Linnæus University
Sweden

Master Thesis

Big data usage in the Maritime industry

A Qualitative Study for the use of Port State Control (PSC) inspection data by shipping professionals.

Author: Dimitrios Ampatzidis
Supervisor: Konstantina Pentarchou
Examiner: Anita Mirijamdotter / Päivi Jokela
Date: 2021-08-05
Course Code: 4IK50E, 15 credits
Subject: Information Systems
Level: Graduate
Department of Informatics

Abstract

Vessels during their calls on ports is possible to have an inspection from the local Port State Control (PSC) authorities regarding their implementation of International Maritime Organization guidelines for safety and security. This qualitative study focuses on how shipping professionals understand and use Big Data in the PSC inspection databases, what characteristics they recognize these data should have, what value they attach to those big data, and how they use them to support the decision-making process within their organizations. This study conducted interviews with shipping professionals, collected their perspectives, and analyzed their sayings with Thematic Analysis to reach the study's outcome. Many researchers have been discussed Big Data characteristics and the value an organization or a researcher could have from Big Data and Analytics. However, there is no universally accepted theory regarding Big Data characteristics and the value for the database users. The research concluded that Big Data from the PSC inspections procedures provide valid and helpful information that broaden professionals' understanding of inspection control and safety need, through this, it is possible to upscale their internal operations and their decision-making procedures as long as these data are characterized by volume, velocity, veracity, and complexity.

Keywords

Big Data, Data Analytics, Port State Control, Shipping, Maritime, Information Systems, Big Data characteristics, Big Data and Analytics value, Qualitative study, Interviews, Shipping professionals.

Acknowledgments

I want to thank each of them who helped me throughout my studies and completing this thesis. Of course, I would like to dedicate a special mention to my supervisor Mrs. Konstantina Pentarchou, who throughout the course guided me academically and constructively in the approach of the work and the issues that arose; as well as the teachers of the program who, through their encouragement and guidance, opened new academic horizons. Also, the contribution of the six respondents from the field of shipping was decisive, and I sincerely want to thank them both for the time they gave me and their willingness to contribute to this work. It would be an omission not to mention my colleagues in the job who showed understanding and support throughout the postgraduate course. Finally, of course, I would like to thank my people, family, and friends, for their moral and emotional support throughout the dissertation; without them, this "trip" would have been unsuccessful.

Contents

List of Tables

List of Figures

Abbreviations

BDA	Big Data and Analytics
BPM	Business Process Management
DPA	Designated Person Ashore
HSQE	Health, Safety, Quality, and Environment
IMO	International Maritime Organization
ISM	International Safety Management
KM	Knowledge Management
KPIs	Key Performance Indicators
LNG	Liquefied Natural Gas
MoU	Memorandum of Understanding
MRM	Management Review Meetings
P&I	Protection and Indemnity
PSC	Port State Control
RQ	Research Question
SMS	Safety Management System
TA	Thematic Analysis
USCG	United States Coast Guard

1. INTRODUCTION

This introduction aims to illustrate the topic, the purpose statement, and the research questions. Moreover, it will build the study's reasoning for the previous studies conducted around this research and find the existing knowledge gaps. For that reason, the deficiencies model will be followed in this chapter, as it is introduced by Creswell and Creswell (2018, p.105). The reasoning process in the following lines would be deductive, following Aristotle's division of two reasoning methods. Following this procedure, it will be feasible for the reader to fully understand the need for this study and the reasoning's escalation.

1.1 Introduction and Research Setting

Due to technology's expansion recently and the entrance of Information Systems in more and more everyday tasks, humanity is in front of a booming amount of data collected daily. Of course, this escalation of the data amount is based on the enhanced data storage possibilities because of the technology's new standards; thus, Wilkes et al. (2008) estimated that between 2005 and 2010, the total amount of disk storage would be almost fourfold. This expanded data storage and the general technology evolution helped organizations and companies to build Information Systems infrastructure and assist data-driven decisions. Statistical modeling and algorithms combined with more data in the maritime industry could benefit the industry because of the recent technology expansion. For example, we could find various applications, as a Bayesian Network model used on shipping accidents (Zhang and Thai, 2016). Kowalska and Peel (2012) introduced a model for detecting the unnormal behavior of vessels with Gaussian Processes. While a few years later, Wang et al. (2018) applied statistical models to find the relationship between fuel consumption rate in alternatives, and previously, Vanek et al. (2013) introduced an application for a data-driven decision-support tool for piracy. Therefore it seems that shipping companies and the maritime industry have taken some steps during the last decade in the direction to use more data in their daily procedures and the fleet's operation.

Big Data provides intelligence to professionals and creates value for businesses. According to Manyika et al. (2011) through McKinsey Global Institute, Big Data may create value in five ways; by providing transparency, enabling experimentation to discover the organization's needs or business management, identifying the industry's "environment", and promoting new business models, concluding that big data support human decision-making. Management in any large organization uses decision support systems (DSS) that require a significant amount of data to support their decision-making process (Sharda et al., 2014, p.38). However, data could not provide value to businesses and organizations independently; they require human knowledge and understanding. Thus, data analytics techniques have been developed. Broadly, the term data analytics refers to statistical algorithms for identifying patterns in the data to translate into business intelligence (Beynon-Davies, 2020, p.155) in every organization. It is understood that systems (DSS) provide their estimations, and the data used in these systems could be described as "Big Data" regarding their characteristics that vary in the bibliography.

Data are used in various daily procedures in the shipping industry. Many databases worldwide gather and disseminate freely all the data regarding fleets' trade, voyages, repairs, insurance, crewing, financial issues, or safety controls. Maritime organizations benefit from the big data provided by such databases by transforming them into business intelligence and supporting their decision-making progress like every other organization (Sharda et al., 2014, pp.57-58). In shipping companies, specific procedures must be undertaken to ensure that their fleets operate according to the safety and security legislation (International Maritime Organization, n.d.). Regarding safety and security control, several databases provide data about the port inspection controls and how the vessels of every shipping company adapt to the safety

and security legislation introduced by the International Maritime Organization. One of these databases is Paris MoU which in 2018 stopped providing all the data about Port State Control (PSC) inspections and their findings (RightShip, 2018) and instead of that produced an annual report on the critical items of the past year only for the detainable items that are not more than approximately 3% of the total cases annually (SAFETY4SEA, 2019).

Combining the fewer and fewer data published with the fast-growing legislation about pollution control and measures to protect the environment, shipping companies strive to adapt to the standards that the International Maritime Organization (IMO) establishes. Additionally, each company cannot access the data from other companies' vessel inspections from other public sources to inform about the trends and better prepare their fleet. Only a few details are open freely to all stakeholders, and with the parameter, the shipping professional searches specifically for a vessel with its identification number or by name in public databases like Equasis.

In more detail, whenever vessels visited ports globally, the local authorities may or may not be inspected for safety and security reasons. These inspections concluded on some findings based on the vessel's condition, documentation, and the crew's performance on drills or operations. The findings are subjective as are coming from human - so the human error existed, and each regime has a different focus. For transparency reasons, the main MoU's decided to publish the PSC inspections' outcomes on public databases in datasets containing all the information about the vessel and the inspection. As a result, shipping professionals are trying to use these databases and extract intelligence either by informing specific details or adjusting big data mining techniques to extract knowledge and help them in the decision-making.

This research will focus on the big data characteristics and the value extracted from public PSC databases usage from the responsible personnel in shipping companies about the vessels' safety and security. This research's outcomes could provide valid and valuable information and intelligence around the topic of Big Data usage, especially in the shipping industry.

1.2 Purpose Statement and Research Questions

The purpose of this qualitative study is to explore the shipping professionals' point of view on PSC inspection data and how they work with the provided information to extract intelligence and value from these data, enhancing the decision-making procedure in the shipping companies. Therefore, this study investigates the characteristics PSC inspection data should have from shipping professionals' perspectives regarding companies' internal use to better prepare fleets for safety and security reasons. Furthermore, the research tries to approach the topic from the professionals' perspective to understand better how raw data from these public PSC inspection databases could have a strategic value for corporations and organizations in the shipping environment. For the above reasons, the following three research questions are established for this study.

RQ1: *What value have the port state control (PSC) inspection data for a shipping company?*

RQ2: *How port state control (PSC) data are used for decision-making?*

RQ3: *What characteristics port state control (PSC) data should have?*

1.3 Topic Justification

In general, it is feasible that Big Data and Analytics could have a positive output in the shipping industry as well; in the last few years, many applications built to help the shipping companies to work more efficiently with the people working onboard and ashore (Varelas and Plitsos, 2020). However, it is necessary to clarify that little research has been done on providing access to data and information in shipping, as the challenge with the lack of Paris MoU data arose in the last few months (RightShip, 2018). Most of the research focused on how big data tools are used in shipping or provided statistical analysis of the ports and regimes' data (Ghalehkhondabi, Ahmadi, and Maihami, 2020; Marine Digital, n.d.). Both companies and the community need to continue the fleets' eco-friendly and human-friendly operation to follow the rules and legislation about safety and security. However, the research on this topic regarding the provided assistance from PSC databases to decision-makers is limited. Finally, existing research does not explicitly examine how shipping executives use that kind of data and information.

Since December 2018, when the Paris Memorandum of Understanding (MoU) alters the volume of the published data for the inspections into its ports, a few things have changed, and the majority of the industry tries to adjust to the new reality (RightShip, 2018). The main issues raised focused on the need for these data for the shipping industry and enhancing the safety culture. There would be an academic and business interest in determining whether such business analytics tools could help the shipping industry and how important it is for the maritime sector to share knowledge and information regarding the PSC inspection outcomes. In this way, it could also be approached whether the equal access of data by PSC inspections would help both the companies that operate the global fleet and society increase safety and security in shipping. Of course, it is interesting to look for the people who handle this information in each company and the decision-making process. All the recent studies around this topic have discussed the importance of these data for mining and analytics reasons (Tsou, 2018) but with no concentration on how shipping professionals and companies handle such databases. This elimination builds a gap in the knowledge; it would be beneficial for both communities – academic and business – to explore the value of these data to vessels' everyday operation and how, by gaining PSC data, safety culture could be enhanced.

This study could build knowledge on how shipping professionals benefit from technological evolution developments and use data from the public databases regarding the vessels' inspections and what characteristics these data should have. Furthermore, the outcomes from this research could be a trigger for more research on the usage of big data in other industries and settings.

1.4 Scope and Limitation

It is crucial to highlight a gap that we see in the existing literature because of the non-examined assumptions or fields that are not explicitly researched on the Big Data usage in the maritime sector and especially which is their value and what characteristics should they have in order to fulfill the needs of the shipping companies. Therefore, it would be of academic interest to approach professionals handling knowledge from the PSC, which could be more focused on their needs to upgrade the safety and security services. However, it is critical to answering some fundamental questions regarding PSC databases and how these structures assist shipping professionals in decision-making and provide knowledge for the entire community. In the following chapter and on the continuation of this study, it would be attempted to research

the characteristics these public databases should have to provide business value to the industry and assist the data-driven decisions within shipping corporations.

The purpose of this qualitative study will be to understand the value for shipping companies of the PSC inspection outcomes data provided by the public platforms. We differ the public from the private Big Data platforms regarding the PSC inspections as there are differences in the structure and, of course, on the motive that exists; as it is rational, the private platforms have a commercial interest, which may alter some characteristics. By the "value," we referred to the meaning that has been given to value on big data by Manyika et al. (2011); thus, the potential of big data to create value could be described only by those actions that are depended on using Big Data to have an outcome (Manyika et al., 2011). The strategic value of Big Data and Analytics (BDA) in businesses is mainly divided into two elements, symbolic and functional (Grover, 2018). Grover implies that symbolic is referred to the value an organization may extract from the positive brand name, while the functional is about the market share and the financial performance. Instead of financial performance or market share, this study would explore the strategic functional value of Big Data in the maritime industry. This research would explore PSC data use in shipping companies with the word value in mind as introduced above.

As it is referred, the research's scope is to explore the data usage from the shipping professionals and their insights for how the published data could be helpful in their decision-making procedure and in general for the industry.

1.5 Thesis Organization

The research structure is crucial for the study's continuation that explains and discusses the procedures; Figure 1 illustrates the connections between the study's chapters.

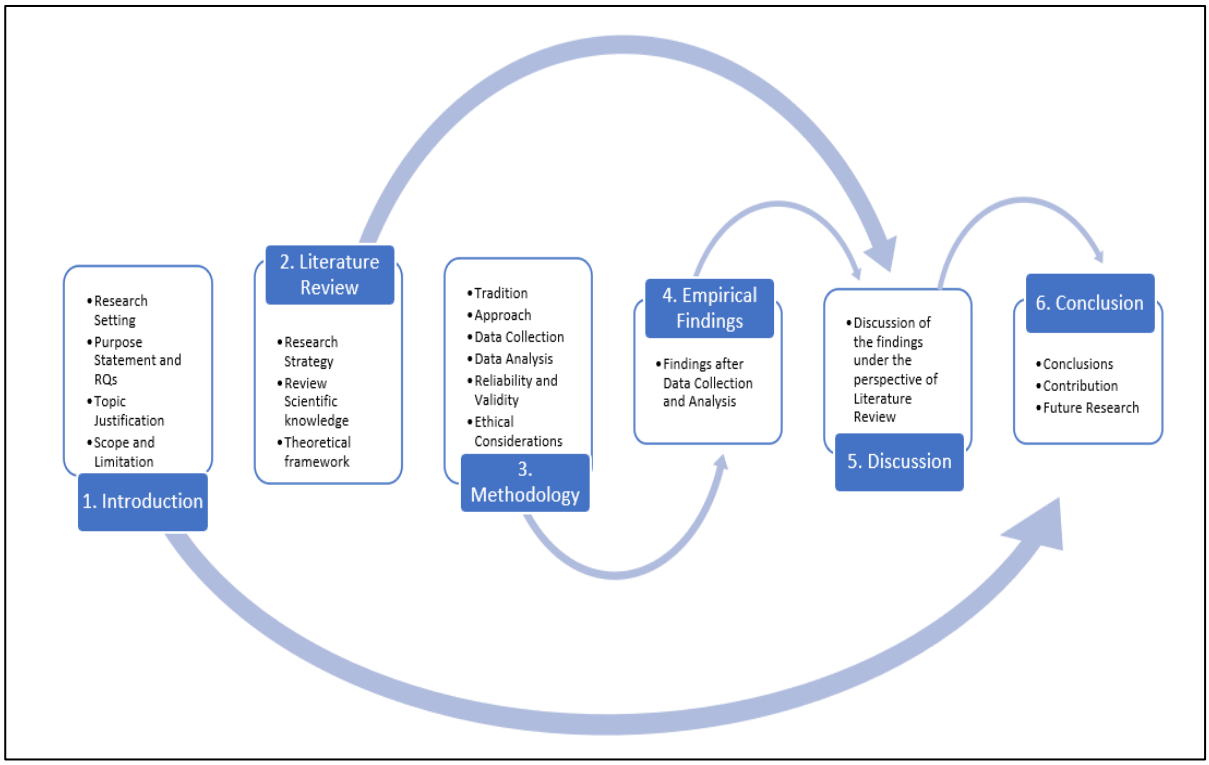

***Figure 1**. Thesis Organization (Source: Created by the author)*

The "Introduction" is critical to establish the topic with the relevant reference on the published bibliography, both academic and business; the first chapter will be connected with the last one,

"Conclusion", in which the outcomes of the empirical findings and the discussion would be provided. The second chapter, "Literature Review", discusses the bibliography around the topic divided mainly into two sub-subjects, Big Data and Shipping Analytics; this chapter provides all the fundamental knowledge for continuing the research. The second chapter is connected with the "Discussion", in which the findings of the research would be analyzed under the perspective of the given theories. Finally, the two median chapters of "Methodology" and "Empirical Findings" concentrated on the procedures that should be followed for the data collection – analysis and presenting the previously collected data.

2. LITERATURE REVIEW

To approach the topic holistically in this study, we first have to address the definitions from the general to the specific; that means through this first chapter of the study, we will discuss big data, and in a second time, we will present the challenges regarding the big data in shipping. The literature review is essential before the research; through this review, it is possible to inform both the study and the reader of the analysis of the results of prior conducted studies, or even use it as a benchmarking tool for the outcomes of the present research (Cooper, 2010; Marshall & Rossman, 2016, cited in Creswell & Creswell, 2018, pp.25-26).

The literature review bridges those topics in the chosen topic, about the value of the PSC data in the shipping industry and the companies' usage. In the following subchapters, the theoretical approach is described as well; in that way, it is feasible to have the study's general direction.

2.1 Search Strategy

Before proceeding to the literature analysis around the subject, it is vital to have a clear plan and approach. As the University of Southern California (2021) refers to the preliminary literature review approaches: argumentative, integrative, historical, methodological, systematic, and theoretical.

The present study's initial aim is through the literature to lay the theoretical foundations on which we seek our answers as to whether the shipping industry needs data from the PSCs and then whether such information is available to the public. Therefore, our bibliographic analysis and approach fall into the category of integrative review. The synthesis of bibliographic reports that refer to or approach our topic to create the necessary background to analyze; the primary goal of the literature review was to find all the possible gaps in the knowledge.

Having defined the general approach is essential, as well, to clearly state the implied strategy to come in contact with all the critical bibliography around the topic. First of all, the search's central elements were the bibliography about the PSC inspection value in both companies and community, how the information is used in companies for preparation reasons, and the business value of data analytics tools to the maritime sector. The titles' search was performed in Linnaeus University's search engine (OneSearch), with the following restrictions: firstly, only peer-reviewed articles were accepted in English; secondly, only full-text titles from 2001 till 2021; thirdly, accepted subjects would be "ships", "ports", "shipping", "shipping industry", "port state control", and "maritime industry"; and finally, the searching phrase was "port state control," and it was required anywhere in the record. The findings were evaluated per title relevance, abstracts, and a more detailed examination of the content. Moreover, the same procedure was followed for the Big Data terms and descriptions; thus, we had an overview of the bibliography around the topic.

The academic bibliography significantly added fundamental knowledge to how data provide value after its usage from industry experts with this study. Many surveys and studies focused on finding the best model to describe the PSC data or extract value from them, but few of those studies concentrated on how these models could help the industry operate with more safety and security for the people and the environment. Investigating with this study, at first, the value and characteristics of the PSC data and then, secondly, collecting the information from responsible people in the industry. This procedure would evaluate the business value from PSC inspections data to the industry and the shipping professionals looking to enhance the value derived from these databases.

2.2 Review scientific knowledge

During this subchapter, the goal is to critically discuss the academic bibliography regarding Big Data and guide the reader on the fundamental elements of this topic; the approach is to divide the literature review into four sections. The first one refers to Big Data's definitions and characteristics so that the reader would contact the basic knowledge of Big Data. The second one is how Big Data could be stored and extract information from these databases. Thirdly, the business organizations' usage and outcome are presented, while the last element refers to Big Data and Analytics in the shipping industry.

2.2.1 Big Data definitions and characteristics

Before exploring the fundamentals elements of the term Big Data, it is critical to identify data's meaning. Data mainly are just symbols that communicate a substance or a piece of information (Beynon-Davies, 2020, p.61), or differently, data that cannot be described by a unit alone (Sharda et al., 2014, pp.57-58; Chen and Zhang, 2014); thus, as it is understood, data could be everything that is used to acknowledge a statement by a human or an information system (IS). The different languages globally are different codings to express the same information, so data coding is essential to transfer the information. Data structures are within data systems; thus, data systems are not static, although they change their elements frequently (Beynon-Davies, 2020, p.70). As mentioned by Gregory Bateson (1972, cited in Beynon-Davies, 2020, p.75), information could be characterized from the difference it makes to something established, highlighting data systems' purpose to provide assistant to the decision-making process.

Big Data was first introduced in the academic and business literature by John Mashey in the mid-1990s (Kitchin and McArdle, 2016), and a few years later, Doug Laney (2001) proposed the main three elements that characterized Big Data; these elements are volume, velocity, and variety. Firstly, data volume referred to the databases' size, as it is understood that it is impossible to have Big Data with just a few data items within. Secondly, the data should vary on their sources, so the database's information would be valid. Thirdly, the data's velocity, the speed of updating the database, and the required information extract.

The word "big" in Big Data defines a volume; existing data is in petabytes and is expected to increase to zettabytes shortly. Social media networking generates the same terabytes of data daily, and that the amount of data is difficult to handle using existing traditional systems. A survey conducted by IBM in 2012 presented that more than half of the approximately one thousand respondents considered data more than one terabyte as big data (Schroeck et al., 2012). Beaver et al. (2010) highlighted in their report that Facebook processes up to one million photos per second. According to previous estimates, Facebook saves 260 billion images using storage over 20 petabytes.

Velocity in Big Data is an idea that deals with the speed data from various sources inserted in the structured database; this feature is not limited to the speed of the incoming data but also to the speed at which they flow the data. The proliferation of digital devices drove the industry to a continually growing data creation rate, which underlined the need for real-time analysis of these data. Wal-Mart, for example, processes more than a million transactions per hour (Cukier, 2010).

Variety refers to structural heterogeneity in a data set. Technological developments allow companies to use different types of data based on how structured these databases are. Structured data, which is just 5% of all existing data (Cukier, 2010), refers to tables in spreadsheets or relational databases. The text, images, audio, and video are examples of unstructured data,

which sometimes lacks the structural organization machines require for analysis. At that point, it is common to be enabled IT professionals to "clean" the data and provide the appropriate format to the dataset. The semi-structured data format, covering a continuum between fully structured and unstructured data, does not comply with strict standards.

Following Laney's three Vs. thinking, many researchers tried to describe the critical elements that characterize Big Data; Kitchin (2014, cited in Beynon Davies, 2020, p.154) mentioned data resolution, data relatability (Boyd and Crawford, 2012, cited in Kitchin and McArdle, 2016), data flexibility, data indexicality, and data exhaustivity (Mayer-Schonberger and Cukier, 2013, cited in Kitchin and McArdle, 2016).

In addition to the three Vs., other Big Data dimensions have been reported. IBM (Lukoianova and Rubin, 2014) suggested Veracity as the fourth V, representing the unreliability inherent in specific data sources. SAS (n.d.) introduced variability and complexity as two additional dimensions of big data; variability refers to data flow rates, while complexity refers to the fact that big data is generated through various sources. Finally, Oracle (2019) presented "value" as the defining feature of big data. The data received in the original form usually have a low value concerning their volume, and that is why Big Data and Analytics techniques should be implemented to extract intelligence from large databases. However, the value could be achieved by analyzing large volumes of such data. Other definitions of Big Data suggest that all data available in any organization could be described as big data and large data sets compared to typical computers' memory (Grandinetti, Joubert, and Kunze, 2015, p.6).

Although many researchers, institutes, and organizations made attempts to build the gaps in the knowledge and provide definitions and descriptions of Big Data, three Vs. theory by Laney enjoys such a reputation in the scientific community by offering a simple and inclusive framework that seems to be the appropriate theory for studying the characteristics of Big Data in the maritime industry. Moreover, shipping professionals with PSC data usually come from the maritime industry without an IT background, so the three Vs. theory would be closer.

Given the possibilities that data storage and management applications offer because of the recent technological expansion, every business could develop capabilities to get as many benefits as possible by improving operations, maximizing asset performance, and controlling maintenance costs (Geng, 2015, p. 619). The International Data Corporation (Needham, 2021) predicted that the big data technology and artificial intelligence services would grow at an annual rate growth of 16.4% for 2021, with yearly expenditures reaching 327.5 billion dollars.

2.2.2 Big data storage, structure, and data analytics

Technological developments have transformed organizations into more data-driven structures in the last few years. Thus, institutes or business organizations maintain in-office large amounts of data for various reasons no matter the subject of their interest; indeed, the storage capabilities have been enhanced recently and the transactions within these datasets through recent technological upgrades. Organizations store data either for legal and documentation reasons or analyze the patterns with revenues, sales, marketing, and general departments with the departments authorized to analyze the business environment.

In Informatics, data structures refer to different ways of organizing and storing data within a computer to be used efficiently. For example, a data set can be stored in a tree structure, table, stack, linked list, pile, or tail. As coming naturally, the extensive volume of big data could be connected with specific infrastructures to help people analyze such databases and export intelligence from them. Different data structures are suitable for different species applications, and some have such a degree of specialization used in special-purpose applications. When we refer to big data, we get away with it from the limits of a computer, and we look for ways of

organization, storage, and management that are both efficient and inexpensive. Such data are collected by various sources such as social networks, network sensors, scientific applications, web texts, and documents with many characteristics, including large size, heterogeneous structures, and complex processing (Grandinetti, Joubert, and Kunze, 2015).

The most popular way to manage large volumes of data is distributed databases (DDB), referred to by Shakhovska et al. (2019). A distributed system can be defined as "a kind of decentralized and distributed architecture network" (Steen and Tanenbaum, 2016). Independent nodes operate simultaneously as providers and resource consumers. In contrast to centralized client-server models, the client node requests access to resources provided by central servers. In a distributed network, the tasks (such as file search or audio/video streaming) are shared between multiple interconnected nodes, each of which offers a portion of the resources (processing power, storage space, network bandwidth) to other nodes, without intermediaries central coordination by servers (Steen and Tanenbaum, 2016). Covering the continually growing storage space requirements and the need for big data processing, new large data platforms are emerging; for example, NoSQL databases arose as an alternative to traditional relational databases, and, on the other hand, Hadoop is introduced as an open-source framework for cheap, distributed software arrays.

Data storage platforms cannot provide intelligence regarding the business issues as they only contain the information items; thus, data analytics techniques should be implemented to extract the professionals' knowledge. Data analytics are distributed into three main elements: descriptive analytics, predictive analytics, and prescriptive analytics (Sharda et al., 2014, pp.60-61; Beynon-Davies, 2020, p.156); each one is using in the business or research entities for different reasons purposes. As a result, data analytics procedures are based on the algorithms used and what we want to investigate. The outcome is weighted for its value; using the "value" term, we describe the output's worth for the researcher or the organization (Beynon-Davies, 2020, p.169). Thus, different data, especially Big Data structures, demand extensive and well-designed storage capabilities and the appropriate data analytics methods for organizations to extract information and intelligence.

The outcome from Big Data and Analytics that companies and institutes enjoy is weighted for its value; as the term "value" is introduced by Beynon-Davies (2020, p.169), the output's worth is for the researcher or the organization. However, organizations are still faced with many challenges regarding using databases on how it would be possible to extract value from stakeholders (Nalchigar and Yu, 2020).

2.2.3 Big Data usage and Outcomes for the business organizations

Palem (2014) provided academic and business literature a roadmap about Big Data and Analytics (BDA) objectives, dividing them into three principal elements; short-term, medium-term, and long-term goals. Short-term objectives aim to define the technology base and build a solid customer base for sustainable revenue generation; medium-term support the functions to maintain the customer base and strengthen their bond with high-quality results; and lastly, long-term goals towards market leadership with innovative solutions strategic partnerships.

For the reasons mentioned above, the information and insights gained from the customers are valid for the business organizations; they could enjoy the strategic value from Big Data, as Qi et al. (2016) mentioned. Palem (2014) introduced a strategy tool known as the Delta Model to develop a customer-centric approach to strategic management with Big Data analytics. The model fits both established organizations and startups because this model targets to connect the business infrastructure with the clients successfully, not to win the competition. While a few years later, Qi et al. (2016) proposed an automatic filtering model to predict online reviews'

usefulness from a product designer's perspective. Based on the classic model combination analysis, the KANO method analyzed online reviews and developed appropriate product improvement strategies. Lastly, Wilberg et al. (2017) developed a model that supports companies developing a data strategy for product development. Besides, with this model, data can support product development by understanding the customers' needs and conditions exposed during their use.

It is feasible that the strategic value to business and corporations could come by understanding in depth the needs of the market as along with the industry's environment; indeed, BDA could allow the professionals to combine their experience and mindset with better knowledge of the patterns that have been found through advanced statistical methods in data analytics operations.

BDA already envisions business operations, reducing the time spent by the professionals to understand the factors hiding behind various business phenomena; thus, it is possible to extract strategic value from using data analytics techniques through this understanding. Dutta and Bose (2015) developed a new framework that provides organizations a holistic map of conception, design and successful implementation of Big Data projects. The proposed framework is divided into three components: integrated exploration process, exploitation of Big Data resources, and identification and measurement of creating business value. Bertei et al. (2015) used a theoretical model incorporating Big Data technology and business process management (BPM) in strategic decision making. While, Elgendy and Elragal (2016) contributed to further theoretical progress of how Big Data technologies help companies improve their finances and environmental performance; using a scientific methodology "Big - Data, Analytics and Decisions" (B-DAD) framework to map the tools, architectures, and analysis of the Big Data during various decision-making phases and provide a general model. It is mainly illustrated that the BDA focuses on how these techniques could be implemented within organizations and companies.

Additionally, many academic or business researchers are trying to find the appropriate way to efficiently analyze Big Data to add "value" or a better meaning to the existing infrastructures. Besides, their research targeted upscale the decision-making procedures within companies using Big Data techniques. Mazzei and Noble (2017) introduced a procedure that recognizes how Big Data improves the organizations' functional possibilities; they shape new industries and are a crucial component of innovative strategies that businesses use to differentiate and jump over barriers in traditional markets. At the same time, Lakoju and Serrano (2017) developed a Big Data strategy framework that helps organizations align their business strategy with Big Data projects to determine their potential value before full implementation. Intezari and Gressel (2017) propose a decision-making typology based on data; they developed a theoretical framework for how systems work knowledge management (K.M.) can integrate Big Data into strategic decisions. A multidimensional framework for creating value through information technology and Big Data scenarios developed by Elia et al. (2019); the framework mentioned previously includes eleven different value directions grouped into five dimensions: information, Transaction, Transformation, Strategy, and Infrastructure.

A summary of Big Data models is provided from the previously conducted research and the accessible bibliography. Firstly, studies and implemented Big Data paradigms highlight examples of already applications in organizations. Secondly, the general theories about Big Data's characteristics and the strategic value they could create would be discussed. However, this study must explore the theoretical base of the previously mentioned paradigms of how the implementation of Big Data could upscale the operations of any business organization.

Some studies show that one of the main challenges for correctly imply business analytics tools and applications is a misunderstanding between the data scientists and the stakeholders (Lavalle et al., 2010a). It is controversial for companies and institutions to effectively imply business analytics, as shown from an extensive survey highlighting that top-performing

companies use analytics more extensively than lower performances (Lavalle et al., 2010b). Thus, Big Data and Analytics (BDA) could provide a competitive advantage to the company as it could be shown all the patterns and connections behind the data (Grover et al., 2018). At last, it could be concluded that Big Data may create value in five ways, as McKinsey Global Institute through Manyika et al. (2011) implies, are by creating transparency, enabling experimentation to discover needs, segmenting populations to customize actions, supporting human decision making with automated algorithms, and innovating new business models.

From the analysis of McKinsey Global Institute (Manyika et al., 2011, p.4) is feasible to find that Big Data can create value in five broadly applicable ways; by creating transparency, enabling experimentation, segmenting populations for marketing or risk-management reasons, supporting the human decision, and innovating new business models. However, the question raised is how Big Data's value could be measured. Manyika et al. (2011, p.7) could guide us by clarifying the requirements for having a value from Big Data; indeed, McKinsey's study mentioned that only actions depended on Big Data utility could be counted as a liable measurement for Big Data's potential to create value and worth. Additionally, the division of the value mentioned by Grover et al. (2018) into functional and symbolic has importance to the perspective this research would examine the meaning of value into shipping organizations; thus, the study would focus on the functional value that Big Data could bring by analyzing the data patterns to make a competitive advantage. As a result, this research will focus on how Big Data could create functional value to shipping organizations with the five ways mentioned by McKinsey Global Institute; thus, this framework could give a holistic approach to Big Data's value meaning.

Manyika et al. (2011, p.59) from McKinsey state that it is possible to analyze large datasets if the researcher or the professional uses automated algorithms on Big Data datasets to assist decision-makers for more data-driven and correct decisions. Indeed, as mentioned above, for Big Data's value, the scope of having and analyzing large datasets is to access knowledge that is not accessible through human experience and understanding of the organization's environment. Big data analysis can support the human procedures for decision-making by highlighting hidden correlations and patterns in data or even underlining some hidden risks (McKinsey Global Institute, 2011, cited in Kubina et al., 2015).

It has an educational value to examine the five factors that McKinsey Global Institute (Manyika et al., 2011) presented that Big Data could create value in organizations.

- **Transparency:** Having Big Data easily accessible to stakeholders among the industries could create value in many sectors. For example, in both the public and private sectors, the separated departments' ability to have all the required information fast could enhance their operations; the better outcome could be connected with better revenues or more wise spending recourses for the organization.
- **Experimentation:** Big Data and Analytics could make it possible to collect with accuracy and detail performance data about employees, customers, or logistics, so with accurate data, the IT department or a data scientist could apply statistical methods to examine the patterns behind the "numbers" and find the more appropriate formula for the business operation.
- **Customize actions:** Big Data could make it possible for every company to analyze better both the risks and the customers to meet more precisely the critical needs and to prepare better for the forthcoming risks or to provide better services to the customers, as it is feasible to separate the populations and find the specific needs.
- **Decision-making:** The decision-makers in companies traditionally are based on their experience, mindset, and knowledge to understand the environment and the factors affecting their services correctly; Big Data and Analytics make it possible with sophisticated analytics and statistical methods to find patterns and prepare better for

the future decision. Moreover, some platforms could not only support but also replace human decisions with automated algorithms.

- **New business models:** Big Data makes many companies establish their presence in a new way to produce more and new services and positively impact the whole community, as Big Data could present new horizons for society and business organizations.

McKinsey Institute (Manyika et al., 2011, p.7) mentioned that it is essential to measure Big Data's value; they underlined that we should consider only the dependent actions on the Big Data for an organization to measure the value.

It is interesting to explore Big Data and Analytics usage in a sector that might be traditional in its uses around the decision-making procedures; in the next chapter, the reader could find Big Data applications within shipping companies.

2.2.4 Big Data in the shipping industry

With a review of the literature with the above-given procedure, it is possible to find that four are the main elements that the international bibliography concerning around the topic of PSC inspections; the value of information, applications for predictive reasons, different aspects of the inspections, and lastly, the access on PSC data.

Port state control inspections took place daily in various ports by the supervisory authorities to ensure that ships and companies comply with all regulations and carry out inspections onboard ships. As the IMO (n.d.) states, the Port State Control inspections are done on the national ports to verify that the vessels are in good condition and operate safely and securely. As it is understood, the PSC inspections are contributed to maintaining safety and security standards in the industry.

Starting from 2018, Paris MoU stopped providing information on port inspections under its supervision, as typically stated in the RightShip (2018) press release. This change created a problem for the companies in the area, which were suddenly found to lack information about the global fleet's course. Also, the lack of information causes confusion among the stakeholders regarding the commercial motives of such moves. Several studies in recent years have focused on the factors that affect PSC inspections' results (Yuan et al., 2020); to make sense of all the information about companies, the information it is necessary to have access to this data themselves, to compare their results with the whole industry.

Derive value in an organization from Big Data is necessary to meet some primary conditions. It is needed for an organization to have the right people to carry out the required processes to extract information from the data and the organization to have a data-driven culture to use them properly; obviously, the data alone does not present anything (Sharda et al., 2014, pp.581-584). However, it is natural that the information that emerges from the data makes sense compared to the business environment; thus, it is possible to explain emerging trends. Consequently, the lack of data leads companies and organizations to draw the wrong conclusions, especially companies whose small fleets do not access crucial information worldwide. Bad decisions can lead either to a reduction in the company's profitability or complacency about global developments -and therefore to moves in an opposite direction of fleet safety and security.

As it is referred to the ISM code in the maritime sector, all companies have to give the responsibility of the implementation of Safety Management System (SMS) in comply with the regulation in a person of the company who is called Designated Person Ashore (DPA), as the DNV GL (n.d.) mention in its training course. This specific position in the company is working with the vessels to ensure that the fleet is safe for sailing and is responsible for communicating with the local port authorities during ships' inspection (Sollien et al., 2014). The in-inspection

vessel is safe for operation, following all the rules; the DPA must follow some PSC authorities' guidelines. For that reason, it is common for the DPA to maintain a database about the previous inspections from PSC or other authorities inspections. For those DPA's who are working with a small fleet is typical to search data about PSC focus from classes, flags, or possible in the MoU's; thus, if the MoU's do not provide about previous inspections on vessels, they lack information, and they are unable to take correct decisions. MoU's information is essential as each MoU concentrates and examines the vessels from a different perspective.

The search for information is considered essential both academically and in shipping; thus, building a complete picture of the research topic is possible. A recent study of findings by Paris MoU inspections from 2018 in the Baltic Sea found that ships' performance is affected by several factors, and management companies need to continually improve their performance (Hassellov et al., 2020). Recent research (Piniella et al., 2020) conducted for the Paris MoU inspections' held till 2014 found that the inspections depended on the inspectors, country, and port. Moreover, the inspector could affect the outcome of an inspection and many other factors; Fu et al. (2020) concentrated on the correlations between deficiencies and ship type, age, deadweight, and gross tonnage with data from Tokyo MoU; Indeed, in this study found that by knowing these correlations it is possible to improve maritime safety. On the other hand, it is possible to reach industry experts' views worrying that organizations, like Paris MoU, started not to provide that kind of information, as Rod Johnson mentioned in his article in 2017 (Johnson, 2017). As drawn from the previous studies, many factors have to be considered regarding the PSC inspection outcome.

Indeed, it seems crucial to have access to specific details on port inspections as there are many differential factors among ports and regimes; thus, the vessel's probability of detention depends on where the inspection is taken place (Knapp and Franses, 2007). A recent study by Ming-Cheng (2018) shows that the provided knowledge after the data analysis of the detainable deficiencies in Tokyo MoU, from 2000 till 2016, could be beneficial for both the shipping company and for the officers that inspect as they would be able to focus on the critical factors that affecting the condition of the vessel. It has to be mentioned that, International Maritime Organization (IMO) has to be credited for enhancing the shipping industry's safety and security, mainly the only rule-maker. However, as Mitroussi (2004) stated, the IMO and the whole community of shipping need to continue the efforts of enhancing the safety of the shipping; in that direction, therefore, it would probably be regulation on the MoUs as well to provide all the detailed data about the inspections on vessels in all ports globally.

First, it is essential to focus on the value PSC data could have in the industry and society in general, as it is concluded that this intelligence is vital for the coming days in shipping. In a recent study by Yan and Wang (2019), the value of the combination in one database about the PSC inspections mentioned could explain how the industry is thinking. Apart from the data format, Heij and Knapp (2018, pp.604- 621) tried to examine if it would be possible to extract intelligence from past deficiencies in PSC inspections to predict future shipping accidents, with an exciting conclusion that the human factor indeed has a significant role in those. It could be understood that it would be possible to prevent harmful situations for the global fleet and the oceans by gaining and analyzing the data correctly from the shipping professionals.

Secondly, as was expected, many researchers and institutes try to extract information from the data and provide intelligence to the shipping experts. Many studies focused on predictive applications of the vessels' risk factor or tried to approach an econometric analysis based on the PSC outcomes (Yang, Yang, and Yin, 2018; Yang et al., 2018; Xu et al., 2007), with those applications vessels managers can have a clear vision of the present value of the risk their ships faces. Data from the Tokyo MoU (2008-2017) was analyzed under the perspectives of grey relational analysis (GRA) with improved entropy weight to identify the nine factors that affect vessel detentions on ports (Chen et al., 2019). These results could be generalized and used in

other regimes to improve the detention rate, which would enhance general safety – as fewer detentions could be connected with more vessels in good condition. Yuhong et al. (2021) proposed a model for a PSC risk assessment, which could benefit both shipping companies and local authorities to improve the outcome and value of these inspections. As it is understood, all these applications and procedures demand data, which could be an alarm to the industry as less data – or even no data – could harm the industry and the community.

Thirdly, many surveys and studies focused on the differences among the PSC inspections; many factors could engage these differences, ranging from different motives to different cultures. Many researchers tried to approach the discrepancies between the regimes and how these differences involved the inspection outcome (Graziano et al., 2018; Ravira & Piniella, 2016); many factors are involved, such as the background and the focus of the inspector. Besides, many countries and ports are independent of MoU's and follow slightly different procedures, as Yuan, Chiu, and Cai (2020) stated that maritime states outside of regional MoU's should be assisted by IMO, given the examples of Taiwan in Tokyo MoU and of Montenegro in Paris MoU. These differences apply more risks in the industry regarding safety and security; thus, it would be interesting for the shipping companies' professionals to access the data to predict the inspections.

Lastly, few surveys and studies have been found regarding data availability in the maritime sector in recent years. It is possible to find an interest in how the data is provided (Meskauskiene, Oorni, and Sell, 2019; Yudono, 2018), focused on how citizens access these data from the maritime industry. It would be exciting for the public audience to research how accessible these data are for the public interest.

2.3 Theoretical Framework

Theories have value, and their role is essential in any research as they explain, predict, understand, and challenge the already existing knowledge of the given topic; the theoretical framework is the structure of the theories mentioned above. For that reason, the three subtopics highlighted in the research questions would be analyzed in the following paragraphs, and the perspective under these theories would be established.

The academic literature discussion in the above paragraphs provides the research with the required theoretical lens to construct the research's continuation and discuss the empirical findings' results. Indeed, this research would attempt to explore how PSC inspection data stored in large datasets internally in companies are used and analyzed by the shipping professionals to create insights and illustrate hidden patterns and risks regarding their fleets' safety and security; thus, how Big Data enhances the decision-making process using McKinsey theory of Big Data contribution to these processes. Laney's theory (2001) on Big Data characteristics and the description of Big Data's value from the McKinsey Institute (Manyika et al., 2011) would be used as a methodological tool to develop the interview guide and gather the insights of the shipping professionals who handle Big Data regarding the Port State Control inspection.

It is critical for the study to connect the research questions and the theoretical lens that would be followed to approach the answers the research investigates. The third research question was referred to the characteristics the PSC inspection data should have, as it is possible understood that question would be connected with the characteristics of the Big Data, as discussed by Laney's 3 Vs. Theory (2001); these characteristics are volume, velocity, and variety. The research would be exploring how that kind of data structure could be described as Big Data databases and if they could be treated like them from both the academic and industry bibliography.

The other two research questions refer to how the PSC data are used for the decision-making process internally in shipping companies and the impact these databases have on the industry. To understand and to explain the meaning of "value" for the Big Data structures is essential to establish a common explanation that would be used along within this research; this could be the description of the business value of Big Data from the McKinsey Institute (Manyika et al., 2011). Under this description, Big Data's value is provided by transparency, experimentation, marketing, or risk management, assist human decisions, and help design new business models.

As we have seen in the literature review, the concept of applying data mining techniques into PSC inspection databases is relatively new to the shipping industry; however, Tsou (2018) used, for example, the Market Basket Analysis as it is proposed by Agrawal et al. (1993, cited in Tsou, 2018) to explore the correlations between the deficiencies and the following outcome on each inspection. Of course, the procedure mentioned above is not the only one that could be applied for data mining reasons, as it could be understood that the requirement to apply data mining techniques is to have a Big Data infrastructure. Besides, PSC inspections data could be seen as Big Data as long as they have the characteristics Laney (2001) insisted, such as volume, velocity, and variety. It could be concluded that these PSC inspection datasets could be seen as Big Data infrastructures, and like these could be treated by both the academic and industry professionals.

Under those perspectives, the PSC data would be analyzed and explore their usage within the shipping industry. Of course, it comes naturally that the generic research design and the following chapters would be structured with the above-mentioned theoretical framework, especially in formating the interview guide. The following Figure 2 shows how this research recognizes the big data infrastructures on PSC inspection datasets, based on the concepts of Laney and McKinsey Global Institute.

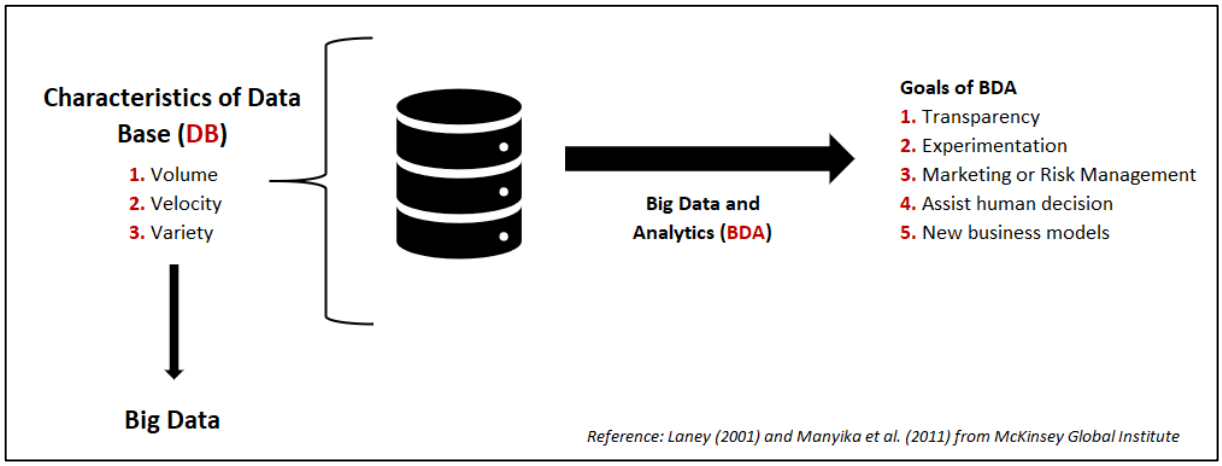

Figure 2. *Illustration of the Theoretical Framework, adapted from Laney (2001) and Manyika et al. (2011) from McKinsey Global Institute. (Source: Created by the author)*

3. METHODOLOGY

This method section aims to discuss the main research designs on qualitative design and establish the chosen one for this study that corresponds to the data and the survey's scope. Although it presents data collection and analysis techniques that work better with the study's data, the reader would understand this research's methodological inquiry in-depth. Finally, in this third chapter, the reader would find the essential steps this study followed to have accurate data and the ethical considerations about the data, the interviewers, and the whole procedure. For the elements mentioned above, the study followed the guidelines of Creswell and Creswell (2018, p.179).

It is crucial to define a clear methodological basis underneath the study; this would help both the researcher and the reader follow clear reasoning in reviewing the gathered information and the provided outcomes or explanations. This framework or theoretical lens gets involved in the whole procedure, as explained by Creswell and Creswell (2018, p.62); the framework is connected with the researcher's perspective, which shapes everything, from the types of questions to how the data would be collected and analyzed. As Creswell and Creswell (2018, p.5) analyzed, three are the main elements considered before research; the philosophical worldviews, the designs, and the research methods. The combination of these three elements brings the research framework underneath the study will be concluded.

3.1 Methodological Tradition

Four are the prominent philosophical worldviews discussed in the literature (Creswell and Creswell, 2018, p.6): postpositivism, constructivism, transformative, and pragmatism; each theory comes with different approaches and understanding of reality. Postpositivism is the traditional research approach, holds a deterministic model that causes determine the outcomes; it is broadly used in empirical research and has been used for theory verification reasons. Constructivist worldview (or interpretivism) is a theoretical stance mentioning that social reality cannot be understood with positivism as many factors affect the outcome; under the constructivist views, the researcher tries to identify multiple participants' meanings to understand the topic and generate a theory that could describe the phenomenon of the research. The transformative point of view, it is essential to underline that this is mainly used for political actions for change-oriented reasons – with this philosophy, to confront and eliminate problems and challenges that minorities face in our world. Finally, pragmatism is concerned with the actions' consequences on a well-established problem; thus, pragmatism required qualitative and quantitative data to have pluralism on the information before concluding.

As understood from the above paragraph, two are the ideal methodological paradigms for qualitative studies; pragmatism and interpretivism. Exploring the main two traditions would allow both the researcher and the reader to fully understand the procedures following this research's following stages.

Conversely, pragmatism in qualitative research focuses on actions and changes in society based on a continuous process of action and not in something stable and structured, as Goldkuhl (2017) mentioned. Thus, pragmatism examines social changes through the actions that lead to those changes; actions and consequences are critical for further development and clarification (Goldkuhl, 2017). It is feasible that pragmatism is concerned with finding the elements that constructed the "why and what happened" behind any social change. On the other hand, interpretivism does not have a single form; different formats existed in Information Systems like conservative and constructivist (Goldkuhl, 2017). Indeed, the terms mentioned above focused on hermeneutic and phenomenological traditions accordingly; the aim here is to present

the main ideas so the reader would become more familiar with the procedure this research follows and the differences with the others. The interpretivism paradigm's main aim is to contact, if not understand, the subjective meanings of persons who participate in the study (Goldkuhl, 2017; Curry, 2018). Besides, as the research investigates the participants' perspective, they should be careful with the generalizations regarding the sample's characteristics examined, as stated by Williams (2000).

It has to be underlined that the constructivist ontology is strongly connected with the interpretivism tradition; as mentioned in Goldkuhl (2017), ontologically given interpretive IS research assumes that the social world is not structured, but more probably is produced by humans through actions. In the constructivist methodological paradigm, the emphasis is given to understand the participants and their meaning to changes or actions. As a result, it is possible to find that the main difference between the main two traditions on qualitative studies on Information Systems is that the one (pragmatism) is concerning the actions that lead to the examined outcome, and the other is regarding the people point of view and feelings about a particular challenge.

Since the purpose of this study is to explore the individuals' thoughts and the meaning they give to data from PSC inspections and how they use that kind of data to extract value, the constructivist worldview could assist this research to come to conclusions on how the information from the MoUs comes to intelligence for the shipping companies from the professionals perspective.

3.2 Methodological Approach

Each worldview could be connected with specific research designs and methods to validate the literature information, and as is highlighted by Creswell and Creswell (2018, p.3), three are the main approaches, the quantitative, the qualitative, and the mixed methods. Quantitative designs are mainly connected with natural science to establish the connections and correlations between the variables, on the other hand, the qualitative designs are appropriate for exploring a phenomenon or a challenge to establish each component; the mixed methods, as it is coming natural, is a combination of the above two.

As discussed previously in this research, in the purpose statement, the focus would be the shipping professional point-of-view on PSC inspection data, resulted in the qualitative methodological approach for this study. Many authors highlight the characteristics of qualitative studies as Creswell (2016, cited in Creswell and Creswell, 2018, p.181), Hatch (2002, cited in Creswell and Creswell, 2018, p.181), and Marshall and Rossman (2016, cited in Creswell and Creswell, 2018, p.181); natural setting, researcher as a key instrument, multiple source of data, inductive and deductive data analysis, participants' meanings, emergent design, reflexivity, and holistic account mentioned in Creswell and Creswell (2018, pp.181-182).

Indeed, we could be sure that the outcome of inspections is the only observable event in port state controls; thus, as we have seen briefly in the introduction, many factors could have a role in this outcome – these factors could be described as the unobservable events. As this study would try to answer how the information from PSC inspections could provide intelligence to shipping companies and enhance safety in the maritime sector, it would be beneficial to proceed beyond the inspections and investigate the existing mechanisms. To explore the professionals' side is essential to have freedom on their thoughts and emotions about it, as the important is to gain insights into their feelings about it (Berg, 1989, cited in Roberts, 2014). It is critical for the success of the study to have a clear "canvas" for researching without predefined ideas and elements in it; thus, following the abductive analysis of Timmermans and Tavory (2012), which

is a qualitative data analysis approach, this study has tried to have the broader and deeper theoretical base possible and develops its academic output throughout the research process.

3.3 Methods and Techniques for Data Collection

The research methods – questions, data collection, analysis, interpretation, and validation – are connected with the aforementioned philosophical worldviews and designs. This subchapter outlines the data collection method, participation involvement, and sampling; for that reason, the selection of the participants and the whole procedure will be discussed.

As analyzed previously, qualitative research may be implied in the data collection procedure with interviews, document analysis, and observation (Creswell and Creswell, 2018, p.188). On the observation, the researcher may take notes of the participants' actions in their natural setting without involving in those actions or with the researcher's involvement regarding their study's requirements. Documentation screening refers to getting public and private documents around the topic to have a holistic view of the bibliography and the procedures that followed. Finally, the third applicable qualitative method for data collection is the interviews; those could be either in-person or focus groups. For this study, semi-structured interviews were chosen concerning the allowed time and the restrictions because of Greece's health protocols due to the recent pandemic. Indeed, interviews could be critical into interpretive and qualitative studies as the research may access the participants' point-of-view; as Myers and Newman (2007) stated for interviews, that could be considered as valid data collection method.

It is rational that every technique and method for data collection has some advantages and limitations; the same applies to interviews. Besides, interviews have usability when it is tough to observe participants in their natural setting when the study is interested in historical data that could be provided from the participants, and of course, it is convenient that the researcher could take control of the questions flow (Creswell and Creswell, 2018, p.188). Apart from the interviews' positive characteristics, it is essential to know the limitations of this chosen technique. It comes naturally from the above that the interviews provide information in a different place than the natural setting of the interviewer; moreover, the information could not be treated as public truth as it is filtered through the personal view of the participant, and finally, the researcher's presence may bias the responses.

3.3.1 Interviewee Selection

The population for this research would be the shipping professionals that work with PSC inspection data for vessels. For having complete and valid information, the study would have some limitations for the interviews participants; thus, the participants should work on a shipping company or a consulting company in the maritime industry and work with that kind of data for at least the last three years. The participants' recruitment would be used the interviewer's business connections with people working in shipping companies. Invitations would be published on social media (LinkedIn) and be sent through corporate emails to the possible participants. All the participants would be informed about the whole procedure. The interview's reflection would be used their participations without referring to them or their company; thus, it is crucial to building a safe environment to express their point-of-view without exposing businesses' initial procedures and outcomes.

The total number of participants in these private interviews should be defined using the bibliography. As defined in Creswell and Creswell (2018, p.186), the number of participants varies depending on the research design in qualitative studies; for example, narratives include

one or two participants, phenomenology from three to ten, grounded theory has twenty to thirty, and case studies four to five participants. In this research, six participants would be interviewed to explore their perspective on PSC inspection data regarding the three research questions provided in the first chapter of this research, as, indeed, this study is not narrative or grounded. In another study, it is possible to find that the ideal number of participants to undergo thematic analysis varies from two to two hundred (Braun and Clarke, 2016a). The aim of the defined number of interviews is saturation, as is described by Charmaz (2006, cited in Creswell and Creswell, p.186); the procedure could be stopped when no new concepts are provided from interviewers' sayings. As a result, following the guidelines from Creswells, and Braun and Clarke, the target of six interviews were examined during the interview procedure if needed to add more participants based on the fields that could be described as closed.

Having decided the population this research would examine is essential to discuss the different characteristics or patterns found among the participants. In all six of the interviews, the participants hold critical positions within the shipping company regarding the safety and security of the vessels and the seafarers; indeed, three were the DPA (Designated Person Ashore, while the other were the HSQE Managers (Health, Safety, Quality, and Environment). Thus, all of them hold managerial positions within the shipping industry helps this study examines the user experience from such databases from professionals with knowledge and responsibility on those PSC data; so, it could be said that their experience or thoughts are valid and worth saying. However, it has to be referred that the participants came from different shipping companies with different trade, in a manner of fleet segments or different areas of interest, that helped on the outcome with different approaches of the same topic of PSC usage public databases. Lastly, it is crucial to clarify that as this study aims to examine how the shipping professionals use these PSC databases, there was no need to examine their sayings concerning their demographic characteristics (such as age and sex); for that reason, such information is excluded from the interview procedure.

3.3.2 Interview Procedure

The main assumptions, as Phillips and Burbules stated (2000, cited in Creswell and Creswell, 2018, p. 8), are that open-ended questions would be used to find the interviewers meanings, based on the fact that humans make sense with the world with their historical and social perspectives, and finally, that the primary generation of meaning would be social.

Because of the participants' availability and the Greek active health protocols due to the COVID-19 pandemic, the interviews took place by Zoom and Microsoft Teams. A semi-structured protocol was followed for the interviews, and in the Appendices section (*Appendix A - B*), it is possible to access the interview guide. Furthermore, Walsham (2006) highlighted that audio-recording could harm the trust between the interviewer and the participant, which was an issue for consideration. Nevertheless, as in this study, the participants' reflections and insights are vital to have the most accurate opinions on PSC inspection data; an agreement between the researcher and the interviewer is signed for confidentiality. Although this study aims to conduct thematic analysis, the interviews would be impossible without the recordings (Nowell et al., 2017). As a result, all the interviews would be recorded and be accessible for the researcher to analyze them and then destroy all the collected data. Before the interviews, all the candidates must sign a Consent Form (*Appendix C*) with all the needed information about the following procedure and the interview protocols. Interviews were conducted in English and Greek language based on each candidate's profile. English is the universal language in the shipping industry, and Greek is the language many professionals speak since many shipping companies are operating from Greek territory.

The study followed the steps described below for the interview process. In April 2021, the researcher contacted the possible participants through the LinkedIn platform and discussed the topic of the study and their contribution to it; out of eleven firstly contacted targets, only seven out of them respond positively, but one of them deny the recording of the procedure so this participations excluded. After the first contact with the final six participants via email, the consent form and a brief description of the research and its aims were provided and organized the day and the time of the interview. During the interview, the interviewees have introduced the topic with a description of the research topic; after that, questions were generally asked about how long they work with that kind of data, providing their PSC databases' experience, and after that, the interview guide was followed. The duration of the interviews varied from 20 to 30 minutes.

To eliminate errors in the interview process, this study followed specific procedures. Precisely, the whole process is monitored, allowing the researcher to scrutinize the interviews and find possible misunderstandings. Besides, to exclude errors connected with subjective factors, all the questions are placed slowly and always asked for feedback regarding participants' understanding. Of course, it has to be underlined that all the participants had the time to answer each question.

3.3.3 Interview Limitations

The interview limitations are divided into two main sections; the first one refers to how the interview was conducted, and the second one is about protecting the sensitive participants' data.

The interview procedure took place under the restrictions for movement due to the recent COVID-19 pandemic in Athens, Greece. Because of these restrictions, this research followed the given protocols, following that all the interviews undergo via teleconference platforms (like Zoom and Microsoft Teams). Apart from the restrictions above, these platforms help the research contact shipping professionals worldwide; thus, one out of six interviews was with a participant based in Asia and working in a major shipping company. Additionally, these platforms provide easy access to the recording process of the meetings, something necessary for their analysis. It is then that this limitation helps the whole procedure with the interviews and allows broadening the horizons of the research with participants out of the Greek territory.

Considering the second part of the discussion in this chapter is essential to highlight that this study explores the usage of PSC databases around the shipping industry without considering the demographics or other characteristics of the professionals. As a result, there is no need for the research to collect data referring to people. As personal data could be described all the information that can be linked with a living person, only the company and the department would be mentioned; thus, as that information is not referring to a living person are not under the GDPR restrictions. Apart from that, a consent form was given to the participants to inform them how the research would treat their interviews and give permission to record their sayings.

3.4 Methods and Techniques for Data Analysis

The methods and techniques used for the data analysis from the collected information from the interviews varied on the steps compared to the quantitative data. As it is understood, the transcripts' produced volume would be required to be winnowing for the screening procedure to produce valid information from the interviews (Guest, MacQueen, and Namey, 2012, cited in Creswell and Creswell, 2018, p. 192).

For the data analysis of the qualitative data gathered from the interviews, Thematic Analysis (TA) extracted the needed information and proceeded with the findings. The scope of the TA is to conclude with initial codes that describe the participants' sayings. Clarke and Braun (2016b; 2006), as well with Nowel et al. (2017), proposed an analysis based on the codes extracted from the interviewers' sayings and then themes exported from them; while Lichtman (2013) introduced a more analytical approach that the initial codes were connecting with the categories, and finally with the discussed concepts. As Lichtman's approach seems more able to describe all the different perspectives of the interviewers, and indeed more analytical is chosen for this study.

First of all, the thematic analysis could be described by Boyatzis (1998, cited in Nowell et al., 2017) that thematic analysis could be the "translator" between a qualitative and a quantitative researcher and assist them to communicate their results efficiently. Indeed, thematic analysis is a highly flexible approach for various studies and inquiries; however, it has to be mentioned that the bibliography on thematic analysis is much less than other well established as grounded theory, ethnography, and phenomenology (Nowell et al., 2017). The technique mentioned above for data analysis has six steps till concluding as described by Lichtman (2013), these are:

- Initial coding,
- Revisiting initial coding,
- Developing an initial list of categories,
- Modifying initial list based on additional rereading,
- Revisiting the categories and subcategories, and
- Moving from categories to concepts.

Indeed, thematic analysis (TA) could be used to analyze small and large datasets from interviews, focus groups, or case studies from one or two participants to sixty or more (Clarke and Braun, 2016b); thus, this could allow this study to explore the sayings from the interviews using TA. The first step for conducting TA is to familiarize with the data; the researcher was familiarized with the data during the transcription and translated from Greek to English of the verbal content (Riessman, 1993 cited in Clarke and Braun, 2006).

In the second phase, the researcher proceeded to the generation of the codes by reading the transcripts. The coding was done manually, and the procedure was firstly the author to structure a coding list based on the research and the interview questions (researcher-derived codes), and secondly, codes arose from the thematic analysis of the sayings (data-derived codes). By the meaning of "codes," this study is referring to the provided definition of Braun and Clarke (2013, p.207) that: "a code is a word or brief phrase that captures the essence of why you think a particular bit of data may be useful". After that, the initial codes were categorized by discovering the typical patterns of information, and then the researcher moved into the next step and identified the emerged themes. Clarke and Braun (2006) suggest reviewing the themes and exploring the validity of the individual themes. As described by Clarke and Braun (2006), firstly, the researcher should evaluate the existing themes with the data and make the additional modifications generated from the missing codes. If the thematic map works, it is possible to move forward to finalize the themes; however, if any problem arose at this stage, the researcher should go back to the beginning of the thematic analysis.

Coming to an end, the final two stages refer to naming the chosen themes and producing the output – report. For reporting reasons of the thematic analysis, this study would follow Creswell and Creswell (2018, p.199), meaning that it would describe the sayings of the participants for each thematic element and compare the findings with the gathered information from the literature.

3.5 Reliability and Validity

Reliability and validity are critical elements for any research, either quantitative, qualitative, or mixed methods. However, there are differences based on the methodological approach; for example, reliability and validation on quantitative or mixed methods connect with these terms' mathematical approach. More clearly, Gibbs (2007, cited in Creswell and Creswell, 2018, p. 199) described the reliability and the validity in the qualitative approach as following: validity could be connected with the accuracy of the findings, and reliability is based on the researcher's approach and how wide in use this approach is in the academic bibliography. This chapter will present all the essential elements of the reliability and validity of the data based on the academic literature, and at the end, it will be present briefly to examine these aspects of this study's present data.

Regarding the procedures that will establish the collected data's validity, the researcher must use multiple approaches (Creswell and Creswell, 2018, p. 200); eight are the most frequently used strategies:

- Triangulate different data sources,
- Member checking,
- Use a rich, thick description to convey the findings,
- Clarify the bias the researcher brings to the study,
- Present negative information,
- Spend prolonged time in the field,
- Use peer debriefing to enhance the accuracy, and
- Use an external auditor to review the entire project.

As is referred above, the more of these strategies used by a researcher, the more is to feel safe about the data's validation. For this study, three strategies will be implied for validation reasons; their choice is based on the study's requirements and the time allowed. First of all, the researcher's bias would be clarified, so the reader would understand how the researcher's background and profile affect its insights. Secondly, the interviews' findings would be discussed both the findings that agree and those that disagree with the assumptions of this research; under this perspective, the reader would be able to explore by his or her own the usability of PSC inspection data from shipping professionals. Finally, the interviewer is in the shipping industry expecting to help this study with his experience in the maritime environment, so it is feasible to have a deep understanding of the in-discussion topics.

Coming to the qualitative reliability, it may be worth referring to Yin (2009, cited in Creswell and Creswell, p. 201) words suggests that documentation during the research process could be the key. For reliability reasons, the transcripts would be checked periodically after concluding the interviews and during the coding process; thus, both the researcher and the reader feel that the possibility of human error is eliminating. Finally, after these procedures, it is possible to interpret that this research is valid and reliable of the data concluding and how it treats these data and information for the interview insights.

3.6 Ethical Considerations

Several ethical aspects are considered prior, during, and after the interviews or during the findings' reporting (Creswell and Creswell, 2018, pp.88-97). This subchapter discusses these aspects, giving the required field for the empirical findings and the discussion.

First of all, the research problem – as it is stated on the introduction chapter – it is well defined, and it could have a positive and a beneficial outcome to the both academic and business

community; as the provided knowledge could be used for future studies or to understand in more depth how the shipping professionals are working with the PSC inspection data. The purpose statement in the first chapter will guide the reader through this study as it explains why this study has significance for the community. Apart from the elements above, the participants voluntary participated and were informed that they could withdraw at any time, and through the whole procedure, they had the freedom to ask about the motives and had the right to withdraw their participation anytime or to deny to conclude their words into the final transcript. The recruitment procedure is critical for the success of the research, so before the selection, the researcher tried to avoid harmful situations for the participants and tried to establish a positive environment (Lloyd and Hopkins, 2015). It has to be underlined that possible harmful questions or information were excluded; the purpose was to make all the participants feel safe to express their point-of-view. Secondly, this research uses them wisely for the public interest without any harmful information being published after collecting the data; with this procedure, this research develops trust between the researcher and the participants (Israel and Hay, 2006, cited in Creswell and Creswell, 2018, p.88). Finally, all the collected data will be secured and stored until the research's completion; after the study's finish, the researcher will keep no records. A consent form (*Appendix C*) will be signed between the interview participants, and the researcher mentioned all the research aspects and how this study uses the interviews.

Apart from the researcher's and interview participants' agreement, it is critical to clarify the "agreement" between the reader of this study and the researcher; thus, the researcher mentioned that the researcher worked in the shipping industry. To have unbiased conclusions on this study, it is essential to clarify that all the findings would be presented, whether they agree or disagree with the author's statements (Neuman, 2009, cited in Creswell and Creswell, 2018, p.96). Besides, this study uses public databases for PSC inspection data, something with no commercial worth; thus, there is no pressure on this outcome. As a result, the experience in the maritime sector would allow the researcher to explore the interviewers' sayings and make the required connections with IS theories and applications.

Concluding this subchapter, it is worth saying that this study tried to respect all the participants who engage with the research and the readers.

4. EMPIRICAL FINDINGS

After the description of the data collection and the analysis that the interview data have undergone through this study is critical to present the findings in a narrative style so that the reader would easily guide himself or herself into the discussion of the results. For such a study, the narrative report of findings seems ideal in the academic bibliography, as Miles and Huberman stated (1984, cited in Creswell and Creswell, 2018, p.209).

This chapter presents the empirical findings after completing the Thematic Analysis, with the procedure mentioned above; twenty-two codes arose under ten categories and five main concepts. To have a holistic perspective of the participants' sayings, the five concepts would be separately analyzed and examined in the patterns within this chapter. In Appendix D, it is possible to find the codes, the categories, and the concepts that the Thematic Analysis provides. After the first close screening of the transcripts from the interviews, we had twenty-two initials codes that arose from the participants. Following the Thematic Analysis' guidelines, we tried to combine the information to code to a more generic category (ten in total) and then to the final five concepts. Those five concepts are the following:

- Knowledge,
- Decision-making,
- Enhance operations,
- Data-driven, and,
- Challenges

With this procedure, we had the advantage of getting deeply into shipping professionals' sayings and analyze them accordingly. In the continuation of this chapter, each theme is discussed separately.

4.1 Knowledge

Shipping professionals' main goal is to understand what is changing in the industry regarding the PSC inspections, and for that reason, they regularly visited such databases. That is why one of the main themes the Thematic Analysis provided was the concept of "Knowledge", meaning all the operations that professionals are doing with the PSC inspection databases to inform themselves and their crews for the upcoming inspections. Their needs for knowledge are mainly focused on the information about the global trends and the transparency they believe should exist on Big Data about PSC inspections. Most of the interviewers (five out of six) mention these two perspectives on using such databases internally.

All the participants agreed that these databases provide transparency, as all the reported findings of the inspections can be found. It has to be underlined here that the main MoU's (like Paris, Tokyo, and USCG) are following a code system regarding the PSC inspections, but the codes of the findings are not internationally shared. However, it has to be mentioned that five out of six interviewees highlighted that it is possible to find all the PSC codes that have been inspected on vessels, but it not even close to describing what precisely happened; with accuracy, he mentioned:

> *"Yes indeed it is possible to find all the codes on these (data)bases, but it is not possible to understand what is happened even if you look about your vessels, as all these codes are describing something with such general words that it could be everything".*

The sixth interviewer did not mention something specific like the above statement. However, mainly this interviewer underlined the importance of having access to such information from these databases with the PSC codes, it is possible to know about *"what is going on, other similar to ours, vessels globally"*.

Half of the interviewees stated that some regimes (e.g., Paris and Riyadh MoU) do not provide the whole picture of the inspection outcomes, as they present either detention or just a few details about the inspection date or the vessel. One of those three connected this fact with the transparency, as mentioned above, and the available data volume; his words could be found below:

> *"The detentions globally is approximately a tiny fraction of the total inspections that taken place daily; having only those data available, you are coming in contact only with the worst scenario, and this is not the issue for everyone, especially for me and my fleet as we do not afraid of detentions."*

As it is feasible, the shipping industry worries about the total volume of the provided data from these databases.

Moreover, the participants stated that, through these platforms, it is possible to combine the needed information about specific projects on focus areas that periodically the regimes have and information about the probability of having inspections on the vessels based on their trade routine characteristics. In more detail, one of them highlighted the importance of such information,

> *"almost every month, the regimes are running some special focused inspections under the IMO; to have intelligence about what they want to find on the vessels and the approach you have to follow the only way is to visit such databases and extract value from their insights."*

As it is feasible, the published proposed procedures by the regimes or the MoUs are beneficial for the general safe sailing of the vessels, and the shipping professionals are alerted through these findings and prepare their vessels following these restrictions. One interviewer shared the same attitude when he commented that *"Indeed, the only way to find what happened last year regarding the safety and security is to 'visit' these databases"*. Besides, PSC inspection datasets provide valid information to the users.

However, four out of six interviewers mentioned the information's variety problem, especially connecting it with the lack of case studies. In more detail, they stated that more or less all the inspection codes reported in those databases are not more than six hundred, while in reality, each of them is connecting with plenty of additional codes underneath describing specific factors about safety and security. Specifically, their words could be described by the below statement, *"The codes are too general on the description, it feels like you face a problem and your friend suggests you not to be panic, it does not help, you neer real assistance; the same applies to these datasets"*. Indeed, these additional codes are never presented in these public databases, so information is missing; as long as a professional does not have all the information cannot be based on these databases for the decision-making process. They insist that it would be much more helpful to have all the additional information for every inspection so they could perform a benchmarking and a risk assessment for all the codes and train their crew more accurately. As one stated that *"by having a deep analysis within each code, it would be possible to have a realistic knowledge of what happened wrong on this vessel, which could trigger us to check our fleet accordingly"*.

4.2 Decision-making

The second concept refers to the decision-making process that is followed internally. In more detail, many participants have underlined the fact that one of the main reasons that they are using such databases is to inform themselves about the differences in MoU's approaches as we have already seen above; however, in some points, they correlate this knowledge to the decision-making process that they do internally. They agreed that this knowledge could easily upscale their ability to make decisions, as they know more factors that affect the outcome of an inspection; so through the Thematic Analysis, we concentrated all these comments under the concept of "Decision-making".

All interviewees stated that these statistics are incredibly beneficial for them on the annual meetings they have internally with the upper management as they can perform benchmarking with other companies, sister-vessels, and the global. From the first interviewer, wordings is possible to understand the value that metrics have on shipping companies, *"all the stakeholders among the shipping industry are extremely interested in KPIs and in general in those numbers, as everyone wants to optimize his or her performance"*. However, they seemed worried about the frequency of these databases and those KPIs because they want to access recent statistics and not for the previous years.

A typical comment regarding the PSC inspection databases was that these datasets have a limited number of inspection findings. As professionals stated within the interview procedure, almost all of these findings are too general in the description and cannot provide intelligence on what happened. During the interviews, the sayings regarding this issue varied from *"six hundred codes for all vessels it is almost nothing"* to more gently stated that *"with the current situation is hard to understand what happened to other vessels and you have to use your imagination and experience to find out"*.

The interviewers agreed that the real meaning of such databases is to enhance the knowledge before the decision-making process. In general, human knowledge and experience are fundamental for every decision, although the human brain faces problems when it has to make lengthy calculations and find patterns behind the "numbers"; for those reasons, databases assist with data-driven decisions. As a result, they stated that the PSC inspection database provides the required information for data-driven decisions as it is feasible to have access to some trends that describe the shipping industry in the last years. More clearly, one said: *"those trends of the KPIs* (that be given from such databases) *are the only way to know if your decline on metrics is affected by your approach or general differentiation from the inspectors"*.

However, they challenged the database creators' opinion that PSC inspection datasets could assist in decisions on their current structure. The main concern is on the format that the data are presented; more specifically, one stated that *"as these platforms are now could provide as knowledge and information, but it is impossible to come to decisions only by using them, as the information included in those is too limited"*. Another one added that *"if those platforms could be connected somehow with what we got internally, it might be possible to use them more automatically"*. Concluding, we could say that shipping professionals have many concerns regarding the utility of PSC inspection databases.

4.3 Enhance operations

The third concept is regarding the operations of the vessels and significantly how PSC inspection databases could enhance the whole procedure. In more detail, all the participants agreed that public databases are valuable and practical when the professionals try to collect information and intelligence about the upscales and downscales in the industry. They agreed

that through the analysis, which is possible through these platforms, any company could get access to generic statistics about the outcome of PSC inspections the last year, and as a result in comparison within previous years; so, it is feasible for anyone to obtain knowledge about the global trends or regimes' trends. One commented that *"databases like the Tokyo MoU or even Equasis could give some 'alerts' on trends as you can find details for similar to your vessels".* However, they all stated that they do not spend time on something like that intelligence is too generic and does not help them understand what is happening behind the uprising of a percentage of an inspection code. More specifically, one interviewer stated:

> *"No professional in any shipping company is interested in whether the code related to the ISM code is having more attention this year, as every year is different about vessels, seafarers, or even shipping manuals; what I would be interested in would be to have something with more depth to analyze, to find specific information".*

It is understood that the shipping professionals focus mainly on how they could use these platforms to obtain more information and knowledge for their fleet specifically, connecting with the second common code within this theme.

An interesting perspective of how these databases could be used came from one interviewer who stated that because of the large fleet that the organization he worked operates, it is possible to have specific information for many ports. He mentioned that

> *"in matters of PSC inspection findings, the company maintains a huge database with all the inspections taken place to our fleet for the last ten years; as a result, in those internally maintained databases, we have specific details even about the inspector name and what he or she found in our vessels the last time we were in his/her port".*

Of course, having personal details in these public databases would be unethical, illegal, and probably out of the scope of these platforms, as something like this does not assist in improving the shipping safety culture. However, it could be understood that private companies are moving based on their interest as they have commercial value in using such databases.

The one interviewer who mentioned the educational perspective of using the database clarified that only the current use them in such a way among the companies he worked. He highlighted that the company built a platform that combines the information from corporations fleet with the data from Tokyo MoU and tries to build case scenarios for the interns and entry-level employers. This interviewer stated that: *"it is beneficial for the new comings on an entry-level job to spend time on such an education, as they could understand the nature of the maritime sector and the company's profile"*. As it is understood, something like that is expensive to be structured as it demands working units from the company, and on the other way, it assumes a large fleet to have a volume of historical data to analyze. As a result, something like that could have value only to the most significant corporations globally. However, such data from PSC inspection outcomes could be used as case studies within the companies to structure and approach theoretical issues for their fleet and, in that way to enhance their readiness. Indeed almost half of the interviewees agreed that *"in shipping, the root-cause analysis is almost on our daily operations, building case studies with PSC data could be enlighting some issues that our fleet has, and we are unable to discover it so far"*.

4.4 Data-driven

The fourth concept is connected with how PSC inspection databases assist shipping professionals and companies to work more data-driven. It is rational that these databases could provide genuine assistance to companies to be more data-driven regarding PSC inspections. Almost all the participants clearly stated that these platforms help them find the needed KPIs to discuss and critically evaluate their results and operational outcomes. In more detail, one of the interviewers stated that *"all the metrics 'speak' to you when you compare them with the industry"*. It is clear that to have a risk assessment of the vessel regarding the possibility of being inspected and having the worst outcome, all companies and professionals need Big Data and Analytics techniques to mining the intelligence from PSC inspection databases regarding their vessels. Only one of the interviewees stated that they do not use such KPIs internally; more clearly, he said that *"We as a company we do not care about what the industry is doing, what we are trying is to make our statistics better in comparison with the previous years".*

Finally, the shipping professionals were worried about how these datasets could assist their decisions in a more data-driven approach, with live metrics or even vessel-specific alerts. They conclude that the lack of depth on such PSC inspection databases makes data-driven decisions impossible to take. Almost all of them agreed with one of their colleague's sayings that *"it is not possible to have data-driven decisions when these datasets, at first do not provide filtered insights or a deep analysis of the inspection codes"*. It is understood that shipping professionals are still based on their experience to understand the external factors that affect an inspection or decide the next step to enhance the safety of their vessels.

4.5 Challenges

The fifth concept refers to the challenges that shipping professionals faced while using PSC inspection databases and the Big Data that they provide. All of them asked about the updates of these platforms and whether they felt happy about the frequency. They agreed that Equasis is quick on inspection reporting, although it has some issues when the vessels change "hands" among companies. While, on the other hand, the databases from Paris and Tokyo MoU are updated almost daily, making them a trusted reference when exploring what is happening. However, the USCG has different procedures, and it is impossible to find the generic picture, as the professional has to search for specific vessels within their database, which makes the operation challenging. In general, their saying could be summarized by the below statement, which one of them underlined, *"the frequency of the updates is not the priority at the moment, especially when we are talking about gaining information for the forthcoming call of our vessel"*.

All participants stated that as it may come naturally, they have more statistics and information about their inspection outcomes than the information could be achieved through public databases. Moreover, one stated that from curiosity, once he tried to found on those public databases from MoUs the inspections for his fleet, he could not extract enough information as it was too generic. In detail, he said that *"A few months ago just from curiosity I have visited the Equasis about a detained vessel we had in Antwerpen, I was completely unable to extract data and intelligence from the way they described the issue"*. However, they all mentioned that they regularly visited a public platform with PSC inspection information, called Equasis, to find specific details and knowledge for their competitors to learn from their mistakes and avoid dire circumstances. They use such Big Data platforms, but six out of six participants agreed on using the provided intelligence from the Equasis and MoUs platforms without the analytics perspective. With more detail, a participant stated that *"Equasis is a really helpful*

database to find specific information about the fleets; the disadvantage is that it presents the raw information for the inspections without any analysis".

The other three comments that the shipping professionals have made during the interviews could be connected somehow. There was a disagreement on whether these databases provide transparency to the maritime sector and at which level. Half of the interviewees stated something close to that *"starting from Paris MoU that do not provide as much data as it is used, other regimes do not have such databases, like Riyadh MoU"*. On the same tone, another one underlined that *"almost in any public database I have ever visited does not provide a general idea of the inspections, you have to search specifically with the IMO number, for example, to find something and this is handy"*. In contrast, some shipping professionals stated that everyone is allowed to find information for everybody through such databases, and that is helpful for professionals to understand what issues other companies are facing and try to avoid such issues.

During the interview process, they explored what is missing from these databases and give them more. Besides, they agreed that different characteristics and more depth on the information included could alert them more as they would extract more intelligence. In more detail, two out of six stated that if they have sub-codes within the primary five hundred codes, it would be feasible to use that kind of information for assistance for preparation reasons. In more detail, they stated: *"If it would be possible to gain in more details why the deficiency took place, in which section of the vessel, for example, it would be possible to give more direct guidance to the officers on the vessels and to have at last a more safe vessel"*. Four of the participants underlined the importance of having a common language between these databases. They mentioned that, for example, the Paris and Tokyo MoU use standard codes, while the United States Coast Guards (USCG) use different codes, which destroys the information as the signals are different. On many occasions, such differences make both the officers from the office and the crew on the board be misleading because they do not fully understand the requirements.

The participants spoke a lot about all of them during the interview discussion and commented on what is missing from PSC databases to help them upscale their business operations. They mainly gave feedback on using these databases; they concentrated on the missing elements based on the data's characteristics. With more details, the interviews enlighten how the shipping professionals understand and use Big Data platforms like the PSC inspection databases provided by public sources; thus, it is interesting how they understand Big Data as they are not familiar with the fundamental base or structure of such platforms. Indeed, the first interviewer stated at the beginning of the procedure that *"It is interesting for us as professionals to participate in such interview procedures, as throughout this, it is possible to think and evaluate the procedures that you follow daily, and you do not give them the required attention"*. As a result, the willingness of this study was to examine this precise understanding.

4.6 Summary

Concluding this chapter, we could have a clearer picture of how the shipping professionals use public PSC inspection databases to assist their daily procedures. The interviews guided this research into five themes. Indeed, PSC inspection databases upscale the knowledge and the understanding of the nature of these inspections for the responsible professionals. This knowledge is connected with the decision-making procedures that shipping companies follow to comply with the internal and external factors that affect their performance. As it is understood, more data-driven decisions help the responsible professionals to enhance their daily operations and upscale the produced outcome. Finally, all the interviewed shipping

professionals raised many concerns regarding the frequency of the updates or the volume and the complexity of the included information. It naturally comes that, like everything else, these public PSC inspection datasets connect with some pros and cons for the shipping industry in general; of course, some results alert us that shipping professionals still want more things from these Big Data infrastructures.

5. DISCUSSION

This chapter discusses the findings after the Thematic Analysis presented in the fourth chapter of this study and tries to find how they answer the Research Questions and how they are connected to the theories presented in the Literature Review chapter; the research questions connected with the thematic analysis, so the reader has a map to guide through this chapter.

This research aimed to explore the shipping professionals' perspective on Big Data regarding the PSC inspections database usage. The focus was on Big Data characteristics and their value for the users and the shipping companies. For the reasons mentioned above, three research questions were placed:

- What value have the port state control (PSC) inspection data for a shipping company?
- How port state control (PSC) data are used for decision-making?
- What characteristics port state control (PSC) data should have?

The first research question refers to Big Data's value on shipping companies, and it connects with how the shipping professionals are using that kind of information. The research shows that Big Data could help shipping professionals inform global trends and look specifically at what triggers deficiencies and detentions on vessels. Indeed, through the Thematic Analysis, the concepts of "Knowledge" and "Enhance operations" connected with the first RQ about the value that it could be gained from using such databases. As a result, we could say briefly that Big Data could have value in a shipping company by providing a better understanding of the maritime environment and enhancing the daily operations of the responsible departments.

The second research question is regarding the usability of Big Data to the decision-making process within shipping companies. This research provides the professionals' insights for a more data-driven mindset during the daily operation in the office as more data-driven decisions help the responsible professionals to enhance their daily operations and upscale the produced outcome. For these reasons, we can see that the second question is correlated with the themes of "Data-driven" and "Decision-making". The research concluded that it is feasible to have a more data-driven approach by upscaling the data and their metrics, which affected shipping professionals' decisions.

In the end, the third research question is about the unique characteristics Big Data should have to be helpful to shipping professionals. The participants' opinions were gathered under the concept of the "Challenges" as shipping professionals critically discussed Big Data's characteristics that they believe would be beneficial for the professionals and their daily operations. Significantly, they raised many concerns regarding the frequency of the updates, volume, and complexity of the included information. Table 1 presents the connection between RQs and the concepts that emerged from this research.

***Table 1**. The connection between RQs and the Main Findings (Source: Created by the author)*

Research Questions	Main Findings - Themes
RQ1: What value have the port state control (PSC) inspection data for a shipping company?	Knowledge, Enhance operations
RQ2: How port state control (PSC) data are used for decision-making?	Data-driven, Decision-making
RQ3: What characteristics port state control (PSC) data should have?	Challenges

It would be interesting to try to come closer to these three research questions using the answers from the shipping professionals who participated in this research; however, it is

essential to highlight that our produced outcome could give a meaning based on interviewers' sayings and notices.

It is coming rational, of course, that the discussion of the empirical findings faced some limitations based on the chosen methodology of this research and the participants' selection. The qualitative research methodology could explain the factors contributing to the PSC inspection databases' successful usage; however, this research alone could not structure a theory around the Big Data characteristics in the shipping world. Combining this study with another one that might include other professionals' opinions within the shipping industry, like insurance officers or brokers, could allow the community to approach the topic holistically.

The theoretical lenses used as a basis for this research are shown in Table 2. Thus the first two research questions can be connected to the five elements theory of McKinsey Global Institute regarding Big Data usage. At the same time, the third research question is trying to discover the characteristics of Big Data that can provide helpful information to shipping professionals under the perspective of Laney's three Vs. theory.

Table 2. *The connection between RQs and the Theoretical Framework (Source: Created by the author)*

Research Questions	Theoretical Framework
RQ1: What value have the port state control (PSC) inspection data for a shipping company?	Five ways that Big Data provides value based on McKinsey Global Institute
RQ2: How port state control (PSC) data are used for decision-making?	
RQ3: What characteristics port state control (PSC) data should have?	Laney's 3 Vs. Theory

The analysis below would follow the concepts that have been introduced from the Thematic Analysis. As a result, the reader of this research can find specific information that arose from how Shipping professionals use big Data. In the continuation of this chapter, the research findings would be discussed critically based on the given Theoretical Framework and the mentioned theories and applications in the Literature review.

5.1 Data and knowledge

Shipping professionals use Big Data for informing reasons about global trends. Indeed, the provided information from such databases could be vital to organizations and businesses for transparency reasons, as in those, everything is being reported, so all the stakeholders could easily find the information they are searching for. Besides, Beynon Davies (2020, p.169) insisted that value from Big Data could be described as the output's worth for the researcher or the organization, and this could be connected with the five-element theory of McKinsey Global Institute (Manyika et al., 2011), which included the "transparency" into the meaning of the business value that could be extracted from Big Data.

According to the Big Data roadmap provided by Palem (2014), the information about the global trends that could be achieved from such databases for shipping companies can be connected with long-term goals from a macroscopic point of view. Thus, shipping companies always want to enhance their safety performance (Hasselov et al., 2020); rationally thinking, it is almost impossible for any company to measure its performance without comparing its results with other companies simultaneously (Yuan et al., 2020). As a result, research suggests that the need for transparency could be achieved through free access to all inspections data for everyone.

Shipping professionals use Big Data from PSC inspection databases to contact other shipping companies' outcomes. As discussed above, transparency is a value from Big Data as the McKinsey Global Institute theory has analyzed it on five elements (Manyika et al., 2011): transparency, experimentation, customized actions, decision-making, and new business models. However, the shipping professionals introduced their need for veracity within Big Data, as veracity was firstly mentioned by IBM (Lukoianova and Rubin, 2014). Indeed, veracity is not included in the 3 Vs. theory from Laney (2001), as only the terms of volume, velocity, and variety are included in this theory, but the research showed us a need for veracity in Big Data. Additionally, the data alone, without an analysis, do not present something with value (Sharda et al., 2014, pp. 581-584), and as Yuan et al. (2020) stated, shipping companies should compare their PSC inspection results with other companies to extract intelligence from this benchmarking and evaluate their process. Therefore, shipping professionals found it critical to have easy access to other companies' performance or, in general, to get informed about their dynamically changing environment. As a result, shipping professionals add a fourth "V" in Laney's theory.

This research proposed that shipping professionals need more information and more complex data to extract intelligence from Big Data. At first, the variety is a characteristic of Big Data introduced by Laney's theory of 3 Vs. (2001); secondly, the reason behind the wanted variety from the shipping professionals is connecting to experimentation and decision-making processes within companies. According to McKinsey Global Institute, the characteristics mentioned above of Big Data are contributed to some ways to extract value from Big Data (Manyika et al., 2011).

As it is stated by Oracle (2019), the value from Big Data usage could come from a large volume of data, as the volume could be connected with the theory of large numbers in mathematics and lead to more safe answers to the questions tried to be approached by the Big Data. However, Yan and Wang (2020) proposed that the shipping industry combine all PSC databases to have more information and the needed data for experimentation reasons or even estimate the outcome of the forthcoming PSC inspection.

Shipping professionals are in contact with the case studies, and root-cause analysis is demanded on many occasions by the public organizations, such as the IMO, or the companies they work with, like the P&I clubs. As a result, as they stated all of them, it came naturally to create a root-cause analysis internally for educational reasons with their co-workers; thus, the same implies with the special projects that the MoU's periodically have in order to enhance the public awareness for specific elements regarding the safety and security culture in the maritime sector.

The shipping industry and the professionals work with intelligence extracted from Big Data for experimentation or the reasons mentioned above. It is possible to connect McKinsey's wordings about experimentation (Manyika et al., 2011), and as a result, to illustrate that indeed the professionals understand that kind of value that could be exported from Big Data. However, the lack of this ability from Big Data, making their operations more complicated and with less information.

5.2 Enhance operations

The second theme supports the value of big data on enhancing operations through knowledge dissemination regarding fleet-based issues, like the ship type or age. Extracting this information is also valuable for them to guide the company's superintendents to "prepare" the vessel for the next inspection, and with this procedure, they ensure that the vessel is safe and ready to sail. As McKinsey Global Institute insists, Big Data could derive value with transparency and

customizing actions (Manyika et al., 2011). The shipping professionals could create different populations from the general sample by applying specific techniques on Big Data to explore different patterns for the interested fleet; this creates value for the shipping professionals to explore their fleet's needs deeply.

Research findings showed that shipping professionals segmented fleet categories because different vessels connect with different PSC inspection outcomes, and as it comes naturally, they are looking for accurate information to decide the next step. Indeed, Fu et al. (2020) found that inspections outcome correlated with ship type, age, and other demographic characteristics of the vessels. At last, shipping professionals could extract value from Big Data techniques on PSC inspection databases by examining the results and trends for the specific population of vessels they are interested in. It seems rational that the shipping industry needs specific PSC information and not just generic trends.

Indeed, shipping professionals have noticed by experience that inspectors' background matters in the outcome of each inspection and the port or the region that the inspection took place. This phenomenon could be connected with Piniella et al. (2020) sayings that inspections outcome depends on the inspector, country, and port. As a result, Big Data procedures employed in PSC inspections could lead companies to make the wisest decisions to prepare the vessels and train better conditions and more depth the crew.

The research presented that PSC inspections Big Data is valuable for the professionals to guide the company's superintendents to “prepare” the vessel for the next inspection, and with this procedure, they ensure that the vessel is safe and ready to sail. This procedure is common to be done while the ship is operating; thus, both the experience of the employees and the information they can have could make them have a risk assessment of the vessels in real-time. With this risk assessment, the superintendent, with the assistance of the company DPO, should handle the situation and make the most appropriate decisions.

The risk management of a challenge through customizing actions and the decision-making that shipping companies’ professionals evolved demand Big Data and Analytics to enhance their outcome and understanding. Indeed, McKinsey Global Institute proposed that Big Data value for businesses and organizations could be connected with customizing actions and decision-making procedures (Manyika et al., 2011). Indeed, research showed that shipping professionals feel much more comfortable now with the information and insights they have through Big Data than what was happening a few years ago when Big Data techniques and, in general, technology has not entered the industry. Significantly, the ability to find specific intelligence for their fleets characteristics makes them possible to make more data-driven decisions, as Fu et al. (2020) specified that correlations existed between deficiencies and ship types, age, and general vessels’ characteristics.

An exciting addition from the research was that shipping professionals with managerial positions use Big Data PSC inspections internally to educate seafarers or officers. McKinsey Global Institute (Manyika et al., 2011) stated that experimentation is one of the five perspectives Big Data could create value for businesses. Indeed, the educational procedures that companies are trying to imply could be read under this perspective. Besides, education could be seen as a long-term goal for any company as it is the required procedure for any organization to create the professionals that would lead the organization in the future; as Palem (2014) stated, this could be achieved through Big Data.

This study concludes that shipping professionals use or want to use the PSC inspections Big Data to conduct case studies internally from the interview process. Case studies are essential to root-cause analysis that shipping companies should perform after any non-clean inspection in many regimes globally as it is a requirement by many stakeholders. As it could be easily understood, the case studies could be connected with the meaning of “experimentation” as

presented by the McKinsey Global Institute (Manyika et al., 2011) as the value of Big Data for organizations.

Shipping professionals through this research insisted that various data would trigger companies and professionals to perform such case studies for preparation reasons as they would have a larger volume and complexity of data. Indeed, as SAS (n.d.) stated, complexity could assist professionals to have more accurate analytics from Big Data use. In the shipping profession, this could enable shipping companies to maintain such datasets more appropriately and perform a more complex and in-depth analysis of the risks they faced, and possibly they would be able to upscale their performance using Big Data.

5.3 Data-driven decision making

The outcomes of this research support Sharda et al. (2014, pp. 581-584), who implied that the data alone does not present anything, and with assurance, we could say that Big Data without the professionals can understand and use them through the decision-making process, is unable to enhance the total outcome. During the research, it was found that shipping professionals use these Big Data on the annual management reviews meetings (MRM) to present the "position" of the company in comparison with the industry or even specific managers with a similar profile to them. Yuan et al. (2020) sayings about the need to compare PSC inspection outcomes among companies to gain intelligence from such data are evaluated from the six shipping professionals, and they share the same ideas.

Indeed, during these meetings (MRM), the connection between Big Data with decision-making seems clear. McKinsey Global Institute proposed the five-elements theory (Manyika et al., 2011) to draw value using Big Data. Thus, customizing actions in risk management, decision-making, and exploring new business models are aspects that those meetings explore with Big Data techniques. The intelligence exported from such PSC databases is beneficial, as they can present the figures of safety and security of the vessels quantitatively and make these meanings more understood among all stakeholders.

Another finding was that shipping professionals spend time searching to find the details they are interested in based on their fleet characteristics. With this, the shipping professionals referred to two perspectives of Big Data; in more detail, they describe the variety as introduced by Laney (2001) and in combination with the complexity by SAS (n.d.). They highlighted the importance of having data that present the reality more clearly. Thus, the shipping professionals agreed to extract more information if they have more depth on the data included in already existing Big Data infrastructures.

In more detail, shipping professionals stated that variety is essential regarding Big Data, and they acknowledge that the most considerable percentage of data is unstructured (Cukier, 2010). Although the existing PSC inspections are the only way to come in contact with such information and knowledge, and as a result, they are unable to have valuable insights from these as long as the data included in them are in such a format. As Hasselov et al. (2020) showed, the ships' performance on PSC inspections is highly dependent on many factors and based on Piniella et al. (2020), inspectors, country, and port are critical on the outcome. All these resulted in shipping professionals having the right to be worried about the decisions they have to make regarding the PSC inspections as they lack information, pinpoints the need for variety on Big Data.

The research showed that shipping professionals are visiting or using PSC inspections Big Data to extract insightful intelligence. Using the meaning of "insightful intelligence", they talk about the metrics and the insights of the patterns discovered by analyzing such information. In parrel with the analysis of Qi et al. (2016) that connect the insights from the customers with the

company's strategic value, it is possible to connect the insights from the MoUs or even specific ports and regimes with the strategic value for shipping companies regarding the PSC inspections. Indeed, those insights could be connected with the decision-making process that McKinsey Institute describes; through the intelligence extracted from the databases, the companies could enjoy more data-driven decisions either with automated algorithms or specific KPIs in the human decision process.

An exciting addition is that shipping professionals use such information and data from public PSC inspection databases to make data-driven decisions for their fleet and company strategy. McKinsey Global Institute (Manyika et al., 2011) stated that Big Data value correlates with the human decision-making process, either automated or without. Moreover, it is convenient that Big Data could be helpful to both micro and macro decisions, as Palem (2014) distributed the short, medium, and long-term goals that Big Data could be assisted the professionals. Shipping professionals agreed that public PSC inspections data are assisted them in making more data-driven decisions by having a better knowledge of what is changing in the industry compared with the volume of knowledge they had a few years ago when the Big Data usage in the shipping sector was not so broad.

Bertei et al. (2015) connected Big Data on business process management (BPM) with the strategic decision-making procedure for companies and organizations, similarly working the Big Data platforms regarding the PSC inspection data in the shipping industry. The decision-making procedure has some steps requiring specific actions from humans and public organizations to take to have an outcome; a decision-making typology like Knowledge Management (KM) proposed by Intezari and Gressel (2017) could be implemented in the shipping sector as well. Combining internally stored data from companies and Big Data regarding shipping issues could provide valid information to companies to have more fundamental knowledge about the business environment. This upscale on knowledge could trigger an upscale on the operation, as professionals stated within this research.

During the research, concerns were raised regarding the volume of the insights they could extract from these PSC inspections Big Data. Besides, professionals' main concern was about the automation of these platforms as they have stated that it is tough to extract the required information and insights. A specific employee or even a whole team working internally for these reasons and shipping professionals are worried about waste of human recourses for having probably not the best picture of the insights that could be extracted. Indeed, the insights extracted from Big Data could be connected with gaining a better position following the competitors and enjoying the strategic value, as Qi et al. (2016) mention, in a functional perspective (Grover et al., 2018). As Oracle (2019) noticed, the value could come only from analyzing a large volume of data, and that is what shipping professionals are trying with few "tools" as these datasets are closer to large lists of incidents than structured databases. Following the meaning that Beynon-Davies (2020, p.169) gives to the "value", it is feasible that shipping companies have to weigh the advantages of using Big Data in comparison to the working hours spent to extract value from them.

Additionally to the management review meetings, the second parameter is the ability of the shipping professionals to find the appropriate KPIs and metrics to compare with their results. As it is mentioned above, plenty of times shipping professionals need to perform benchmarking with their safety performance through PSC inspection outcomes, as at first this obligatory in some times for commercial reasons, and secondly, it is the only way to access and evaluate their performance (Sharda et al.,2014; Hasselov et al., 2020; Yuan et al., 2020). Indeed, this opportunity could be connected with McKinsey's sayings about decision-making and evaluating new business models (Manyika et al., 2011). Through Big Data techniques on PSC inspections, shipping companies can evaluate their procedures as it is the only access to what their competitors have done in the previous years and how to access their successful adjunctions.

Mazzei and Noble (2017) stated that through Big Data, an organization could improve functional operations; thus, shipping professionals agreed that their operational abilities had been enhanced recently in parrel with their ability to extract intelligence from such databases.

Risk management could be done by customizing actions to prevent harmful situations or being prepared on the best affordable level. Indeed, as McKinsey Global Institute insisted, this could be happened by segmenting the population into separated samples with unique and well-described characteristics. Professionals used Big Data techniques within shipping companies using PSC inspections to populate the data based on the shipping professionals' criteria.

Less risk could be connected with more accurate and wise budget spending of any company or even person; correct use of the budget allows shipping companies to expand their trade routes and enjoy the market's leadership. As a result, it could be connected so that the risk management could enable the professionals to map and structure new business models and find more revenues. Manyika et al. (2011) from McKinsey Institute connects Big Data usage's earned value with the “new business models”.

5.4 Challenges

Indeed, the main challenge that shipping professionals are faced is not far away from the problems that other industries are faced; that is the fact that approximately only 5% of all existing data is structured (Cukier, 2010). This fact could be connected with all the needed actions professionals have to access knowledge and intelligence from data; however, the first step is to access the data. This research showed that shipping professionals need more data and information about the PSC inspection outcome than they already have, so they generally require more depth through the Big Data. The research mentioned that Paris MoU, for example, published only the detainable inspections. As a result, professionals and managers within shipping companies lack information as the detentions are only a tiny fraction of the total inspections. However, through those concerns, it is presented to us that variety on both publishing resources and the data as they are, have a significant role on the value that the user of Big Data could have.

Another aspect of the significance of knowledge sharing through Big Data databases is the need for a common “language” among Big Data infrastructures; by the common language, we are referring to the same signals or codes. For example, Paris and Tokyo MoU are sharing almost the same definitions on the PSC inspections. At the same time, on the opposite side of the Atlantic, the United States of America uses completely different codes. This phenomenon builds more difficulties for the shipping professionals as they cannot combine the databases and extract knowledge from this combination. Their sayings could be connected with on “V” of Laney’s theory, that Big Data should have volume, should be contained from large datasets.

An exciting addition from this research came from plenty of shipping professionals who mentioned that internally they are not using daily such information and data regarding the PSC inspections from Big Data because of the frequency of the provided insights. The frequency of the updates on the data is essential; many daily inspections took place worldwide, and while the outcomes presented to the public with a delay could be connected with wrong decisions from the side of the shipping professionals. Indeed, the velocity is one of the three Vs. established through Laney’s theory (2001), and the shipping professionals highlighted this factor of Big Data.

Through the recent technological upgrades worldwide, it is possible to find that many organizations have to use massive datasets with many data and information inserted every second that it passed. For example, Wal-Mart with more than one million transactions per hour (Cukier, 2010), or even Facebook with more than one million photos per second (Beaver et al.,

2010); as it is evident by the shipping professionals, the volume of the daily “transactions” on PSC inspections is not even close to these numbers. This results in the professionals' need to have more recent updates on the insights from Big Data. However, it has to be said that all of them prioritize the needs and mention that if the complexity of data does not come, there is no need for more frequent updates as long as these few data presented now are valid only for informative reasons. There is no need to visit these platforms, again and again, to extract either manually or automatically deep analytics from Big Data.

The research showed that shipping professionals need more information and more complex data to extract Big Data. Variability and complexity are often mentioned during the interviews, with similar characteristics as those introduced to SAS's business and academic community (n.d.). Besides, the shipping professionals stated that the volume is essential, but they have half of the picture if the data are not getting in-depth. The one participant who stated that his company, due to its size, can use such data for academic reasons highlighted the importance of detailed data to extract intelligence and transform all the metrics and the intelligence into decisions as McKinsey Global Institute states (Manyika et al., 2011). The value of Big Data is to enhance human decisions, either automatically or manually; as far as professionals cannot use such platforms for the decisions they have to make, these platforms could not be counted as valid Big Data structures that could produce value.

One of the main challenges shipping professionals face is that not all the regimes or MoUs present a database with all the PSC inspection outcomes. Indeed, this phenomenon could be connected with many difficulties that shipping professionals face regarding using such databases. As a result, they cannot use such information easily as Big Data, and many use just as large datasets for informative reasons. As there is no possibility to combine the information from all the regimes, it is impossible to have a clear picture of the inspections globally and prepare accordingly, as Yan and Wang mention (2020). Although, the missing data from the MoUs affect the transparency that Big Data could bring to the maritime sector. Both McKinsey Global Institute (Manyika et al., 2011) and Qi et al. (2016) agreed that the (strategic) value could be enjoyed from any organization with information and insights. As it is feasible, it is impossible to have valuable metrics and insights if the information is missing.

Piniella et al. (2020) provide an analysis to the shipping sector pinpoints that the inspection outcome varies depending on the inspector, country, or port. As a result of not having the outcome of non-detainable inspections in one of the sector's most prominent regimes, professionals have much less data to analyze and provide intelligence. Indeed, the research shows that shipping professionals try to overtake this challenge by assuming that Paris MoU has similar approaches with the Tokyo MoU; as illustrated clearly, at first, this is coming out of their experience, and it is risky for its validity. McKinsey Global Institute’s theory regarding the value the organization could extract from Big Data discusses the severity of the transparency (Manyika et al., 2011). Indeed, Big Data infrastructures as those that came from PSC inspections could provide “governance” to the industry, but it is essential to have easy access to all the data to be able to analyze them and extract intelligence from them, as is stated by Oracle (2019) and Sharda et al. (2014, pp. 581-584).

The absence of filters and tools in those PSC inspection databases to extract the needed information seems controversial. Shipping professionals have to transform Big Data into different formats (usually Excel) to use specified filters for extracting the information they are investigating. Qi et al. (2016) presented the value that automatic filtering models have for predicting online reviews; something similar may value these infrastructures that shipping companies use internally.

McKinsey Global Institute proposed the five-elements theory for “calculating” the extracted value from Big Data into businesses and organizations. Through this research, the reader can find that shipping professionals often referred to these five meanings (transparency,

experimentation, customized actions, decision-making, and new business models). It seems clear that Laney's 3Vs theory describes the fundamental meanings of Big Data and structures a basis for further analysis, as the research showed that shipping professionals require more perspectives than just volume, velocity, and variety. Besides, they have introduced complexity and veracity as fundamental characteristics of PSC inspections Big Data.

Coming to an end with discussing the themes exported from the Thematic Analysis, it seems that shipping professionals, no matter their non-relevance with the IT concepts, found the need for characteristics specified by Laney's three Vs. theory (Laney, 2001). However, the professionals recognize the need to extract value for the PSC inspections Big Data, as McKinsey Global Institute stated (Manyika et al., 2011) for functionality (Grover et al., 2018).

The following and last chapter of this study will present the project's conclusions of how shipping professionals use PSC inspection databases and how they get value from Big platforms regarding the port state control inspections.

6. CONCLUSION

As mentioned in this study's format, the Conclusion chapter has been connecting with the Introduction. The conclusions' scope is to present briefly the project's procedures that followed and the outcomes that have been raised. The following two subchapters refer to the study's contribution and proposals for future research based on the assumptions or the limitations this study had and for evaluation reasons for the outcomes that had been concluding.

6.1 Conclusions

The recent technological expansion and innovative ideas bring new standards for organizations and companies. It is now possible to maintain large datasets within simple computer structures and analyze the collected data for extracting intelligence from them so that human decisions could be based more on data. Indeed, more and more information be collected and maintained in large datasets with plenty of structured or unstructured data; those formats are usually called Big Data.

In the last few years, Big Data implemented in the shipping sector for vessel operations; more data concerned the daily procedures that ships follow could be collected and organized. As it is rational, such databases and platforms are developed by the public and private sectors. Apart from the common interest these platforms have in shipping, these datasets mainly differ on the produced outcomes and the scopes; public datasets are mainly concerned with reporting what has taken place globally to inform the stakeholders, while the private ones have a commercial perspective. A typical example of such a database is the public databases of PSC inspections that MoU's or institutions, like Equasis, maintain for informative reasons.

The scope of this research was to investigate how Big Data could positively influence the shipping industry by examining how shipping professionals use them. We have tried to explore the shipping professionals' use of these public databases regarding the Port State Control (PSC) inspections Big Data and what characteristics and value they are looking for. In detail, as stated in the first chapter, this study aims to examine the three research questions for the utility of PSC inspection databases by the shipping professionals, focusing on characteristics, the value extracted from these, and how Big Data is used for the decision-making.

For the reasons mentioned above, we placed three research questions for the value, the Big Data characteristics, and PSC data within these datasets to assist them in the decision-making. Through this research, qualitative interpretative research was conducted to investigate shipping professionals' opinions and points of view. An interview guide was constructed based on the theories that received high appreciation in the academic literature – three Vs. theory (Laney, 2001) and McKinsey Global Institute on Big Data value (Manyika et al., 2011); so the data collection was done through semi-structured interviews. After the six interviews, we performed Thematic Analysis following Lichtman's (2013) three Cs methodology to explore their sayings and group their meanings to help us explore more efficiently and appropriately the theories discussed. The Thematic Analysis provides us with five concepts on "Knowledge", "Enhance operations", "Decision-making", "Data-driven", and "Challenges"; within those five concepts, the participants' sayings could be summarized and analyzed. With this procedure, it was feasible to come closer to our outcome with this study.

After completing the Thematic Analysis of the empirical findings was possible to answer the Research Questions that have been placed. The first RQ regarding the value that PSC inspection data could bring to a shipping company; the findings support that these databases provided valid and helpful information that broadened their understanding of the PSC inspections procedures. Through this, it is possible to upscale their internal operations. Regarding the second RQ about how the PSC data are used for decision-making, we have seen that the larger volume of data

assists them better as they can be informed about the maritime trends and make them more data-driven, therefore more accurate on their estimations and preparations. Finally, the third RQ on the Big Data characteristics of these PSC inspection datasets, the research shows that shipping professionals are looking for volume, velocity, variety, veracity, and complexity to gain as much as possible. At last, shipping professionals underlined the lack of complexity on the data and that this simplicity could not assist them in the best way.

In conclusion, it might be fair to say that through this study, we came closer to the needs of shipping professionals regarding the inspections from the authorities and explore their requirements for the Big Data from PSC inspection databases for enhancing their ability to make more data-driven decisions through their daily routine. The already existing datasets indeed assist shipping professionals much better than they are used without them; however, it is vital to upscale the Big Data's service derived from these public PSC inspection databases, helping the companies and society have more safety and secure vessels on the seas.

6.2 Contribution

This study attempts to understand better Big Data usage in a matter of characteristics they should have and the value extracted from them based on the shipping professionals' perspective. The practical contribution focus on the professionals who use such databases. By reading this study, a professional can better understand the PSC inspections operations and enhance his or her ability to make the right choices with the assistance of the data found on these platforms.

The research could enhance future research on Big Data's field regarding the characteristics and value extracted from data analytics. Although, the evaluation of two well-established theories on Big Data through the shipping professionals' opinion could be enabled academics to research in more depth the value of Big Data and Analytics in the maritime sector. Concluding, we could say that the study's outcome for the professionals' preferences on the Big Data characteristics in those PSC databases could provide an alternative combination of characteristics that would describe Big Data infrastructures.

6.3 Future Research

This study had some limitations based on the shipping professionals that have been interviewed and collected their sayings, as we already have discussed through this study. Many other professionals evolved in the shipping sector and PSC inspection databases, such as insurance companies or shipping agents and brokers. As a result, it would be beneficial for academic and business literature to have a broader view of this topic. Besides, apart from the qualitative design that this project had, it would be interesting to research a mixed-methods approach used as an evaluation tool for Big Data characteristics within these datasets. It is feasible with this procedure to generate a more structured theory of Big Data characteristics in PSC inspection databases if we have more professionals' opinions via interviews and questionnaires. Besides, interviews with more professionals would give a holistic idea regarding the PSC inspection outcomes databases. With the addition of questionnaires in future research, constructing a theory for the Big Data characteristics within these databases would be much closer to the industry needs. Finally, similar research within the private Big Data PSC inspections platforms would be helpful for the industry as it will make it possible to compare the results.

References

Beaver, D., Kumar, S., Li, H.C., Sobel, J, and Vajgel, P., 2010. *Finding a Needle in Haystack: Facebook's Photo Storage.* [pdf] Facebook. Available at: <https://www.usenix.org/legacy/event/osdi10/tech/full_papers/Beaver.pdf > [Accessed 17 May 2021].

Bertei, M., Marchi, L., and Buonchristiani, D., 2015. Exploring Qualitative Data: The use of Big Data technology as support in strategic decision-making. *The International Journal of Digital Accounting Research*, [e-journal] 15, pp. 99-126. 10.4192/1577-8517-v15_4.

Beynon-Davies, P., 2020. *Business Information Systems*. 3rd ed. London: RED GLOBE PRESS.

Brown, V. and Clarke, V., 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, [e-journal] 3(2), pp.77-101. 10.1191/1478088706qp063oa.

Brown, V. and Clarke, V., 2013. *Successful qualitative research: A practical guide for beginners*. London, UK: SAGE.

Braun, V. and Clarke, V., 2016a. (Mis)conceptualizing themes, thematic analysis, and other problems with Fugard and Potts' (2015) sample-size tool for thematic analysis. *International journal of social research methodology*, [e-journal] 19(6), pp.739-743. Available at: <https://doi-org.proxy.lnu.se/10.1080/13645579.2016.1195588> [Accessed 17 May 2021].

Brown, V. and Clarke, V., 2016b. Thematic analysis. *The Journal of Positive Psychology*, [e-journal] 12(3), pp.297-298. Available at: <https://doi-org.proxy.lnu.se/10.1080/17439760.2016.1262613> [Accessed 17 May 2021].

Chen, C.L.P. and Zhang, C-Y., 2014. Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, [e-journal] 275, pp.314-347. Available at: <https://doi-org.proxy.lnu.se/10.1016/j.ins.2014.01.015> [Accessed 17 May 2021].

Chen, J., Zhang, S., Xu, L., Wan, Z., Fei, Y., and Zheng, T., 2019. Identification of key factors of ship detention under Port State Control. *Marine Policy*, [e-journal] 102, pp.21-27. Available at: <https://doi-org.proxy.lnu.se/10.1016/j.marpol.2018.12.020> [Accessed 17 May 2021].

Creswell, J.W. and Creswell, J.D., 2018. *Research Design: Qualitative, Quantitative & Mixed Methods approaches*. 5th ed. Los Angeles: SAGE.

Cukier, K., 2010. Data, data everywhere: a special report on managing information. *Economist*. [online] Available at: <https://www.economist.com/special-report/2010/02/27/data-data-everywhere> [Accessed: 06 March 2021].

Curry, D.S., 2018. Interpretivism and norms. *Philosophical Studies*, [e-journal] 177, pp. 905-930. Available at: <https://doi-org.proxy.lnu.se/10.1007/s11098-018-1212-6> [Accessed: 18 May 2021].

DNV GL, n.d. Designated Person Ashore (DPA) Training Course. *DNV GL*. [online] Available at: <https://www.dnvgl.com/training/designated-person-ashore-dpa-training-course-11078> [Accessed: 17 February 2021].

Dutta, D. and Bose, I., 2015. Managing a Big Data project: The case of Ramco Cements Limited. *International Journal of Production Economics*, [e-journal] 165, pp. 293-306. Available at: <https://doi-org.proxy.lnu.se/10.1016/j.ijpe.2014.12.032> [Accessed: 18 May 2021].

Elgendy, N. and Elragal, A., 2016. Big Data Analytics in Support of the Decision Making Process. *Procedia Computer Science*, [e-journal] 100, pp. 1071-1084. 10.1016/j.procs.2016.09.251.

Elia, G., Polimeno, G., Solazzo, G., and Passiante, G., 2020. A multi-dimension framework for value creation through big data. *Industrial marketing management*, [e-journal] 90, pp. 508-522. Available at: <https://doi-org.proxy.lnu.se/10.1016/j.indmarman.2019.08.004> [Accessed: 18 May 2021].

Fu, J., Chen, X., Wu, S., Shi, C., Wu, H., Zhao, J., and Xiong, P., 2020. Mining ship deficiency correlations from historical port state control (PSC) inspection data. *PLoS ONE*, [e-journal] 15(2). https://doi.org/10.1371/journal.pone.0229211.

Geng, H., 2015. *Data center handbook*. Hoboken, NJ: John Wiley & Sons Inc.

Ghalehkhondabi, I., Ahmadi, E., and Maihami, R., 2020. An overview of big data analytics application in supply chain management published in 2010-2019. *Production*, [e-journal] 30. http://dx.doi.org/10.1590/0103-6513.20190140.

Goldkuhl, G., 2017. Pragmatism vs. interpretivism in qualitative information systems research. *European Journal of Information Systems*, [e-journal] 21(2), pp. 135-146. Available at: <https://doi-org.proxy.lnu.se/10.1057/ejis.2011.54> [Accessed: 18 May 2021].

Grandinetti, L., Joubert, G.R., and Kunze, M., 2015. *Big Data and High Performance Computing*. [e-book] Amsterdam: IOS Press BV. Available at: <https://lnu-se-primo.hosted.exlibrisgroup.com/primo-explore/search?vid=primo-custom-lnu&lang=en_US> [Accessed 18 May 2021].

Graziano, A., Cariou, P., Wolff, F.C., Mejia, M.Q., and Schroder-Hinriches, J.U., 2018. Port state control inspections in the European Union: Do inspector's number and background matter? *Marine Policy*, [e-journal] 88, pp.230-241. Available at: <https://doi-org.proxy.lnu.se/10.1016/j.marpol.2017.11.031> [Accessed: 18 May 2021].

Grover, V., Chiang, R.H.L., Liang, T-P., and Zhang, D., 2018. Creating Strategic Business Value from Big Data Analytics: A Research Framework. *Journal of Management Information Systems*, [e-journal] 35(2), pp.388-423. Available at: <https://doi-org.proxy.lnu.se/10.1080/07421222.2018.1451951> [Accessed: 18 May 2021].

Hassellov, I.M., Simopoulou, N., Sundblad, E.L., and Larsson, K., 2020. *Use of port state control inspection data from the Paris MoU to assess pressure from shipping on the marine environment: A study of ships operating in the Baltic Sea region. Lighthouse Swedish Maritime Competence Centre*. [pdf] Lighthouse, Swedish Maritime Competence Centre. Available at: <http://lnu.diva-portal.org/smash/get/diva2:1438927/FULLTEXT01.pdf> [Accessed: 06 February 2021].

Heij, C. and Knapp, S., 2018. Predictive power of inspection outcomes for future shipping accidents – an empirical appraisal with special attention for human factor aspects. *Maritime Policy & Management*, [e-journal] 45(5), pp.604-621. Available at: <https://doi-org.proxy.lnu.se/10.1080/03088839.2018.1440441> [Accessed: 18 May 2021].

Indian Register of Shipping (IRS), 2018. *Technical Circular No.: 041/2018*. [pdf] IRS. Available at: <https://www.irclass.org/technical-circulars/guidance-on-management-of-port-state-control-inspections/> [Accessed: 25 March 2021].

International Maritime Organization (IMO), n.d. *Port State Contro*l. [online] IMO. Available at: <https://www.imo.org/en/OurWork/MSAS/Pages/PortStateControl.aspx> [Accessed: 03 February 2021].

Johnson, R., 2017. *Why Port State Control is like toothpaste: You can't put it back in the tube*. [online] SAFETY4SEA. Available at: <https://safety4sea.com/port-state-control-like-toothpasteutm_sourcesafety4seautm_mediummajors/> [Accessed: 06 February 2021].

Intezari, A. and Gressel, S., 2017. Information and reformation in K.M. systems: big data and strategic decision-making. *Journal of Knowledge Management*, [e-journal] 21(1), pp. 71-91. 10.1108/JKM-07-2015-0293.

Kitchin, R. and McArdle, G., 2016. What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society*, [e-journal] 1(10). https://doi.org/10.1177/2053951716631130.

Knapp, S. and Franses, P.H., 2007. A global view on port state control: Econometric analysis of the differences across port state control regimes. *Maritime Policy and Management*, [e-journal] 34(5), pp.453-482. Available at: <https://doi-org.proxy.lnu.se/10.1080/03088830701585217> [Accessed: 18 May 2021].

Kowalska, K. and Peel, L., 2012. Maritime anomaly detection using Gaussian Process active learning. In: ISIF (International Society on Information Fusion), *15th International Conference on Information Fusion*. Singapore 9-12 July 2012. IEEE.

Kubina, M., Varmus, M., and Kubinova, I., 2015. Use of big data for competitive advantage of company. *Procedia Economics and Finance*, [e-journal] 26, pp. 561-565. https://doi.org/10.1016/S2212-5671(15)00955-7.

Lakoju, M. and Serrano, A., 2017. Saving costs with a big data strategy framework. In: IEEE, *2017 IEEE International Conference on Big Data*. Institute of Electrical and Electronics Engineers Inc., pp. 2340-2347. 10.1109/BigData.2017.8258188.

Laney, D., 2001. *3D Data Management: Controlling Data Volume, Velocity, and Variety*. Technical report, META Group.

Laney, D., 2012. *Deja VVVu: Gartner's Original" Volume-Velocity-Variety" Definition of Big Data*. [online] AIIM Community, Available at: <https://community.aiim.org/blogs/doug-laney/2012/08/25/deja-vvvu-gartners-original-volume-velocity-variety-definition-of-big-data> [Accessed: 20 March 2021].

Lavalle, S., Hopkins, M.S., Lesser, E., Shockley, R., and Kruschwitz, N., 2010. *Analytics: The new path to value*. [online] MIT Sloan Management Review. Available at: <https://sloanreview-mit-edu.proxy.lnu.se/projects/analytics-the-new-path-to-value/> [Accessed: 04 March 2021].

Lavalle, S., Lesser, E., Shockley, R., Hopkins, M.S., and Kruschwitz, N., 2010. *Big Data, Analytics, and the Path From Insights to Value*. [online] MIT Sloan Management Review. Available at: <https://sloanreview.mit.edu/article/big-data-analytics-and-the-path-from-insights-to-value/> [Accessed: 04 March 2021].

Lloyd, J., and Hopkins, P. (2015). Using interviews to research body size: methodological and ethical considerations. *Royal Geographical Society (with IBG)*, [e-journal] 47(3), pp. 305-310. Available at: < https://doi-org.proxy.lnu.se/10.1111/area.12199> [Accessed: 04 July 2021].

Lukoianova, T., and Rubin, V. (2014). Veracity Roadmap: Is Big Data Objective, Truthful and Credible? *Advances in classification research online*, [e-journal] 24(1). http://dx.doi.org/10.7152/acro.v24i1.14671.

Lichtman, M. (2013). *Qualitative research in education: a user's guide*. 3rd ed. Thousand Oaks: SAGE Publications, pp. 241-268.

Macionis, J. and Gerber, L., 2011. *Sociology*. 7th ed. Toronto: Pearson Canada.

Marine Digital. (n.d.). *Big Data in Maritime: How a shipping company can effectively use data*. [online] Marine Digital. Available at: <https://marine-digital.com/article_bigdata_in_maritime> [Accessed: 21 March 2021].

Mazzei, M.J. and Noble, D., 2017. Toward the development of a big data analytics capability. *Information Management*, [e-journal] 53(8), pp. 1049-1064. Available at: <https://doi-org.proxy.lnu.se/10.1016/j.im.2016.07.004> [Accessed: 18 May 2021].

Meskauskiene, V., Oorni, A., and Sell, A., 2019. Transparency driven public sector innovation: Smart waterways and maritime traffic in Finland. *Lecture Notes in Business Information Processing*, [e-journal] 341, pp.331-350. Abstract only. Available through: Springer Link <https://link.springer.com/chapter/10.1007/978-3-030-11395-7_27> [Accessed: 18 May 2021].

Mitroussi, K. (2004). Quality in shipping: IMO's role and problems of implementation. *Disaster Prevention and Management*, [e-journal] 13(1), pp.50-58. https://doi.org/10.1108/09653560410521698.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., and Byers, A.H., 2011. *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.

Myers, M.D. and Newman, M., 2007. The qualitative interview in IS research: examing the craft. *Information and Organization*, [e-journal] 17(1), pp. 2-26. Available at: <https://doi.org/10.1016/j.infoandorg.2006.11.001> [Accessed: 19 May 2021].

Nalchigar, S. and Yu, E., 2020. Designing Business Analytics Solutions: A Model-Driven Approach. *Business & information systems engineering*, [e-journal] 62(1), pp.61-75. Available at: <https://doi-org.proxy.lnu.se/10.1007/s12599-018-0555-z> [Accessed: 19 May 2021].

Needham, M., 2021. *IDC Forecasts Improved Growth for Global AI market in 2021*. [online] IDC. Available at: <https://www.idc.com/getdoc.jsp?containerId=prUS47482321> [Accessed: 19 May 2021].

Nowell, L.S., Norris, J.M., White, D.E., and Moules, N.J., 2017. Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *International Journal of Qualitative Methods*, [e-journal] 16(1). https://doi.org/10.1177/1609406917733847.

Oracle Corporation 2019. *What is Big Data?* [pdf] Oracle Corporation. Available at: <https://www.oracle.com/a/ocom/docs/what-is-big-data-ebook-4421383.pdf> [Accessed: 07 March 2021].

Palem, G., 2014. *Formulating an Executive Strategy for Big Data Analytics*. [pdf] Technology Innovation Management Review, 4(3), pp. 25-34. Available at: <https://www-proquest-com.proxy.lnu.se/publiccontent/docview/1614470827/fulltextPDF/C93127BDC1954B1APQ/1?accountid=14827> [Accessed: 19 May 2021].

Paris MoU, n.d. *Organization*. [online] Paris MoU. Available at: <https://www.parismou.org/about-us/organisation> [Accessed: 10 February 2021].

Pernice, D. and Debyser, A., 2020. *Maritime transport: traffic and safety rules*. [online] European Parliament. Available at: <https://www.europarl.europa.eu/factsheets/en/sheet/125/maritime-transport-traffic-and-safety-rules> [Accessed: 03 February 2021].

Piniella, F., Alcaide, JI, and Rodriguez-Diaz, E., 2020. Identifying stakeholder perceptions and realities of Paris MoU inspections. *WMU Journal of Maritime Affairs*, [e-journal] 19(1), pp.27-49. Available at: <https://doi-org.proxy.lnu.se/10.1007/s13437-020-00193-0> [Accessed: 19 May 2021].

Pies, K.H.,(2015. *Big Data Analytics: A practical guide for managers*. [e-book] Boca Raton, Florida: CRS Press. Available through: EBSCOhost Ebooks <https://web-a-ebscohost-com.proxy.lnu.se/ehost/detail/detail?nobk=y&vid=1&sid=1dd794f4-2b9e-45a7-8c49-f40a41ab1f7d@sdc-v-sessmgr01&bdata=JnNpdGU9ZWhvc3QtbGl2ZQ==#AN=957542&db=nlebk> [Accessed: 19 May 2021].

Qi, J., Zhang, Z., Jeon, S., and Zhou, Y., 2016. Mining customer requirements from online reviews: A product improvement perspective. *Information & management*, [e-journal] 53(8), pp. 951-963. Available at: <https://doi-org.proxy.lnu.se/10.1016/j.im.2016.06.002> [Accessed: 19 May 2021].

Ravira, F.J. and Piniella, F., 201). Evaluating the impact of PSC inspectors' professional profile: a case study of the Spanish Maritime Administration. *WMU Journal of Maritime Affairs*, [e-journal] 15(2), pp.221-236. Available at: <https://doi-org.proxy.lnu.se/10.1007/s13437-015-0096-y> [Accessed: 19 May 2021].

RightShip, 2018. *Paris MoU: withdrawal of data*. [online] RIGHTSHIP. Available at: <https://www.rightship.com/resources/news/paris-mou-withdrawal-of-data/> [Accessed: 02 March 2021].

SAS, n.d. *Big Data: What it is and why it matters*. [online] SAS. Available at: <https://www.sas.com/en_us/insights/big-data/what-is-big-data.html> [Accessed: 07 March 2021].

SAFETY4SEA, 2019. *Global PSC Scorecard: Detentions of 2019 First Semester*. [online] SAFETY4SEA Available at: <https://safety4sea.com/cm-global-psc-scorecard-detentions-of-2019-first-semester/> [Accessed: 03 April 2021].

Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., and Tufano, P., 2012. *Analytics: the real-world use of big data: How innovative enterprises extract value from uncertain data*, *Executive Report.* [pdf] IBM Institute for Business Value and Said Business School at the University of Oxford. Available at: < https://www.bdvc.nl/images/Rapporten/GBE03519USEN.PDF> [Accessed: 19 May 2021].

Shakhovska, N., Boyko, N., Zasoba, Y., and Benova, E., 2019. Big data processing technologies in distributed information systems. *Procedia Computer Science*, [e-journal] 160, pp. 561-566. https://doi.org/10.1016/j.procs.2019.11.047.

Sharda, R., Delen, D., and Turban, E., 2014. *Business Intelligence and Analytics: Systems for Decision Support.* 10th ed. London: Pearson Education Limited.

Steen, M. and Tanenbaum, A.S., 2016. A brief introduction to distributed systems. *Computing*, [e-journal] 98, pp. 967-1009. Available at: <https://doi-org.proxy.lnu.se/10.1007/s00607-016-0508-7> [Accessed: 20 May 2021].

Sollien, K., Jensen, T. E., and Schoyen, H., 2014. Designated person ashore-how do they perceive their role in the organization, and to which extent they are legally exposed in case of an accident? *Journal of Maritime Research*, [e-journal] 11(3), pp.3-12. Available at: <https://www.researchgate.net/publication/290844837_Designated_person_ashore-how_do_they_perceive_their_role_in_the_organization_and_to_which_extent_are_they_legally_exposed_in_case_of_an_accident> [Accessed: 20 May 2021].

Timmermans, S. and Tavory, I., 2012. Theory Construction in Qualitative Research: From Grounded Theory to Abductive Analysis. *Sociological Theory*, [e-journal] 30(3), pp.167-186. https://doi.org/10.1177/0735275112457914.

Tsou, M.C., 2018. Big data analysis of port state control ship detention database. *Journal of Marine Engineering & Technology*, [e-journal] 18(3), pp.113-121. Available at: <https://doi-org.proxy.lnu.se/10.1080/20464177.2018.1505029> [Accessed: 18 May 2021].

University of Southern California, 2021. *Research Guides*. [online] University of Southern California. Available at: <https://libguides.usc.edu/c.php?g=235034&p=1559822> [Accessed: 08 February 2021].

Vanek, O. Jakob, M., Hrstka, O., and Pechoucek, M., 2013. Agent-based model of maritime traffic in piracy-affected waters. *Transportation Research Part C: Emerging Technologies*, [e-journal] 36, pp.157-176. Available at: <https://doi-org.proxy.lnu.se/10.1016/j.trc.2013.08.009> [Accessed: 20 May 2021].

Varelas, T. and Plitsos, S., 2020. Real-Time Ship Management through the Lend of Big Data. In: Big Data Service, *Sixth International Conference on Big Data Computing Service and Applications*. Oxford, United Kingdom, 3-6 August 2020. 2020 IEEE.

Wang, S., Ji, B., Zhao, J., Liu, W., and Xu, T., 2018. Predicting ship fuel consumption based on LASSO regression. *Transportation Research Part D: Transport and Environment*, [e-journal] 65, pp.817-824. Available at: <https://doi-org.proxy.lnu.se/10.1016/j.trd.2017.09.014> [Accessed: 20 May 2021].

Walsham, G., 2006. Doing interpretive research. *European Journal of Information Systems*, [e-journal] 15(3), pp. 320-330. Available at: <https://doi-org.proxy.lnu.se/10.1057/palgrave.ejis.3000589> [Accessed: 20 May 2021].

Wilberg, J. Triep, I., Hollauer, C., and Omer, M., 2017. Big Data in Product Development: Need for a data strategy. *Portland International Conference on Management of Engineering and Technology: Technology Management for the Interconnected World, Proceedings*, pp. 1-10. Abstract only. Available through: IEEE Xplore <https://ieeexplore.ieee.org/document/8125460> [Accessed: 20 May 2021].

Wilkes, J., Hoover, C., Keer, B., Mehra, P., and Veitch, A., 2008. *Storage, data, and information systems.* 5th ed. Palo Alto, CA: Hewlett-Packard Laboratories.

Williams, M., 2000. Interpretivism and Generalisation. *Sociology*, [e-journal] 34(2), pp. 209-224. Available at: <https://doi-org.proxy.lnu.se/10.1177/S0038038500000146> [Accessed: 20 May 2021].

Yan, R., and Wang, S., 2019. Ship Inspection by Port State Control – Review of Current Research. *Smart Innovation, Systems and Technologies*, [e-journal] 149, pp.233-241. Abstract only. Available at: Springer <https://link.springer.com/chapter/10.1007/978-981-13-8683-1_24> [Accessed: 20 May 2021].

Yang, Z., Yang, Z., Yin, J., and Qu, Z., 2018. A risk-based game model for rational inspections in port state control. *Transportation Research Part E: Logistics and Transportation Review*, [e-journal] 118, pp.477-495. Available at: <https://doi-org.proxy.lnu.se/10.1016/j.tre.2018.08.001> [Accessed: 20 May 2021].

Yang, Z. Yang, Z., and Yin, J., 2018. Realizing advanced risk-based port state control inspection using data-driven Bayesian networks. *Transportation Research Part A: Policy and Practice*, [e-journal] 110, pp.38-56. Available at: <https://doi-org.proxy.lnu.se/10.1016/j.tra.2018.01.033> [Accessed: 20 May 2021].

Yuan, C.C., Chung, W.H., Cai, C., and Sung, S.T., 2020. Application of Statistical Process Control on Port State Control. *Journal of Marine Science and Engineering*, [e-journal] 8(10), p.746. https://doi.org/10.3390/jmse8100746.

Yuan, C.C., Chiu, R.H., and Cai, C., 2020. Important Factors Influencing the Implementation of Independent Port State Control Regimes. *Journal of Marine Science and Engineering*, [e-journal] 8(9), p.641. https://doi.org/10.3390/jmse8090641.

Yudono, A., 2018. *Magnifying Democracy and Sovereignty in Indonesian Maritime Governance Through Open Marine Spatial Data Practice*. [pdf] IOP Conference Series: Earth and Environmental Science, 156(1). Available at: <https://iopscience.iop.org/article/10.1088/1755-1315/156/1/012030/pdf> [Accessed: 20 May 2021].

Yuhong, W., Fan, Z., Zhisen, Y., and Zaili, Y., 2021. Incorporation of deficiency data into the analysis of the dependency and interdependency among the risk factors influencing port state control inspection. *Reliability engineering & system safety*, [e-journal] 206. Available at: <https://doi-org.proxy.lnu.se/10.1016/j.ress.2020.107277> [Accessed: 20 May 2021].

Zhang, G. and Thai, V.V., 2016. Expert elicitation and Bayesian Network modeling for shipping accidents: A literature review. *Safety Science*, [e-journal] 87, pp.53-62. Available at: <https://doi-org.proxy.lnu.se/10.1016/j.ssci.2016.03.019> [Accessed: 20 May 2021].

Appendices

Appendix A – Interview Guide (English)

1. How does your company use the PSC data provided by public sources?
2. What can you say about getting information regarding shipping industry trends worldwide through bid data platforms?
3. How the big data utilization supports or can be connected to innovation strategies in your company?
4. How are you feeling about the transparency of the PSC inspection information globally?
5. How the big data provided help or support the decision-making in your company?
6. What do you think should change regarding the data provided that could better support decision-making?
7. What do you think about the volume of the PSC inspection data provided by the public databases? What should change?
8. How do you think about the PSC inspection database updating procedures? What should change?
9. How do you feel about the variety of information provided through open platforms? What should change?

Appendix B – Interview Guide (Greek)

1. Πως η εταιρία σας χρησιμοποιεί τα δεδομένα από τις επιθεωρήσεις των λιμενικών αρχών που δίνονται από τις δημόσιες πηγές;
2. Τί έχετε να σχολιάσετε σχετικά με την πρόσβαση και απόκτηση πληροφοριών σχετικά με τις τάσεις τις ναυτιλίας από πλατφόρμες που χρησιμοποιούν μεγάλα δεδομένα;
3. Πως η χρησιμοποίηση μεγάλων δεδομένων υποστηρίζει ή μπορεί να συνδεθεί με αναπτυξιακές στρατηγικές μέσα στην εταιρία σας;
4. Πως αισθάνεστε για την διαφάνεια των δεδομένων από τις επιθεωρήσεις των λιμενικών αρχών σε παγκόσμιο επίπεδο;
5. Πως τα μεγάλα δεδομένα βοηθάνε ή υποστηρίζουν τις διαδικασίες αποφάσεων στην εταιρία σας;
6. Τι θεωρείτε ότι πρέπει να αλλάξει σχετικά με την παρεχόμενη πληροφόρηση από τις υφιστάμενες πήγες για την καλύτερη υποστήριξη των διαδικασιών αποφάσεων στην εταιρία σας;
7. Ποια είναι η γνώμη σας σχετικά με τον όγκο της παρεχόμενης πληροφόρησης από τις δημόσιες πηγές για τις επιθεωρήσεις των λιμενικών αρχών; Τί θα θέλατε να αλλάξει;
8. Ποια είναι η άποψη σας σχετικά με τη συχνότητα επικαιροποίησης των βάσεων δεδομένων για τα αποτελέσματα των επιθεωρήσεων των λιμενικών αρχών; Τί θεωρείτε ότι πρέπει να διαφοροποιηθεί;
9. Πως αισθάνεστε για την ποικιλία των παρεχόμενων πληροφοριών από τις υφιστάμενες πλατφόρμες; Τί θα μπορούσε να αλλάξει;

Appendix C – Informed Consent Form for the interviews

Linnaeus University
Department of Informatics
Master Program in Information Systems

Informed Consent Form

Title of research: Big Data usage in the Maritime industry: A Qualitative Study for the use of Port State Control (PSC) inspection data by shipping professionals.

Researcher: Dimitris Ampatzidis
Email: da222td@student.lnu.se | d.ampatzidis@hotmail.gr
Tel: +30 698 0055 568

Supervisor: Konstantina Pentarchou
Email: Konstantina.pentarchou.extern@lnu.se

Aim: The purpose of this qualitative study is to explore the shipping professionals' point of view on PSC inspection data and how they work with the provided information to extract intelligence and value from these data, enhancing the decision-making procedure in the shipping companies; thus, this research conducted as a part of the Master Thesis mentioned above.

Procedure: Your participation in this research includes:
- Brief presentation of the study's purpose
- An interview will last for about half an hour; the interview will be recorded under your permission.

Dangers: The participation in this research does not include any danger, while the interviews' material will be used exclusively by the researcher for his research purposes.

Confidentiality: The information you are going to give will be held safe and confidential by the researcher. Only the researcher and his supervisor will have access to the research data. All the data of the research will be held until its completion and will be deleted afterward. Your name will not be revealed, and it will not be linked with any of the findings. After the completion of the procedure, the recording and the consistent form would be destroyed.

Voluntary participation and the right of withdrawal: Your participation in this research is voluntary. You can refuse to participate or withdraw at any time, even during the interview. This decision will not affect your relationship with the department personnel.

Right to ask questions: You have the right to ask any questions by contacting the researcher before, during, and/or after the interview.

Consent: *I agree to participate in the research: " Big Data usage in the Maritime industry: A Qualitative Study for the use of Port State Control (PSC) inspection data by shipping professionals", conducted by Dimitris Ampatzidis.*

I agree to be audio-recorded (check the box):

☐Yes
☐No

***I understand the purpose of this research and my rights*.**

Date:

Participant:

Appendix D – Thematic Analysis

Raw Data	Codes	Categories	Concepts
6x Interviews	Inform about the global trends	Information	**Knowledge**
	Look for fleet-based issues	Preparation	**Enhance operations**
	Educating reasons	Training	**Enhance operations**
	Management Reviews	Reporting	**Decision-making**
	Prepare the fleet based on inspection reports from these databases	Preparation	**Enhance operations**
	Find the KPIs to compare with company's	KPIs	**Data-driven**
	Transparency on big data for PSC inspections	Information	**Knowledge**
	Different characteristics of data could bring different decisions	Decisions	**Decision-making**
	Frequency of the updates	Updates	**Challenges**
	More depth and complexity on the data	Lack of data	**Challenges**
	Common language among databases	Understanding	**Challenges**
	Not easily accessible information within the datasets	Access	**Challenges**
	Published procedures are helpful to understand their operations	Understanding	**Knowledge**
	Insightful intelligence from the public databases	Decisions	**Decision-making**
	PSC big data and risk management	KPIs	**Data-driven**
	Data-driven decisions	Decisions	**Decision-making**
	Variety of the information	Information	**Knowledge**
	Data able to use in case studies	Training	**Enhance operations**
	Databases as they are now, they do not provide assistance to decisions	Decisions	**Decision-making**
	Many regimes and MoUs do not provide any database	Lack of data	**Challenges**
	Paris MoU presents only detentions and that present half of the picture	Lack of data	**Challenges**
	The datasets do not have filters so you cannot access specific information	Access	**Challenges**
	22 Codes	***10 Categories***	***5 Concepts***

Lnu.se

Linnæus University
Sweden

Faculty of Technology
SE-391 82 Kalmar | SE-351 95 Växjö
Phone +46 (0)772-28 80 00
teknik@lnu.se
Lnu.se/fakulteten-for-teknik