flow meter, GPS) [14]. Using the auto-mode detection system, the crew would not be required to update the mode every time the ship changed operational state [14]. The vessel operational profile will be created by analysis of real-time data of fuel consumption, distance travelled and speed. The auto-mode detection system will function without human intervention and will provide a summary report of fuel consumption for the individual engines, ship running hours and emissions in different modes. The onboard and onshore members of staff would be able to use this information to measure the vessel operational performance and KPI. This system will assist ship operators in meeting the EU MRV regulation by monitoring the fuel consumption and emissions for different vessel modes.

## 5. Upcoming EU MRV Regulation

Approximately 180 Mt of CO<sub>2</sub> were emitted by European maritime transport activities in 2010 and the emissions are expected to increase in the near future [15]. The European Union regulation 2015/757 on the MRV (monitoring, reporting and verification) of CO<sub>2</sub> emissions from maritime transport was adopted by the European Council and Parliament and entered into force on 1st July 2015[16]. This regulation covers shipboard CO<sub>2</sub> emissions rather than other greenhouse emissions and also requires ships to monitor data on cargo carried and transportation work[15]. The key aim of this regulation is to provide reliable information of CO<sub>2</sub> emissions in the maritime sector to define the emission reduction target for the EU [16]. This regulation applies to all ships greater than 5000 gross tonnage except naval and fishing vessels and ships undertaking more than 300 voyages within the reporting periods [16].

Ship operators will be required to produce monitoring plans to monitor data on a per-voyage and annual basis. Fig.5 shows the regulation timeline and data required for reporting. The monitoring plan, emission reports and the issuance documents of compliance will be accredited by third party verifiers [15]. The verifiers will be independent of the company or operator of the ship concerned and be accredited by a national accreditation body according to European Commission Regulation No.765/2008 [15]. EU MRV will be applicable for all ships regardless of their flag:

- Intra-EU Voyage
- Voyages from the last non-EU port to the first EU port of call (incoming voyages)
- Voyages from an EU port to the next non-EU port of call (outgoing voyages)

Ship operators are required to use one or more of four methods for monitoring: Bunker Delivery Note, tank sounding, flow meters and direct emission measurements [16]. CO<sub>2</sub> emissions from the shipping industry relate to

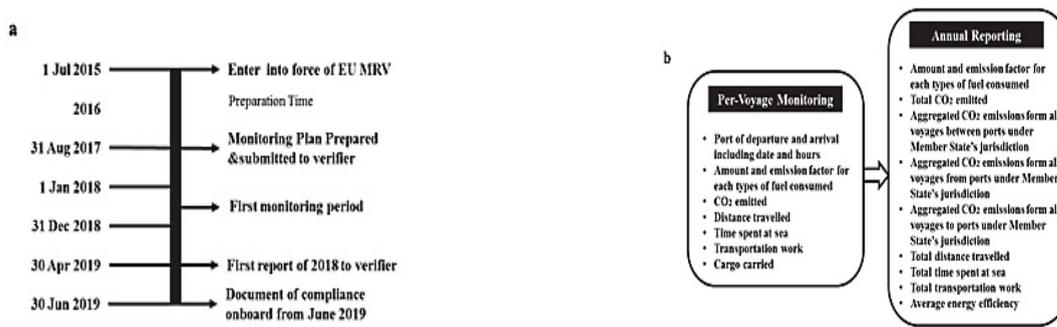


Fig. 5: (a) EU MRV timeline [16] ; (b) required data for MRV regulations [15].

the type of fuel used and the total consumption [4]. CO<sub>2</sub> emissions can be calculated based on fuel consumption and appropriate the carbon emission factor for the fuel type being consumed.

## 6. Conclusions

This paper illustrates how big data analytics have the potential to create a significant impact in the shipping industry. It also covers the associated applications and obstacles to the implementation of big data in the industry. The shipping industry will be faced with ever-increasing requirements for safety, environmental and efficiency performance beyond 2020 [13]. New emerging technologies will be applied for efficient and safety operation. More data-oriented regulations are expected to come into force in the near future. The EU MRV of regulation is expected to reduce the emissions by up to 2% which will lead to a net cost reduction of up to €1.2 billion by 2030 [15]. So it will be an opportunity rather than a threat as this regulation will help the shipping industry to move into the big data era.

## Acknowledgment

The authors would like to acknowledge the support of Royston Ltd, Innovate UK and KTP Project 9738.

## References

- [1] Lloyds Register, QinetiQ, University of Southampton, Global shipping technology 2030, UK, 2015.
- [2] Global Shippers Forum, Maritime Emissions, UK, 2014.
- [3] European Commission, Integrating maritime transport emissions in the EU's greenhouse gas reduction policies, Brussels, 2013.
- [4] International Maritime Organisation, Third IMO GHG study 2014, London, 2015.
- [5] DNV-GL, Maritime EU-MRV Regulation, Hamburg, 2015.
- [6] Oxford Economics, The economic value of the EU shipping industry (2014) 6.
- [7] A.E. Hassaniien, A.T. Azar, V. Snasel, J. Kacprzyk, J. H. Abawajy, Big Data in Complex Systems-Challenges and Opportunities, Springer, London, 9 (2015).
- [8] M. Bulger, G. Taylor, R. Schroeder, Data-Driven Business Models: Challenges and Opportunities of Big Data, University of Oxford (2014)19-50.
- [9] S. Mao, Y. Zhang, V. M. Leung, Big Data Related Technologies-Challenges and Future Prospects, Springer, London, 2014.
- [10] IBM, The Four V's of Big Data, 2016.
- [11] ISO, Big Data-Preliminary Report 2014, Switzerland (2015) 10-13.
- [12] C. Fields, Risk Futures: Smart Ships, 2015.
- [13] DNV, Shipping 2020, Norway, 2014.
- [14] I. Zaman, K. Pazouki, S Coleman, R. Norman, S. Younessi, Performance Monitoring And Automatic Operating Mode Detection Based On The Statistical Analysis Of Ship Data, RINA Energy Efficient Ship Conference, Rotterdam, 2015.
- [15] European Commission, Proposal for a regulation of the European parliament and of the council on the monitoring, reporting and verification of carbon dioxide emissions from maritime transport and amending Regulation (EU) No 525/2013, Brussels, 2013.
- [16] Lloyds Register, European Union Regulation on Monitoring, Reporting and Verification of Carbon Dioxide from Ships, Southampton, 2015.