Mälardalen University
School of Innovation Design and Engineering
Västerås, Sweden

Thesis for the Degree of Master of Science in Computer Science
120.0 credits

CREATION OF AN OPERATIONAL DESIGN DOCUMENT FOR THE AUTONOMOUS SHIPPING INDUSTRY

Samir Huseynzade
she22005@student.mdu.se

Examiner: Cristina Seceleanu
Mälardalen University, Västerås, Sweden

Supervisor: Peter Backeman
Mälardalen University, Västerås, Sweden

Company supervisor: Pontus Boström,
Groke Technologies, Turku Finland

October 4, 2023

Abstract

The advancement of Artificial Intelligence technologies has paved the way for the increasing use of autonomous systems in various fields, including air, land, and sea. Maritime is an important domain for applying autonomous systems, as collisions at sea can result in significant losses. However, the development of autonomous systems in the maritime domain involves the collection of vast amounts of data from sensors on vehicles, which poses significant challenges in understanding the types of data collected from vehicle sensors and the corresponding data system requirements due to the large amount of data collected. To address this issue, this thesis proposes the use of Operational Design Domain (ODD), inspired by the autonomous car domain, to structure and manage the data. An ODD taxonomy was designed with the help of BSI PAS 1883 and ASAM OpenODD concepts. This thesis presents an initial draft of the ODD taxonomy for the maritime domain, followed by the implementation of ODD samples on actual data. The work involved converting unstructured data into ODD-friendly data. The results of this thesis include an ODD taxonomy, a framework, and a use case demonstrating the application of these concepts to actual data. These results provide an example of how to manage large amounts of data in the autonomous maritime environment and facilitate further testing of autonomous systems in this domain.

Table of Contents

1. Introduction

In the maritime industry, the implementation of safe autonomous systems is of great interest due to the high cost of collisions in the maritime domain. Additionally, human operators face difficulties monitoring vessels in extreme weather conditions, which can lead to potentially dangerous situations [1]. In order to achieve autonomy in the maritime environment, the first step is to provide a situational awareness system [2], which is offered by companies such as Groke Technologies [3].

Situational awareness in the maritime domain is attained through the installation of sensors onboard the vessel. These sensors serve the purpose of perceiving the surrounding environment and gathering relevant data. By capturing various types of information, such as the vessel's position, speed, heading, weather conditions, and nearby objects, these sensors provide a comprehensive view of the vessel's surroundings at any given moment. The collected data from these sensors is then processed and analyzed to extract meaningful insights about the maritime environment. Through advanced algorithms and techniques, the sensor data is interpreted and transformed into actionable information. This information can include real-time updates on the vessel's position in relation to navigational hazards, other vessels, and critical landmarks, as well as weather conditions, sea state, and potential risks. The goal of situational awareness is to enhance the understanding of the vessel's operating context, allowing the crew, operators, or autonomous systems to make informed decisions and take appropriate actions to ensure the safety and efficiency of maritime operations. By continuously monitoring and assessing the surrounding environment through sensor data, situational awareness enables early detection of potential risks, facilitates proactive navigation, and supports effective response to changing conditions.

However, the sensor units in a vessel collect a large amount of data, including information from cameras, GNSS, radar, and LIDAR sensors, which must be fused to make sense of the environment in which the vessel is operating [4]. As the amount of data collected by various sensors increases, the task of understanding the content of the collected data becomes more challenging. Difficulties begin to arise in determining what type of data has been collected and what further data is needed to improve the system's performance. This problem limits the system from exploring and utilizing its collected data to understand its system better and improve its functionality. Unfortunately, there are not yet any publicly available solutions to this problem in the maritime industry.

To address this issue, this thesis explores solutions from the autonomous car domain, which has already produced production-level autonomous systems. In this research, the Operational Design Domain (ODD) concept was employed, which is utilized in the autonomous car domain for describing the environment in which the vehicle operates [5]. This concept enables development companies to define the domains in which their system will operate and employ these ODDs to analyze their dataset's diversity. This also helps them to identify areas where their system may be lacking in certain domains and gain valuable insights.

The adaptation of the ODD concept in this thesis offers a new perspective on managing data for autonomous systems in the maritime environment. By utilizing the ODD concept, companies and researchers can better structure and store their data in a scalable manner, enabling future analytics to be performed on the data to test the effectiveness of their systems. This, in turn, can lead to the development of better autonomous systems and situational awareness technologies for the maritime environment.

The ODD taxonomy and framework presented in this thesis can serve as a valuable resource for researchers and companies working on autonomous vehicles or situational awareness systems in the maritime environment. The taxonomy offers a clear and concise description of the maritime domain, providing a foundation for the development of autonomous systems that can operate safely and efficiently in this environment. Additionally, the framework provides guidance on how to apply the ODD concepts to actual data, helping to alleviate the complex data issues that arise from collecting and storing large amounts of data from sensors.

Furthermore, the authorities responsible for creating standards for autonomous vessels can also benefit from this thesis. The discussion of design documents in this paper can serve as a starting point for developing standards that are specifically tailored to the maritime environment. By using the ODD concept, authorities can provide guidelines for the design and operation of autonomous vessels, ensuring that they meet safety and performance standards.

The focus is on creating an Operational Design Document (ODD) taxonomy for the maritime domain, which can be used to describe the environment and state of the ship. The taxonomy provides a structured way of categorizing and organizing the data collected by the sensor units on a vessel, which can be used to facilitate the analysis of the data. The thesis aims to demonstrate how the ODD taxonomy would work on actual data, by using example data from Groke Technologies, a company that provides situational awareness systems for the maritime environment.

The thesis recognizes the significance of demonstrating the practical utilization of the ODD concept. It aims to showcase how the ODD framework can be effectively applied in real-world scenarios. To achieve this, the thesis utilizes data provided by Groke Technologies, a specific case study that allows for the practical implementation and evaluation of the ODD concept. By working with actual data, the thesis provides tangible evidence of the applicability and usefulness of the ODD framework in a practical setting.

In addition, the thesis acknowledges that applying the ODD concept is not without its challenges. It recognizes that there may be obstacles and complexities associated with implementing the ODD framework in practice. The thesis takes a proactive approach by acknowledging these challenges and making them a central focus of the research. By identifying and addressing the challenges, the thesis aims to enhance the effectiveness and reliability of the ODD concept in practical applications.

This thesis makes several contributions to the field, which are outlined below:

- Structured way for describing collected sensor data for situational awareness systems: One of the key contributions of this thesis is the development of a structured approach for describing the collected sensor data in the context of situational awareness systems. By defining a systematic framework, this thesis enables a more organized and standardized representation of sensor data, ensuring consistency and facilitating effective analysis and interpretation of the data. This structured approach enhances the overall understanding of the maritime environment and supports better decision-making processes.

- Initial design of ODD taxonomy for the maritime environment: Another important contribution of this thesis is the creation of an initial design for the ODD taxonomy specific to the maritime domain. ODDs play a crucial role in defining the operational boundaries and characteristics of autonomous systems, including ships. By designing an ODD taxonomy tailored to the maritime environment, this thesis provides a foundation for accurately describing and categorizing the operating conditions and constraints of autonomous ships. This taxonomy serves as a valuable resource for future work in creating ODDs for specific maritime applications.

- Use case demonstration of exploring sensor data using ODD: Additionally, this thesis presents a practical use case demonstration of exploring sensor data using the ODD framework. By applying the designed ODD taxonomy to a specific dataset, the thesis showcases how ODDs can be utilized to effectively analyze and visualize sensor data for situational awareness purposes. This demonstration highlights the potential of ODDs in enhancing the understanding of sensor data and extracting valuable insights about the maritime environment. It serves as a practical example of how the developed structured approach and ODD taxonomy can be applied in real-world scenarios.

2. Background

The maritime industry, often regarded as one of the oldest and most significant industries in the world, plays a vital role in global trade, transportation, and economic development. It encompasses a wide range of activities related to the navigation, operation, and management of ships and vessels across seas, oceans, and waterways. The maritime industry is characterized by its complexity, as it involves the navigation of diverse types of ships, such as cargo vessels, tankers, cruise ships, fishing boats, and offshore platforms, among others. In addition to the diversity of ships, the maritime industry is confronted with numerous obstacles and challenges. Navigating through busy shipping lanes, congested ports, and narrow waterways requires skilled seamanship and adherence to international maritime regulations. Furthermore, the industry must address environmental considerations, including weather conditions, sea states, and the protection of marine ecosystems. Moreover, the maritime industry is subject to various risks and hazards, such as collisions, grounding, piracy, and natural disasters. The safety and security of vessels, crew members, cargo, and the marine environment are of paramount importance in this industry. Therefore, comprehensive navigation systems, communication protocols, and operational procedures are implemented to mitigate risks and ensure the smooth and efficient movement of ships. Furthermore, the maritime industry serves as a critical enabler of international trade, facilitating the transportation of goods and commodities across continents. Given the vital role it plays, the maritime industry is subject to continuous advancements and innovation. Emerging technologies, such as autonomous vessels, unmanned surface vehicles, and digitalization, are reshaping the industry and offering new opportunities for efficiency, safety, and sustainability. These developments require a comprehensive understanding of the complex maritime environment and the implementation of advanced systems and practices.

The maritime industry is beginning to put efforts into autonomous shipping [6]. There are several researchers and companies that are working on creating solutions to enable autonomous shipping using Artificial Intelligence(AI) approaches. To use AI systems to their full capacity, a substantial amount of data needs to be collected. Thus, data collection and data analysis play a huge role in the performance of these systems [7]. Data collection results in a massive volume of data and making sense of the collected data becomes a complicated task. In a maritime environment, the collected data is usually a mix of outputs from multiple sensors [8]. These sensors encompass a range of technologies, including Lidar (Light Detection and Ranging) [9], Radar (Radio Detection and Ranging) [10], Camera, GNSS (Global Navigation Satellite System) [11], AIS (Automatic Identification System) [12] receivers, and others [1]. In a complex system, categorizing the sensor outputs and analyzing the content of the collected data becomes a challenge. To systematically group the collected data and categorize it, there is a need for a scalable technique. While this topic hasn't been extensively researched in the maritime environment due to the recent interest in autonomous shipping within the industry, it has been well-studied in the autonomous car industry. In the autonomous car industry, this challenge is addressed by employing Operational Design Domains (ODDs) to define the specific operating environment for the vehicles.

2.1. Operational Design Domain

Autonomous vehicles have been at the center of the applications of artificial intelligence technologies. One of the main use cases of autonomous vehicles is the smart cities [13]. In today's world, autonomous cars have gained significant recognition as one of the most prominent types of autonomous vehicles. Extensive research and the efforts of various companies have yielded promising results in the development and implementation of autonomous car technology [14]. When it comes to autonomy in cars, there are six levels of autonomy, from level 0 to level 5 where level 0 is when the system provides only information to the driver like warnings and signals and level 5 is when the vehicle is fully automated and it can drive itself without human intervention [15].

The testing and validation of autonomous vehicles pose a significant challenge, given that the environments in which these vehicles operate are subject to variation. It is not guaranteed that an autonomous vehicle that has been developed and tested in one country, such as the USA, will operate as intended in another country, such as Germany, due to country-specific characteristics. Therefore, it is necessary to define the expected behavior of autonomous vehicles in different domains. ODD is defined as “operating conditions under which a given driving automation system or feature thereof is specifically designed to function, including, but not limited to, environmental, geographical, and time-of-day restrictions, and/or the requisite presence or absence of certain traffic or roadway characteristics.” [16]. ODD is widely employed to define the boundaries of vehicles in different domains and to query scenario databases, which are collections of various scenarios gathered during autonomous vehicle development.

Generally, ODDs are expressed with three main attributes: Scenery, Environmental Conditions, and Dynamic Elements [17]. These top-level attributes can have nested structures meaning several iterations of sub-attributes are allowed. Each attribute focuses on describing the domain from a different perspective. *Scenery* describes the static elements in the designed domain, such as attributes describing navigable area structures or domain-specific elements. *Environmental Conditions* describe the environmental elements that may affect the vehicle’s state, such as weather or illumination. Finally, *Dynamic Elements* describe the subject vehicle or the movements of other objects.

Multiple approaches exist for designing the ODD of the system. Standards such as Asam ODD [18], BSI PAS [17], Wise [19], and others address the ODD design process, each offering distinct perspectives on tackling the problem. One of the examples of ODD based on the Wise standard can be seen on the Listing 1.

```yaml
Operational_Design_Domain:
  Road_Structure:
    Type: Highway
    Number_of_Lanes: 2
    Speed_Limit: 80 mph
  Road_Users:
    User_Type: Commuter
  Environmental_Conditions:
    Time_of_Day: Rush_Hour
    Weather: Cloudy
    Visibility: Moderate
  Animals:
    Type: None
  Road_Objects:
    Type: Traffic_Signals
  Pedestrians:
    Presence: None
```

Listing 1: Example ODD document based on Wise standard

The sample ODD provided above is visualized in YAML format and utilizes the taxonomy offered by the WISE lab. It categorizes the surrounding information into different aspects, including “Road_Structure“, “Road_Users“, “Environmental_Conditions“, “Animals“, and other categories. Under the “Road_Structure“ category, the ODD specifies that the road type is a highway with 2 lanes and a speed limit of 80 miles per hour. The “Road_Users“ category describes the type of user. The “Environmental_Conditions“ category indicates that the time of day is rush hour, the weather is cloudy, and the lighting is natural. The “Animals” category specifies that there are no animals present in the operational design domain. The “Other_Objects“ section includes the presence of traffic signs and no pedestrians in the surrounding environment. This sample ODD demonstrates the representation of different aspects of the operational design domain

2.2. Groke Technologies

The focus of this thesis is on designing an ODD for the autonomous shipping industry, with the help of the company where the research was conducted. The research was carried out in collaboration with Groke Technologies, a company that specializes in the development of Artificial Intelligence-based situational awareness systems for maritime vehicles. Groke Technologies specializes in the development of situational awareness systems designed to enhance the working capabilities of crew members in the maritime industry. Groke Technologies was established in October 2019 in Finland with the aim of becoming the leading provider of state of the art situational awareness systems and related autonomous navigation solutions and helping our customers to move to the forefront of maritime digitalization.” [3] Their flagship product in the situational awareness solution domain is known as Groke Pro.

Groke Pro is a comprehensive system that comprises several components working in harmony to provide an advanced situational awareness experience. The system consists of a sensor unit, a central unit, and an application deployed on a tablet device for visualizing the collected information. The sensor unit serves as the primary data collection mechanism, employing a range of sensors to gather pertinent information about the vessel’s surroundings. These sensors include GNSS (Global Navigation Satellite System) receivers for accurate positioning, an inertial measurement unit for measuring vessel dynamics, an AIS (Automatic Identification System) receiver for vessel identification and tracking, as well as both thermal and daytime cameras for visual monitoring of the environment.

The central unit plays a crucial role in processing the data collected by the sensor unit. It performs various computational tasks, such as data fusion, filtering, and analysis, to extract meaningful insights from the raw sensor data. By employing sophisticated algorithms and techniques, the central unit transforms the collected data into actionable information that can be effectively utilized by the crew members.

The tablet application acts as the user interface for the Groke Pro system, providing a map-based visualization of the gathered information. The application presents the processed data in a user-friendly and intuitive manner, enabling crew members to easily comprehend and interpret the situational awareness information. Through the tablet application, users can access real-time updates on vessel positions, nearby objects, potential hazards, and other relevant information necessary for safe and efficient navigation. These components can be seen from the Figure 1.

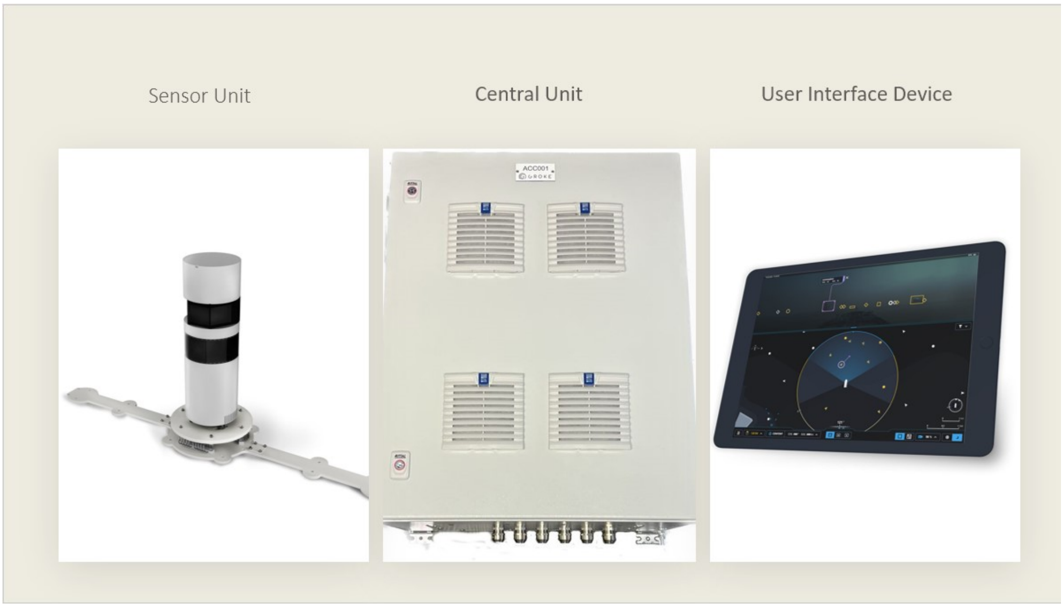

Figure 1: Groke Pro product components [20]

Throughout the thesis, the data collected from the Groke Pro product was utilized to analyze and explore various aspects related to operational design domains and situational awareness in the maritime industry.

2.3. Elastic Search

In this thesis, Elastic Search was employed due to its existing utilization by the company for data storage and querying purposes. Elastic Search is a distributed multi-tenant, full-text search engine [21]. This engine offers rapid data search capabilities. Data is stored as JSON objects referred to as documents [22], which can be organized and stored within an index by grouping them together. These indexes are split into shards for scalability purposes visualized in Figure 2 [23].

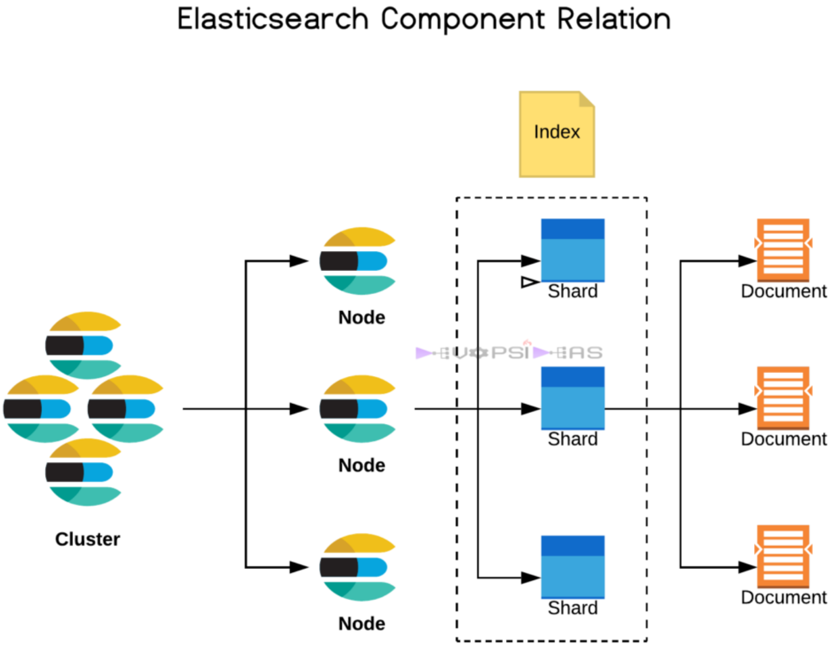

Figure 2: Elastic Search Architecture [24]

When working with Elastic Search, querying documents from an index is done using Elastic Search Query Language, which is a powerful and flexible JSON-based language specifically designed for searching and retrieving data from Elastic Search indexes. The Elastic Search Query Language documentation, available at the [25], provides detailed information on the syntax, usage, and available options. The provided example taken from documentation [26] query showcases the usage of the ”match” keyword in the Elastic Search Query Language. The query aims to retrieve documents from the index based on the presence of the term ”daytime” in the ”message” field.

```
GET /_search
{
  ”query”: {
    ”match”: {
      ”message”: {
        ”query”: ”daytime”
      }
    }
  }
}
```

By executing the above query, Elastic Search will search through the specified index for documents where the ”message” field contains the term ”daytime”. The search results will include all the relevant documents that match this criterion. Similar queries were used in this thesis to extract relevant information from the dataset.

Listing 2 illustrates the creation of complex queries using the Elastic Search query language. The query demonstrates how to retrieve specific entries from the dataset based on multiple criteria, showcasing the flexibility of Elastic Search for nuanced data retrieval.

In this example, the query aims to fetch entries that match various conditions: it seeks entries from the nighttime, specifically between January 1, 2023, and July 1, 2023. Moreover, it excludes entries with rainy weather conditions and gives preference to those situated in harbor or coastal areas, indicating navigable areas of potential interest.

The query structure is nested within a ”bool” clause, where ”must” conditions are used for mandatory matches, ”should” conditions indicate optional matches, and ”must_not” conditions exclude specific matches. The final retrieved entries will be the ones that meet the defined criteria.

Additionally, the query restricts the result set to 10 entries and sorts them in descending order based on the ”updated” field. Elastic Search assigns scores to results based on their match to the query conditions, and the retrieved entries are presented in the order of descending scores.

```
GET /_search
{
  ”query”: {
    ”bool”: {
      ”must”: [
        { ”match”: { ”time_of-day”: ”night” } },
        { ”range”: { ”updated”: { ”gte”: ”2023-01-01”, ”lt”: ”2023-07-01” } } }
      ],
      ”should”: [
        { ”match”: { ”navigable_Area_Type”: ”Harbor” } },
        { ”match”: { ”navigable_Area_Type”: ”Coastal Area” } }
      ],
      ”must_not”: [
        { ”match”: { ”weather”: ”rain” } }
      ]
    }
  },
  ”size”: 10,
  ”sort”: [
    { ”updated”: { ”order”: ”desc” } }
  ]
}
```

Listing 2: Complex query using Elastic Search Query Language

3. Related Work

Related works have been conducted in the field of autonomous cars with the aim of describing data in various ways using ODD. A paper summarizing the widely used ODD designs can be found in [27]. In this thesis, the focus is on analyzing four main works - BSI PAS 1883, WISE, AVSC, and OpenODD - that describe the domain in different ways for autonomous cars.

3.1. British Standards Institution

BSI PAS 1883 [17], which is provided by the British Standards Institution, is focused on providing a definition for ODD in the context of road infrastructure in Britain. The standard provides a detailed taxonomy and presents the ODDs in a tabular format. It includes several attributes to describe the domain, including road type, road class, lane type, and weather. The work emphasizes the importance of unambiguous descriptions of ODDs to ensure the safe operation of autonomous cars. BSI PAS 1883 not only focuses on defining ODD for the context of road infrastructure in Britain but also emphasizes the relationship between stakeholders and ODD. The ODD is structured into three main attributes: Scenery, Environment, and Dynamic Elements. The Scenery attribute describes the physical characteristics of the environment that the vehicle operates in, such as the road geometry, lane markings, and road signs. The Environmental attribute describes the external factors that may affect the vehicle's operation, such as weather and lighting conditions. The Dynamic Elements attribute describes the behavior of other road users, such as pedestrians, cyclists, and other vehicles. This thesis is based on BSI PAS 1883 to develop the ODD taxonomy for the maritime environment. While BSI PAS 1883 was developed with the context of road infrastructure in mind, the general structure of Scenery, Environment, and Dynamic elements can be applied to the maritime domain as well. Therefore, this thesis is based on a similar structure to BSI PAS 1883 to describe the ODD taxonomy for the maritime environment, but with the necessary modifications to map the domains to the maritime context. The mapping between the road and maritime domains is explained in detail in section 7.2., which includes the different types of entities in each domain and how they are related. By using BSI PAS 1883 as a reference, the ODD taxonomy in the maritime environment was developed in a structured and organized manner, which provides a better understanding of the environment in which the vessel is operating.

3.2. Waterloo Intelligent Systems Engineering lab

The WISE standard is a comprehensive taxonomy developed by the Waterloo Intelligent Systems Engineering (WISE) lab at the University of Waterloo in Canada [19]. This standard aims to provide a framework for the on-road safety of Automated Driving Systems (ADS) by defining the ODD of these systems. The WISE standard takes into account various aspects of the ADS, such as sensor capabilities, functional limitations, and environmental conditions. It also includes a taxonomy of the interactions between the ADS and other road users, such as pedestrians, cyclists, and other vehicles. The WISE standard uses the Operational Road Environment Model (OREM) to describe the environment in which an automated driving system operates. The OREM is designed to include only the necessary details of the environment that are relevant to the system's operation while excluding unnecessary details that can potentially overload the system. This approach aims to provide a simplified and manageable representation of the environment to the system. While it provides useful concepts such as the Operational Road Environment Model (OREM), which excludes unnecessary details about the environment, the standard's focus on road structure and specific objects like users and animals makes it more difficult to map to the maritime environment.

3.3. Automated Vehicle Safety Consortium

Automated Vehicle Safety Consortium (AVSC) is an industry program of SAE Industry Technologies Consortia [28]. The AVSC standard focuses on providing a conceptual framework and lexicon for describing an ODD for automated vehicles [29]. The goal of this standard is to promote safe automated vehicle deployment by providing best practices for ODD description. The standard suggests a bottom-up approach to the design of the domains which emphasizes identifying the capabilities and limitations of the autonomous vehicle first, followed by identifying the environmental and operational factors. The AVSC standard mainly focuses on the on-road environment and provides details on road structures, road surface conditions, and weather conditions. However, this standard may not be the best fit for describing the maritime domain. This is because the AVSC standard is developed primarily for the on-road environment, and it may not cover all the necessary factors that are unique to the maritime domain.

3.4. The Association for Standardization of Automation and Measuring Systems

The Association for Standardization of Automation and Measuring Systems (ASAM) which is an international non-profit association that develops standards for the automotive industry published ASAM OpenODD Concept Paper for the design of ODD [18]. As it is a concept paper, it does not provide a concrete implementation or standard for the ODD design. Therefore, the validity of the paper might be questioned as it is not a fully developed standard. However, it brings some innovative ideas to the table that can be incorporated into the ODD design. When designing operational design domains, one of the proposed concepts in ASAM OpenODD is to incorporate ontologies [30]. The ASAM OpenODD Concept Paper emphasizes the importance of creating operational design domains that can work in both simulation and real systems. This means that the ODDs should be programmable, adaptable, and flexible enough to be used in various environments and situations. The paper discusses the limitations of using a tabular approach for describing ODDs and suggests the use of OpenSCENARIO 2.0 language [31] instead. OpenSCENARIO is a language for describing complex traffic scenarios in a machine-readable way and is also developed by ASAM. Unlike other papers that mainly focus on the design and lexicon of the ODD, this paper focuses on the representation of the ODD for both simulation and real systems. In this thesis, many aspects of this concept paper were utilized, such as defining the domain in a machine-readable format, which is a crucial aspect of programmable ODDs.

3.5. Operational Envelope

There has been some initiative on creating the ODD concept in the maritime domain as well. Operational Envelope (OE) concept has been introduced by Rødseth, Lien, Lars, and Nordahl [32]. The OE concept addresses the limitations of ODDs, particularly their failure to account for human intervention. The paper proposes that the OE concept can cover this gap by considering the autonomous system's capabilities and limitations, environmental factors, and the presence of human intervention in the decision-making process. The paper discusses the concept of the Operational Envelope (OE) in the maritime domain as a complementary concept to the ODD. While the ODD focuses on the technical capabilities and limitations of an automated system, the OE considers human intervention in the control of the vessel. However, for the specific problem of describing the domain, the paper does not provide significant input, as it does not offer a taxonomy or a structured framework for describing the OE.

In addition, while the concept of Operational Envelope is relevant and provides insights into the design process, it differs from the specific focus of this thesis, which is the application of ODDs to analyze collected sensor data for a navigational awareness system in the maritime domain. Therefore, the thesis does not directly compare its solution to existing ones like Operational Envelope, as they address different aspects and perspectives within the broader context of autonomous navigation systems.

Therefore, it may not be directly applicable to the scope of the thesis, which aims to provide a taxonomy for the ODD in the maritime domain. However, the concept of OE can be considered as a complementary aspect to the ODD for further research in autonomous vessel technology.

4. Problem Formulation

The primary objective of this thesis is to contribute to the creation of safer and more efficient maritime transportation by exploring the potential of autonomous navigation systems. The thesis recognizes that the development of autonomous navigation systems poses several challenges and obstacles, and it aims to address some of these issues. The ultimate aim is to help accelerate the progress of autonomous navigation technology and increase the adoption of these systems in the maritime industry.

In order to evaluate the capability of autonomous systems, it is important to analyze the data that the system has been exposed to. This helps to determine if the data covers a majority of the possible situations that the system may face. The diversity of the data directly affects the performance and the quality of the system. For example, if a system has been trained with data that only contains situations from a river environment, the same system might not work as expected in a lake environment. It is important to keep track of the quality of the data as more data comes in. However, this can become challenging as the amount of data increases. Therefore, there is a need for a scalable way to monitor the diversity of data, ensuring that the system is trained on a wide range of scenarios, and is capable of performing well in a variety of situations.

The research questions this thesis is trying to answer are listed below.

- How can the diversity of collected data for autonomous maritime systems be analyzed in a scalable way?
- Using existing solutions from the autonomous driving domain, how can the domains be designed for the autonomous maritime system to operate in?
- What challenges arise when applying domain describing techniques from the autonomous car domain in the autonomous shipping domain?

The first research question aims to address the issue of understanding the performance of autonomous maritime systems by evaluating the diversity and quality of the collected data. Analyzing the diversity of the data in a scalable way can provide insights into whether the system is capable of handling a wide range of situations and conditions, which can help in identifying the limitations of the system. Applying ODD concepts to actual collected sensor data can also give a better understanding of the system's performance and whether it is functioning as expected.

The second and third research questions are related to the design of the domains where the autonomous maritime system is expected to operate. By using existing solutions from the autonomous driving domain, a taxonomy can be designed to describe the domains where the system can operate. However, there may be challenges in applying the domain describing techniques from the autonomous driving domain in the autonomous shipping domain due to differences in the environment and operating conditions. Addressing these challenges can help in creating a more accurate and reliable taxonomy to describe the system's operating domains. Overall, these research questions aim to provide insights and solutions to improve the safety and efficiency of autonomous maritime transportation.

5. Method

In this section, multiple methods were utilized to achieve the research objectives. The first method involved conducting a literature study on the automated driving domain using various search engines such as Google Scholar, IEEEXplore, and ACM Digital Library. This enabled the identification of the existing literature on the subject matter.

The primary keywords utilized during the research encompassed terms such as "Autonomous Navigation", "Autonomous Cars", and "Autonomous Vehicles". During the literature review, the concept of ODD was discovered as a potential solution to the research problem. To gain a deeper understanding of ODD design concepts, a more detailed literature study was performed, which led to the identification of various ODD design concepts from different organizations.

The main ODD concepts used and practiced in the industry included BSI PAS 1883, WISE, AVSC, and ASAM OpenODD, each approaching the problem from different perspectives and proposing their own approach and representation.

In order to select the most suitable base ODD framework, a thorough comparison was conducted based on the specific requirements of this thesis. The requirements for this thesis were established through collaboration with industry professionals from Groke Technologies. Extensive discussions and meetings were held with experts in the field of autonomous maritime technologies to gain a deep understanding of the challenges and needs in the industry. Their valuable insights and feedback were instrumental in defining the scope and objectives of the thesis.

However, designing the taxonomy was not a one-time activity but rather an iterative process that required constant review and adjustment. Feedback from the company was crucial in this regard as it helped to identify areas where improvements were needed. The iterative approach also allowed for the taxonomy to be refined and expanded as more data became available, and new insights were gained from the research process. The taxonomy was designed to be flexible and adaptable, so it could be updated as the technology and the maritime environment evolve over time. The iterative approach to taxonomy design ensured that the taxonomy was fit for purpose and aligned with the specific requirements of the company. This, in turn, helped to ensure that the ODDs generated from the taxonomy were accurate, reliable, and effective in guiding the operation of autonomous vessels in the maritime environment.

6. Ethical and Societal Considerations

6.1. Ethical Considerations

In this section, the research ethics issues that may arise in the nature of this thesis will be discussed, along with the appropriate ways to deal with these issues. Firstly, the ethical issue related to data management in the thesis context is the ownership, privacy, and security of the data collected from real-life scenarios, which belongs to a private company.

To address the data management ethical concern, the thesis was conducted within the company's network, ensuring that the data remained within the company's premises. By keeping the data in-house, the thesis project aimed to protect the privacy and ownership rights of the company and the individuals associated with the collected data. Transporting the data outside of the company network can raise concerns regarding data security, potential unauthorized access, and compliance with data protection regulations. By keeping the data within the company's network, the thesis project mitigated these risks and maintained control over the data. Additionally, all data processing and analysis were performed within the company's network, ensuring that sensitive information was handled and processed in a secure environment. This approach helps maintain the confidentiality of the data and prevents any unauthorized use or disclosure. Furthermore, the thesis project respected the company's rights and obtained consent for using the final results derived from the data in the thesis.

Secondly, addressing bias in data and approach is an important ethical consideration in research. In the context of the thesis, there are several steps taken to address this issue:

Alternative Solutions: The existence of other solutions is mentioned in the thesis, indicating a recognition of the potential for bias and a commitment to exploring different perspectives. By acknowledging alternative approaches, the thesis promotes a more comprehensive understanding of the research problem and encourages critical thinking.

Decision-Making Process: The thesis explains the decision-making process behind selecting the ODD from BSI as the basis for the research. By providing insights into the rationale and considerations involved in this decision, the thesis promotes transparency and accountability. This allows readers to evaluate the approach and assess any potential biases.

Iterative Nature of the Solution: The thesis clarifies that the proposed solution is not intended to be a production-ready solution, but rather an initial draft that can be improved over time. By acknowledging the need for further development and improvement, the thesis demonstrates a commitment to ongoing refinement and addresses potential biases that may arise from an incomplete or early-stage solution.

6.2. Societal Considerations

The societal considerations of this thesis are significant, particularly in relation to the safety of autonomous maritime technologies. Given the world's dependence on maritime transportation, ensuring the safety and reliability of these technologies is of utmost importance. Any accidents or disruptions during the logistics process in the maritime environment can have severe consequences, including shortages of essential resources and significant economic impacts. By focusing on the safety aspects of autonomous maritime technologies, this thesis addresses a critical societal need and contributes to the overall well-being and efficiency of global maritime operations.

7. Description of the work

This section is focused on the detailed description of the work that has been done in this thesis. The section begins by discussing the project requirements, outlining the key objectives and goals that need to be achieved. Following this, the selection process for choosing the appropriate ODD framework as the foundation for the research is explained, taking into account various factors and considerations.

Next, the modifications made to the chosen ODD framework are thoroughly explained, detailing the attributes that were altered or introduced to tailor the framework specifically for the autonomous shipping industry. Each design decision is justified with comprehensive reasoning.

The newly created taxonomy for the autonomous shipping industry is then introduced, presenting a holistic view of the ODD attributes and their significance in the maritime context. The thought process behind each attribute design is extensively discussed, ensuring a clear understanding of the taxonomy's purpose and applicability.

This section proceeds to demonstrate the mapping of the Navigable Area Type attribute to actual sensor data, highlighting the challenges encountered in this process and the strategies employed to address these challenges effectively.

Furthermore, the conversion of ODDs into Elastic Search database queries is explained in detail, offering insights into how queries are created from the ODD attributes. This step allows for efficient data extraction and visualization, making the process more accessible and practical for future use.

Lastly, the thesis delves into a real-world use case with Groke Technologies, showcasing the application of the ODD taxonomy on actual data collected by the Groke Pro system. The results of this application are presented, and the effectiveness of the ODD framework in analyzing and classifying the data is discussed.

7.1. Requirements

In this section, three main requirements were identified by the industry experts. The requirements are demonstrated below.

- Req1: The solution should be scalable and easy to maintain. This means that the solution should be able to handle a large amount of data and be flexible enough to adapt to any changes or updates that may be needed.
- Req2: The solution should be easy to integrate with the user's own database and query engine, which was Elastic Search on this occasion. This would allow the company to easily access and analyze the data collected by their navigational awareness system.
- Req3: Attributes of the designed solution should represent measurable units from their sensors. This means that the data collected by the system should accurately represent the capabilities of the system, and be measurable and verifiable.

7.2. Selection of the base ODD framework

In the field of autonomous maritime vehicles, there is a lack of established taxonomy compared to autonomous cars. As there are limited numbers of production-level autonomous ships, there is a need for a taxonomy that can define the key attributes. In this thesis, a new taxonomy has been introduced specifically for autonomous maritime vehicles, taking into consideration the unique characteristics of the maritime environment. The taxonomy has been designed as an initial version that can be extended and improved upon as the field of autonomous maritime vehicles continues to evolve.

In order to create an ODD taxonomy for an autonomous maritime environment, the existing taxonomies in the autonomous car domain have been used as an inspiration. The main reason for this was the lack of an established taxonomy for autonomous ships. Thus, the research began by exploring the existing taxonomies in the autonomous car domain such as BSI PAS 1883, WISE, and AVSC.

To determine which ODD concept to adopt, industry experts from Groke Technologies were interviewed to gain insight into the necessary specifications and requirements for the final ODD design. A comprehensive analysis of these standards was conducted, and they were compared based on the specified requirements. The primary objective was to ensure that the data was easily accessible and that informative queries could be performed on it to comprehend the extent to which their data covered the system's capabilities and to identify areas where additional data was required. This evidence can be in the form of an ODD, which describes the conditions under which the vessel can operate. By using the ODD taxonomy designed in this thesis, the company can create their own ODDs that can be used to certify their systems. This provides a measurable unit for the authorities to evaluate the system and to determine whether it meets the required safety and efficiency standards. Therefore, the designed ODDs can potentially open up further certification possibilities for the companies and provide a standard for the authorities to evaluate autonomous vessels. The data was also required to accurately represent the system's capabilities. Another critical feature was the ability to describe the vessel's location, whether it was in a zone such as a harbor or open sea.

Among these, WISE primarily focused on obstacles on the road such as animals, pedestrians, and other vehicles, which are not as relevant to the maritime environment. On the other hand, AVSC provided a conceptual framework and a lexicon that would be essential for the further work of this thesis. After evaluating and comparing these taxonomies, it was decided that BSI PAS 1883 would be the most suitable taxonomy for the maritime environment, as it provided a comprehensive and practical set of attributes that could be used as a starting point for creating a taxonomy for autonomous ships. However, the transition from the autonomous car domain to the maritime environment required significant modifications to the taxonomy, as there are fundamental differences between the two environments. An additional resource that was considered during the research for this thesis was the concept paper from ASAM called ASAM OpenODD. The main focus of OpenODD is to provide a programmable ODD language for simulation systems, which is particularly useful as the ODDs can be utilized by different programs in the autonomous system. The concept of OpenODD was explored to determine its potential relevance and usefulness for the creation of an ODD taxonomy in the autonomous maritime domain. The combination of the taxonomy provided by BSI PAS 1883 and the machine-readable format from Open ODD ensures that the domain can be represented in a clear, concise, and standardized way that can be used in various applications. This approach also allows for the easy integration of the ODD into different systems and the sharing of the ODD between different stakeholders.

7.3. Modification process of the base ODD framework

To generate an ODD, it is necessary to have a taxonomy that includes all possible attributes relevant to the ODD. During the design of the ODD taxonomy for autonomous maritime vehicles, the attributes that could be measured by the commonly available sensor units were included. This was because the inclusion of other attributes that cannot be measured would not be useful in generating an accurate ODD for the autonomous vessel. By including only the measurable attributes in the ODD taxonomy, the resulting framework could be more effectively applied in real-world scenarios. This approach also helped to ensure that the ODD accurately represented the operational limits of the autonomous vessel and provided a realistic framework for safe and effective autonomous navigation.

In the development of the ODD taxonomy for autonomous maritime vehicles, the *Scenery* attribute was kept with some modifications made to the sub-attributes as seen in Figure 3. *Zones* attribute was retained but the sub-attributes were converted into a field that represents the geographical location of the vehicle. This was necessary because the maritime environment does not have predefined zones like roads or highways, which are present in the autonomous car domain. Instead, the location of the vessel and its proximity to coastlines and other navigational hazards are more relevant factors to consider.

The Drivable Area attribute in the original taxonomy was converted into Navigable Area, as navigation is a more suitable term for the maritime environment. The Navigable Area Type subattribute was added to represent frequently occurring environments such as Harbour, Channel, River, Open Sea, and Coastal Area. The Navigable Area Signs subattribute was added to represent traffic signs in the maritime environment, such as Buoys and Lighthouses. The Navigable Area Surface subattribute was added to represent the different surface conditions of the water, such as

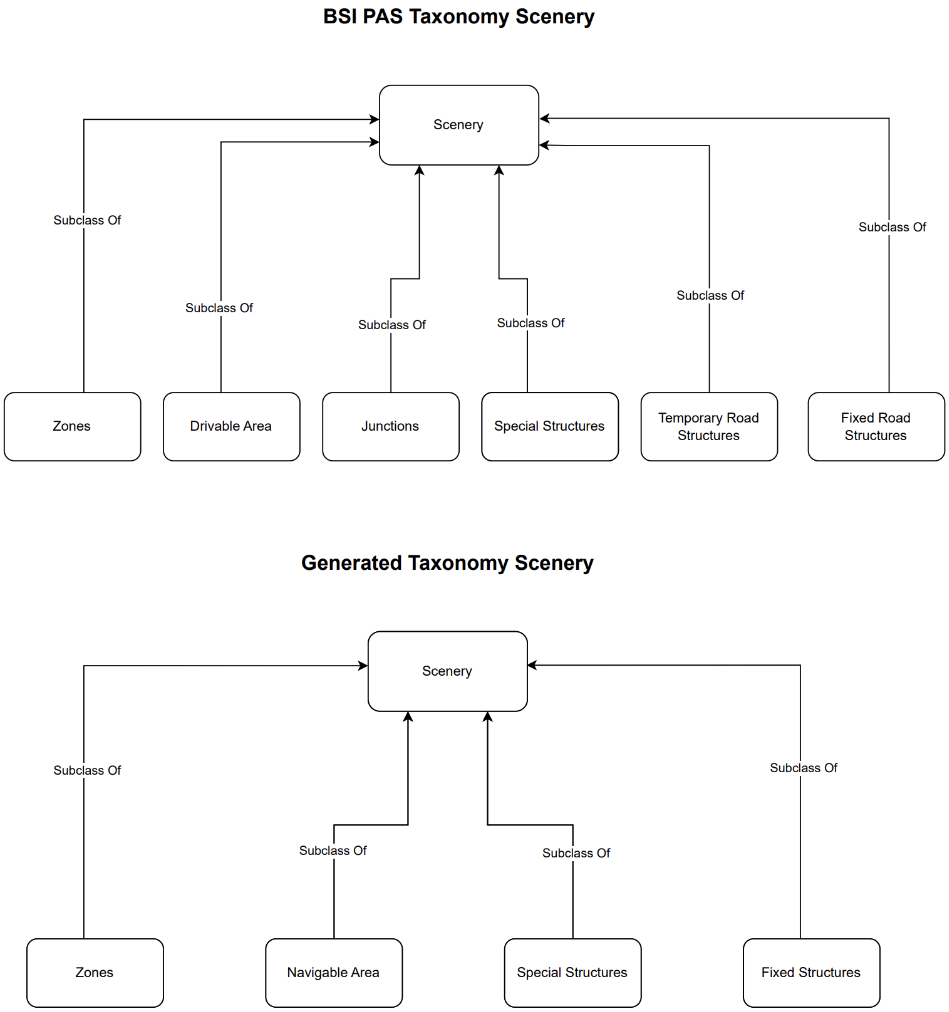

Figure 3: Taxonomy conversion from BSI PAS for Scenery attribute

Ice during cold conditions.

In the process of adapting the taxonomy from the autonomous car domain to the autonomous maritime domain, some subattributes such as Junctions and Temporary Road Structures were removed as they were not applicable in the maritime environment. The Fixed Road Structures subattribute was modified to represent static objects in the maritime environment that should be avoided, and the Special Structures subattribute was added to represent objects specific to the sea, such as offshore oil platforms. Overall, the modifications made to the Scenery attribute of the ODD taxonomy were necessary to ensure that it is relevant and suitable for the unique challenges of the maritime environment. All the modified subattributes can be seen from Table 1.

BSI PAS 1883	**Generated Taxonomy**
Zones	Zones
Drivable Area	Navigable Area
Junctions	-
Special Structures	Special Structures
Temporary Road Structures	-
Fixed Road Structures	Fixed Structures
Weather	Weather
Particulates	Particulates
-	Temperature
-	Waves
Illumination	Illumination
Connectivity	-
Traffic	Traffic
-	Flow Rate
-	Agent Type
Subject Vehicle	Subject Vessel

Table 1: Altered subattributes between BSI PAS 1883 and Generated Taxonomy

Environmental conditions are an essential component of the ODD taxonomy, and in this thesis, only the attributes that could be measured by the sensor unit were included. The categorical representation of the values has been removed, as it is easier to maintain the attributes with the interval of the measurement units. The subattributes for particulates have been changed to sea fog, water droplets, and humidity, as these are the common particulates in the maritime environment. In addition, temperature and sea temperature have been added as subattributes since the temperature can affect the state of the sensors during extreme cold or hot situations.

Furthermore, waves have been added as a separate attribute because the accuracy of the sensor units is affected as the vehicle begins to roll in different conditions. Object detection models may struggle to detect other objects in the sea when the camera cannot see the object in its usual state. Finally, for *Illumination*, ship light and aid light have been added as part of the artificial illumination attributes because of the domain's specificities. These changes have been made to align the ODD taxonomy with the maritime environment and ensure that the taxonomy contains all the relevant environmental conditions that affect autonomous maritime vehicles.

Regarding *Dynamic Elements*, motion and navigation have been added to the subject ship attribute as ships can move in different directions and can be in different states, unlike cars.

7.4. Taxonomy

To describe an ODD, a structured taxonomy is needed to represent the different aspects of the environment in which the autonomous system will operate. Taxonomy helps to create a standardized approach to describe the ODD and provides a clear understanding of the system's capabilities and limitations. By using a taxonomy, it becomes easier to identify which attributes are relevant to a particular ODD and which are not.

The final designed taxonomy can be seen in Figure 4, 5, and 6. The taxonomy comprises of three main attributes - Scenery, Environmental Conditions, and Dynamic Elements. The Scenery attribute describes the location of the ship, whether it is in the open sea or in the harbor and other static objects in the maritime environment. The Environmental Conditions attribute explains the weather, sea conditions, and visibility in the area. The Dynamic Elements attribute deals with the other objects in the environment such as other vessels, buoys, and floating objects. It is noteworthy that these attributes are not mandatory fields, and any combination of these attributes can be used to generate ODD documents.

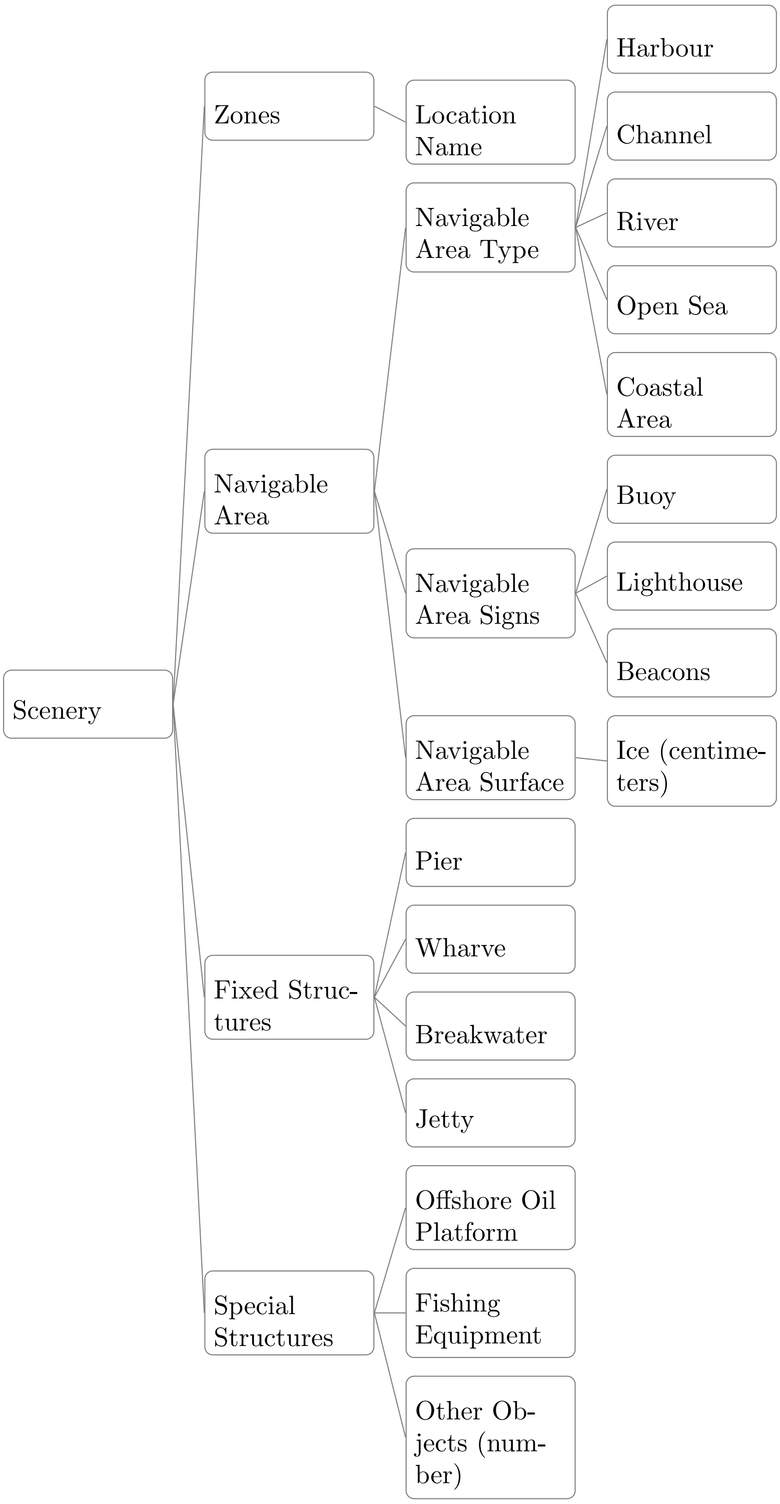

Figure 4: Designed ODD taxonomy for autonomous maritime domain. Scenery attribute.

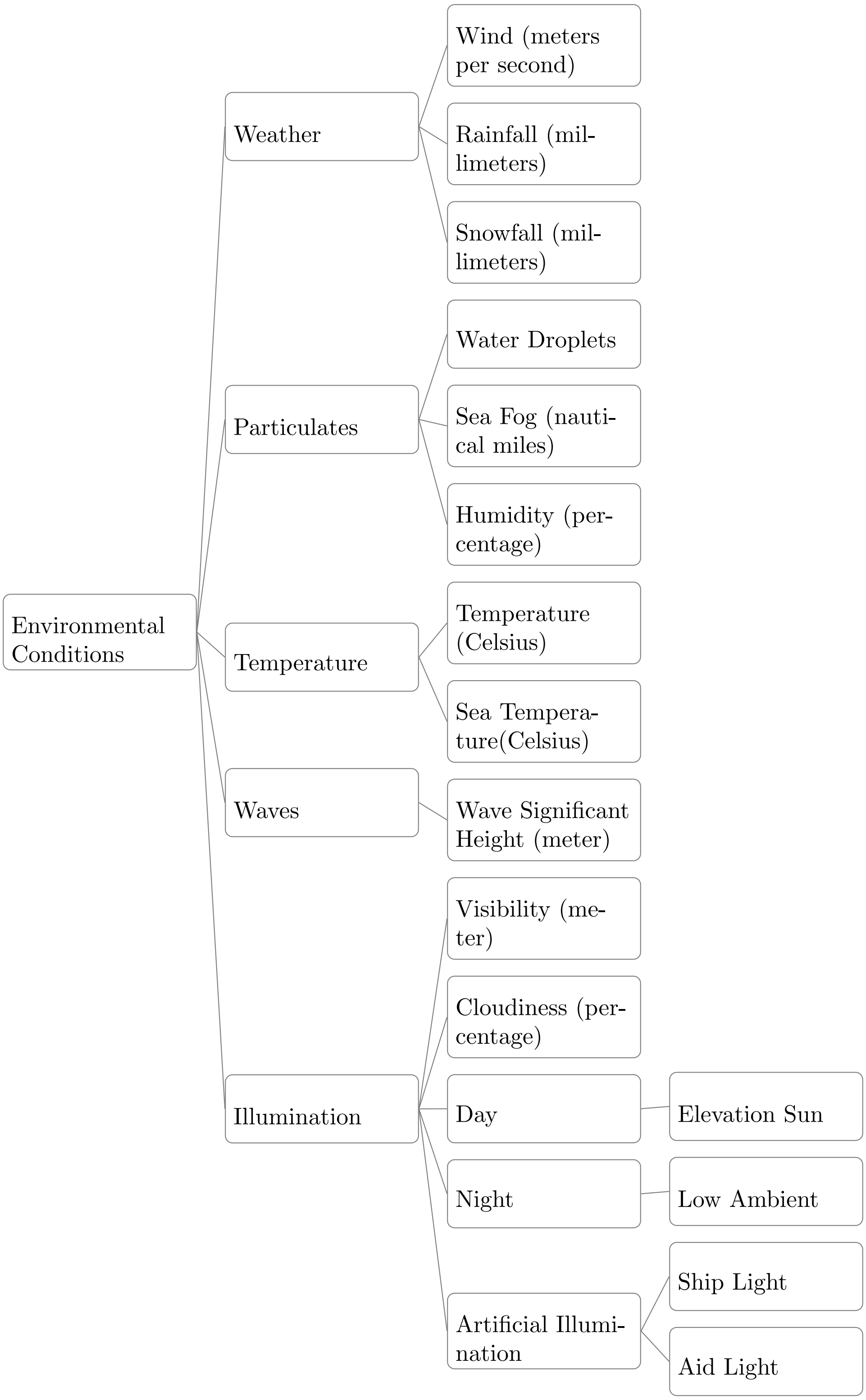

Figure 5: Designed ODD taxonomy for autonomous maritime domain. Environmental Conditions attribute.

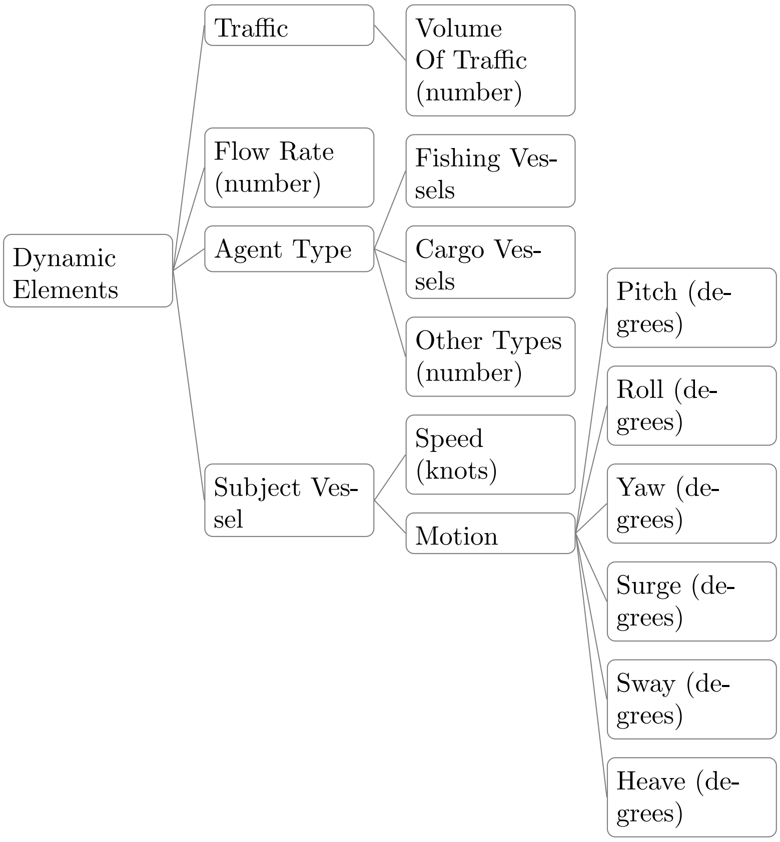

Figure 6: Designed ODD taxonomy for autonomous maritime domain. Dynamic Elements attribute.

7.5. Attributes

7.5.1 Scenery

Scenery plays a crucial role in ODD design as it describes the static objects present in the domain. Detailed visualization of *Scenery* can be seen from Figure 4. *Zones* attribute is one of the key attributes used in describing the operational design domain of autonomous maritime systems. This attribute provides information about the geographic location of the vessel, including the country or region where it is operating. This information is crucial in defining the ODD as it helps to determine the specific environmental and regulatory conditions that the vessel is likely to encounter. By understanding the zone attribute of a vessel's ODD, it becomes possible to create more accurate and effective navigation plans, as well as identify potential risks or limitations associated with operating in specific regions or countries. Overall, the zone attribute plays a critical role in ensuring that autonomous maritime systems are designed and operated in a safe, effective, and compliant manner.

Navigable Area attribute is an important element of the ODD for autonomous maritime systems. It defines the water area where the vessel is operating, and it can affect the vessel's state, behavior, and performance. For example, a vessel operating in a canal with narrow dimensions and sharp turns might require different navigation strategies than a vessel operating in a river with wide dimensions and strong currents. Therefore, it is essential to specify the *Navigable Area* attribute in the ODD to ensure that the autonomous system is designed and tested to operate effectively in the intended environment.

Navigable Area attribute consists of three subattributes, *Navigable Area Type*, *Navigable Area Signs* and *Navigable Area Surface*. The subattribute *Navigable Area Type* allows us to categorize the vessel's surroundings into distinct zones such as Harbour, Channel, River, Open Sea, and Coastal Area. Understanding the type of environment is crucial for effective navigation and decision-making, as each zone comes with its unique challenges and regulations.

Navigable Area Signs plays a pivotal role in describing the traffic signs present in the operating environment. The presence of Buoys, Lighthouses, and Beacons can indicate navigational hazards, safe routes, and points of interest. This information is essential for understanding the density of the traffic in different areas, enabling the vessel's crew to make informed decisions on route planning and collision avoidance.

Moreover, the *Navigable Area Surface* subattribute focuses on describing the possible water surface conditions, particularly in winter when ice can be present. Measuring the surface conditions, especially the presence of ice, is critical for ensuring safe navigation and avoiding potential dangers posed by ice formation. This information can be obtained using public APIs about the weather, allowing the vessel to adapt its course and speed based on the current surface conditions.

Fixed Structures represent the structures that can appear in the maritime environment. These include pier, wharve, breakwater, and jetty. The ability to detect and recognize fixed structures is crucial for the safe and efficient operation of autonomous vessels. Therefore, it is important to include the Fixed Structures attribute in the taxonomy for describing the vessel's environment.

Special Structures attribute is of significant importance as it provides valuable information about the presence of structures in the maritime environment that are not considered traditional traffic signs. These structures, such as offshore oil platforms and fishing equipment, can have a significant impact on navigation and safety. Understanding the location and characteristics of these special structures is crucial for vessels to navigate safely and avoid potential hazards.

Additionally, this attribute includes the subattribute as *Other Objects*, which accounts for uncommon objects that may appear in the environment. These objects might not fit into predefined attributes but can still impact navigation. *Number of Objects* provides information on how many of these uncommon objects can be expected in a particular environment, offering a measure of their potential prevalence.

7.5.2 Environmental Conditions

Environmental Conditions attribute is an important aspect of the ODD taxonomy, which describes the environmental conditions in which the autonomous maritime system operates. Detailed visualization of *Environmental Conditions* can be seen from Figure 5. One of the subattributes under the environmental conditions attribute is *Weather*, which refers to the possible weather situations in the maritime environment. These weather situations can include wind, rainfall, or snowfall, which can significantly impact the performance of the autonomous system. To represent these weather situations, the subattribute weather contains interval values that can represent the measured unit as a value. The unit of measurement is defined next to the property names in the parenthesis. The intervals for these units are not defined, thus if the taxonomy is applied, users need to specify their own intervals for these attributes based on the domain knowledge. By including this subattribute in the taxonomy, the autonomous maritime system can be designed to operate optimally under different weather conditions.

Particulates subattribute mainly describes possible particulates that can decrease the performance of sensors in the maritime environment such as water droplets, sea fog, or humidity.

Temperature subattribute is divided into two subattributes: temperature and sea temperature. This division is necessary because the temperature difference between water and air can sometimes affect the performance of sensors. Temperatures can be expressed both in Celsius and Fahrenheit.

The *Waves* subattribute plays a critical role in ensuring the safety and accuracy of the vessel's operations in the maritime environment. The *Wave Significant Height* refers to the measurement of the height of significant waves in the surrounding waters. Waves can have a significant impact on a vessel's stability, maneuverability, and sensor performance, making it essential to monitor and consider this attribute.

When the *Wave Significant Height* is high, it can pose serious challenges and dangers to vessels. Large waves can cause a vessel to pitch, roll, and heave, which may lead to loss of control and increase the risk of capsizing or grounding. High waves can also affect the performance of onboard sensors, such as cameras and LiDAR, impacting the quality and reliability of the data collected.

Illumination subattribute is a comprehensive attribute that encompasses various aspects of lighting conditions in the maritime environment. It includes both natural and artificial illuminations that can occur during the vessel's operation time. One of the key components of this subattribute is *Visibility*, which describes the level of visibility in the surrounding environment. The visibility level is crucial as it directly affects the performance of cameras and human vision. It can be predicted based on weather information, enabling vessels to prepare for varying visibility conditions.

Additionally, *Cloudiness* is considered as a separate attribute within *Illumination*. While cloudiness can impact visibility, it also affects the overall weather situation. As weather conditions can significantly influence vessel operations, it is important to account for cloudiness as a distinct factor.

The distinction between *Day* and *Night* attributes is also essential. Daytime and nighttime conditions can have different effects on the performance of sensors and the crew's ability to navigate safely. During the daytime, the *Elevation Sun* attribute becomes relevant as the position of the sun can impact visibility and the direction of light sources.

Between daytime and pitch dark, there is a transitional situation where there is still some visibility, which is addressed using the *Low Ambient* attribute. This attribute acknowledges the conditions where there is a limited level of illumination, which may be relevant during dusk or dawn.

The subattribute *Artificial Illumination* focuses on lighting sources that are not naturally present in the environment. It includes lights from other ships, referred to as *Ship Light*, and lights from traffic signs, known as *Aid Light*. Considering artificial illumination is crucial as it helps vessels be aware of additional sources of light that may impact their visibility and safety.

7.5.3 Dynamic Elements

Dynamic Elements attribute plays a crucial role in describing the traffic and the state of the ship in the maritime environment. It contains several subattributes that provide valuable information for vessel operators to make informed decisions and ensure safe navigation. Detailed visualization of *Dynamic Elements* can be seen from Figure 6.

Traffic subattribute consists of *Volume Of Traffic*, which indicates the intensity of vessel presence around the subject vessel. This information is vital as it helps determine the likelihood of collision with other vessels. A high volume of traffic may require the vessel to exercise caution and adjust its course accordingly. On the other hand, the *Flow Rate* subattribute focuses on the average number of vessels present on an hourly basis, providing insight into the typical traffic patterns in a given area.

The *Agent Type* subattribute is responsible for describing the types of vessels that can be expected in different environments, such as fishing vessels or cargo vessels. This information is essential, as different vessel types have varying characteristics in terms of speed and maneuverability. Knowing the types of vessels present allows the crew to anticipate the behavior of other vessels and take appropriate measures to avoid potential hazards. The *Other Types* subattribute acts as a placeholder and can be customized to specific use cases, allowing for a more flexible and adaptable taxonomy.

Furthermore, the *Subject Ship* subattribute describes the state of the main vessel where the crew and the system are located. *Speed* is a crucial factor that directly impacts safety and the performance of sensors. A vessel's speed determines how quickly it can respond to changing conditions and avoid potential collisions.

The *Motion* subattribute provides detailed information about the ship's movements, including *Pitch*, *Roll*, *Yaw*, *Surge*, *Sway*, and *Heave*. These motion parameters are crucial as they affect the vessel's stability, sensor accuracy, and crew comfort. Understanding the vessel's motions enables the crew to make adjustments to ensure a stable and safe navigation experience.

7.6. Navigable Area Type mapping

The task of mapping Navigable Area (NA) information from taxonomy to the presented data poses a significant challenge due to the inherent complexity of the data. According to the original NA taxonomy, a ship's operating area is defined, which is difficult to discern in actual sensor data. The available data only provides the ship's location in latitude and longitude format, which is insufficient to determine the type of potential areas, such as Open Sea, Canal, River, or others. To address this issue, a modification to the Navigable Area Type is proposed, which includes two zones: ”Free Zone” and ”Cautious Zone”. These zones are calculated based on the location of the ship and the nearby coastlines, thus providing a more accurate representation of the navigable areas. The algorithm 1 describes the pseudocode on how the zone classification algorithm was implemented in this thesis.

Algorithm 1 Determining Navigable Area Type (Free Zone or Cautious Zone)

```
    Input:(lat,lon)
$coastLinePoints$ = Load Coastal Points from Dataset With 10m deviation
$ballTreePoints$ = Convert $coastLinePoints$ into BallTree data structure for faster search
$closestPoints$ = Get $k$ closest points in the $ballTreePoints$ based on $(lat, lon)$ where $k = 10$
$polygonPoints$ = Create Polygon from $closestPoints$
if $(lat, lon)$ is inside $polygonPoints$ then
    return "CautiousZone"
end if
return "FreeZone"
```

If there is a coastline within 10km distance from the ship then this is considered as the ship is in a dangerous zone. The example ”Cautious Zone” visualization can be seen in Figure 7. In Figure 7, the location of the ship is denoted by a blue dot, and a polygon is formed by connecting the 10 closest coastal points to the ship. If the ship is located within this polygon, it can be inferred that the ship is partially in the cautious zone, as it is surrounded by land on both sides.

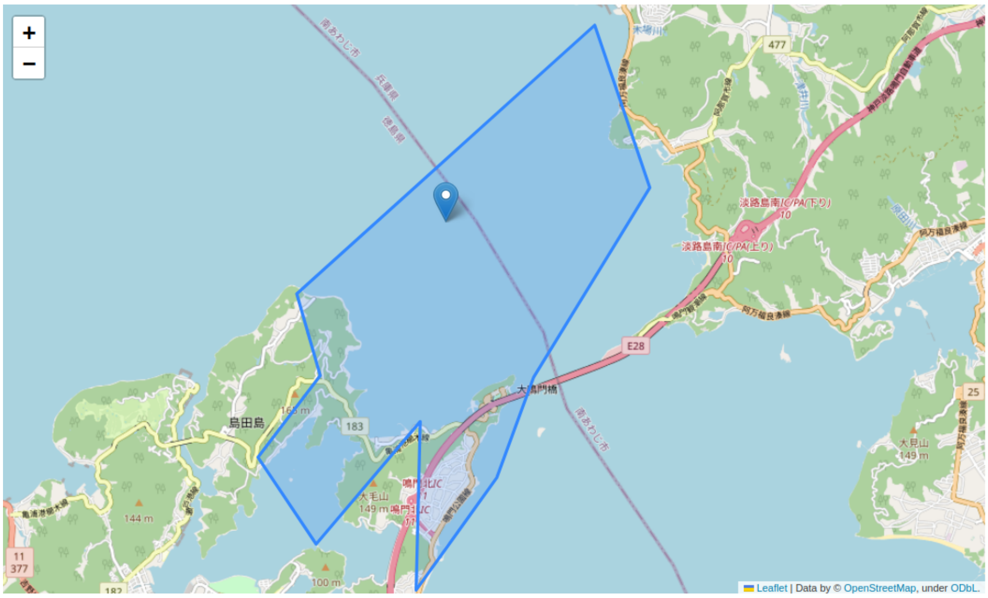

Figure 7: The ship is surrounded by land within 10km, Cautious Zone

Otherwise, the navigable area type is labeled as ”Free Zone”. The demonstration of ”Free Zone” can be seen in Figure 8. In Figure 8, the coastlines are located more than 10km away from the subject ship, and the ship is not situated within the constructed polygon. This algorithm is only used for experimental purposes as the accuracy of the algorithm has not been measured. Coastline information has been retrieved from [33]. The coastline information is stored in GeoJSON format which includes points within 10m deviation of the coasts. The coordinates of the coasts have been extracted from GeoJSON objects and they have been stored in a flat structure. These coordinates represent the coastlines of the whole world, thus in order to search efficiently in the coordinates, the k nearest neighbors algorithm was used to store the coordinates where k=10 in an easily accessible way by using BallTree [34] from sklearn [35]. With this approach, the search for the nearest coastlines was optimized.

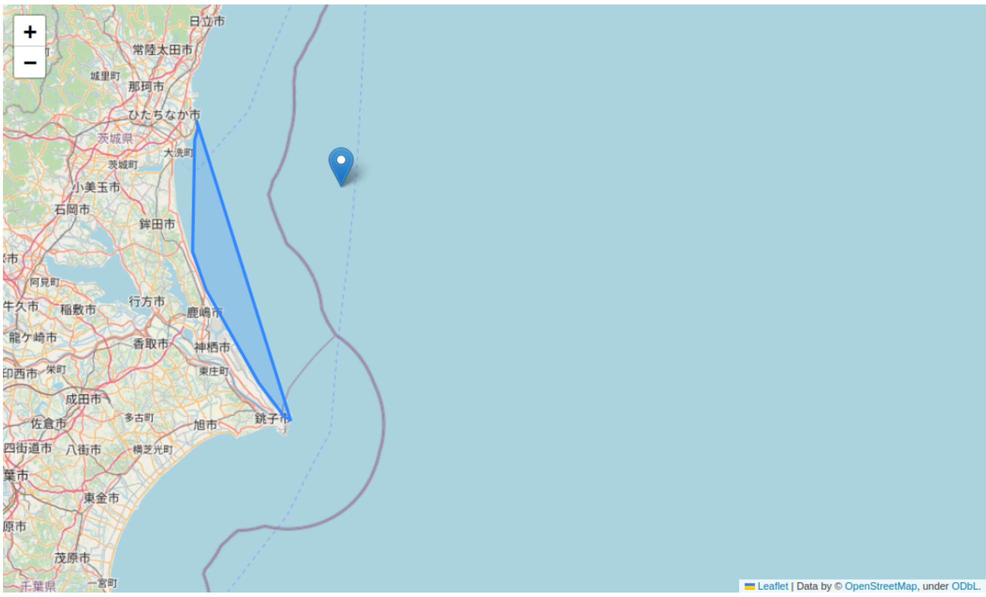

Figure 8: The ship is not surrounded by land within 10km, Free Zone

7.7. Extraction of data into Elastic Search

This thesis relies on data about the ship and its surroundings, which was provided by the company. Due to company regulations, the content and structure of this data can not be published fully. Initially, the data was in rosbag format, which contains sensor information from various sensor units. These rosbag files are recordings of a few minutes of sensor readings. The data was preprocessed already with the extracted information available in JSON format. However, the sheer volume of data produced by sensors posed a challenge in terms of data management and analysis. Thus, specific fields were extracted from the original data that were relevant for understanding the state of the ship and for creating an ODD mapping. Additionally, as the sensor files were recordings of sensor readings, the extraction of the maximum and minimum recordings of the relevant fields from the dataset was performed. Once data is extracted, it is vital that the data can be queried easily both individually and using Operational Design Documents. Thus each detection of the ships is stored as separate documents where certain queries like Listing 3 can be performed. In the Listing 3 all the detected objects from the ego ship where it is in ”Free Zone” with a distance range from 3000 to 400000 nautical miles can be collected. This enables for querying ODDs as an Elastic Search query can be constructed with the definition of ODDs.

```
”query”: {
    ”bool”: {
        ”filter”: [
            {
                ”bool”: {
                    ”must”: [
                        {
                            ”match”: {
                                ”area”: ”Free Zone”
                            }
                        },
                        {
                            ”range”: {
                                ”distance_from_sensor_unit”: {
                                    ”gte”: ”3000”,
                                    ”lt”: ”400000”

                                }
                            }
                        }
                    ]
                }
```

```
            }
        ]
    }
}
```

Listing 3: Example Elastic Search Query for Extracted data

7.8. Practical Application of Generated ODD: A Use Case with Groke Technologies

In this section, the practical application of the ODD taxonomy to the dataset collected by Groke Pro at Groke Technologies was demonstrated. The goal was to analyze a specific portion of the dataset and utilize generated ODD taxonomy to classify the data and extract valuable information. By achieving this, the company can establish automated pipelines that regularly run these ODDs on their updated dataset. This process allows them to assess the diversity of their collected data and identify areas where data is insufficient, prompting them to collect more data in those specific domains. Such an approach aids the company in enhancing the quality of its data and ensuring a comprehensive understanding of its system's operational domains.

Classifying all the data points in the company's vast database was not feasible for this project, so only a limited amount of data was experimented with. The data was filtered to include only ships with diverse and interesting situations, and these were collected into a single Elastic Search index for further visualization.

In total, 1063 entries were created in the Elastic Search index, representing each detection made by the system in a real-life scenario. However, 576 of these entries were found to have missing speed values of the vessel, making them unsuitable for further analysis based on the ODD criteria.

```
Dynamic_Elements:
    Subject_Ship:
        Speed: [0, 0]
```

The excerpt describes the creation of an ODD that focuses on a specific operational scenario, namely a ship parked in a stationary position. The ODD was created using a YAML format, which is a human-readable data serialization language. The excerpt shows the specific ODD that was created for this scenario, which contains information about the dynamic elements in the scenario, such as the "Subject_Ship" and its "Speed" attribute. The "Speed" attribute is defined as a range of values, [0, 0], which indicates that the ship is stationary, not moving in any direction. By creating this ODD, the thesis demonstrates how the ODD concept can be used to define specific operational scenarios for autonomous ships and how it can be applied in real-world scenarios.

The results from the ODD created for the parked situations highlighted the importance of having a specific ODD for different scenarios. The images captured from the parked situations, as shown in Figure 9a and Figure 9b, look different from each other despite belonging to the same ODD with a speed interval of [0,0]. This indicates that the ODD used was either too generic or not well-defined enough to capture the nuances of different situations within the same domain.

To address this issue, a more detailed ODD could be created to capture specific parameters relevant to the situation. This could include factors such as lighting conditions, weather, proximity to other vessels, and the presence of nearby land areas. Companies can generate their own ODDs using the taxonomy created in this thesis, which would allow them to create more specific and detailed ODDs tailored to their specific needs.

```
Dynamic_Elements:
    Subject_Ship:
        Speed: [0, 0]
    Traffic:
        [0, 3]
Environmental_Conditions:
    Weather:
        Night: dark
```

The new ODD presented in the code snippet above includes more specific criteria that further define the operational design domain. The additional criteria include traffic density and weather conditions, specifically the time of day as indicated by the "Night: dark" attribute. These additional criteria can help to differentiate between different situations that may occur within the same speed range.

By defining the ODD more precisely, it becomes easier to analyze the data and draw meaningful insights from it. It also helps to ensure that the system is only operating within a safe and reliable range of conditions, as defined by the ODD. It is worth noting that the specific criteria included in an ODD will vary depending on the specific application and use case of the system. Therefore, it is important to carefully consider the relevant factors and ensure that the ODD is well-defined and tailored to the specific needs of the system.

The designed ODD framework offers flexibility to the users to add or remove attributes from the ODD to customize their ODDs as per their requirements. This allows for a better understanding of the surrounding environment and helps in collecting relevant data for the autonomous systems. For instance, if a specific attribute does not provide any useful information for the particular application, it can be removed from the ODD, while a new attribute can be added to capture the relevant information. This flexibility in the ODD framework allows for the creation of more effective and accurate ODDs, which in turn enhances the performance of autonomous systems in real-world scenarios. Thus, the designed ODD framework provides the users with a tool to customize the ODDs according to their needs and improve their understanding of the domain.

(a)

(b)

Figure 9: Parked situation data

The newly created ODD provides a more detailed and specific description of the domain in which the ship is expected to operate. The added attribute of "Environmental Conditions: Weather: Night: dark" specifies that the ship should be in a dark environment, which is a more realistic scenario for a parked ship during the night. Additionally, the "Traffic: [0, 3]" interval suggests that the ship should be in an area with a low to moderate amount of traffic, which is more relevant to a parked ship in a harbor or a bay.

As a result of this more specific ODD, the images returned by the query (Figure 10a and Figure 10b) appear to be more similar to each other, which suggests that the ODD is better defined and more suited to the task at hand. This demonstrates the importance of creating tailored ODDs for specific scenarios and the flexibility of the ODD framework to incorporate various attributes from the taxonomy to create a more accurate representation of the domain.

(a)

(b)

Figure 10: Parked situation new query data

8. Results

The result section of this thesis reports on the data visualization process and data classification performed in the project. One of the major challenges encountered during the visualization process was the presence of missing data due to malfunctioning sensors or network connection issues. This is a common occurrence in maritime systems, especially when ships travel to remote areas or face severe weather conditions. As a result, some of the data had to be ignored in the analysis.

In order to showcase the effectiveness of ODDs in classifying data points, a total of 1063 data points were analyzed in this study. The analysis focused on two specific attributes: Speed and Roll. The attribute "Speed" refers to the speed of the ship, which indicates the movement of the vessel in the maritime environment. The sensor units installed on the ship collect data related to the ship's speed during its operation.

The attribute "Roll" represents the standard deviation of the Roll angle of the ship. The Roll angle refers to the side-to-side motion of the ship, indicating its stability in the water. The standard deviation is calculated over a 2-minute interval, capturing the variation in the Roll angle during that time period. By examining the Speed and Roll attributes of the data points, the study aimed to identify any potential anomalies or outliers that may have been overlooked initially. The ODDs provided a framework to define the expected ranges or patterns for these attributes, allowing for the detection of data points that deviated from the norm.

These ranges were determined based on the recommendations and insights provided by industry experts, who possess extensive knowledge and experience in the maritime domain. The distribution of these ranges and how many data points are in each attribute can be seen from Table 2.

To categorize the Speed attribute and provide a meaningful representation of the ship's movement, four distinct ranges were established. The distribution of data points is effectively visualized in Table 2. This table provides a comprehensive overview of the number of data points within each category, allowing for easy analysis and comparison.

Speed				Roll		
Standing	Slow	Medium	Fast	No roll	Minor	Major
104	12	203	168	275	531	257

Table 2: Data distribution across Speed and Roll Domains

Firstly, the first range, known as "Standing," encompasses situations where the ship's speed is recorded as 0 knots. This indicates that the ship is stationary or in a static position, such as when it is anchored or docked. Within this range, a total of 104 data points were observed.

The second range, referred to as "Slow," includes speeds ranging from 0 to 5 knots. This range captures scenarios where the ship is moving at a relatively slow pace, such as during maneuvering, navigating through confined areas, or performing maintenance activities. A total of 12 data points have been observed within this range.

The third range, labeled as "Medium," covers speeds between 5 to 11 knots. This range signifies moderate speed, typically observed during normal cruising or transit operations. Within the Speed section, the "Medium" category exhibits the highest number of data points, totaling 203 observations.

Lastly, the "Fast" range represents speeds exceeding 11 knots. This range corresponds to situations where the ship is moving at a relatively high velocity, such as during open-water navigation, long-distance voyages, or high-speed transits. In addition, the "Fast" range within the Speed section encompasses a total of 168 data points, indicating a significant number of observations within this category.

Speed	Roll		
	No roll	Minor	Major
Standing	66	35	3
Slow	0	12	0
Medium	0	135	68
Fast	2	105	61

Table 3: Data distribution across Speed and Roll Domain Combination

The Roll attribute, which represents the standard deviation of the ship's Roll angle over a 2-minute interval, is categorized into three distinct ranges: "No Roll," "Minor," and "Major."

The first range, "No Roll," indicates situations where the ship exhibits minimal or no rolling motion. This range corresponds to scenarios where the standard deviation of the Roll angle is negligible or close to zero, suggesting a relatively stable and balanced position of the ship. The "No Roll" range encompasses a significant number of data points, totaling 275 observations.

The second range, "Minor," encompasses situations where the ship experiences minor rolling motion. This range is characterized by Roll standard deviations between 0.05 and 0.4. The minor rolling motion signifies slight variations in the ship's orientation, which can be attributed to factors such as waves, currents, or minor disturbances in the maritime environment. Among the different ranges, the "Minor" range stands out with the highest number of data points, amounting to a total of 531 observations. This indicates that the vessel experiences minor roll motion more frequently compared to the other ranges.

The third range, "Major" represents instances where the ship encounters significant rolling motion. This range is defined by Roll standard deviations exceeding 0.4, indicating pronounced fluctuations in the ship's Roll angle. Major rolling motion is typically associated with adverse weather conditions, rough seas, or other external factors that cause substantial movement of the ship. Surprisingly, the "Major" category exhibits a comparable number of data points to the "No Roll" range, with a total of 275 observations. This suggests that the vessel encounters significant roll motion events to a similar extent as situations where no roll occurs. The presence of a significant number of data points in the "Major" range highlights the importance of addressing and understanding these high-roll scenarios.

The ODD framework plays a crucial role in helping users understand their data better by providing a structured approach to analyzing and classifying different domains of operation. In Table 3, the data points are categorized based on their Speed and Roll combinations, allowing for a clear visualization of the available data. The table showcases the number of data points within each Speed and Roll combination. For instance, in the "No roll" category, there are 66 data points for the Standing speed, while the Slow, Medium, and Fast speeds do not have any observations without a roll. Similarly, the "Minor" category has 135 data points for the Medium speed and 105 data points for the Fast speed. Additionally, the "Major" category comprises 68 data points for the Medium speed and 61 data points for the Fast speed. Utilizing ODDs enables us to obtain definitions for attributes like Roll, Slow, Medium, etc., making it effortless to modify these definitions as needed.

One of the benefits of using the ODD framework is that it allows users to identify gaps or areas with limited data coverage. By examining the table, users can easily spot combinations where there are no data points, indicating a lack of observations in those specific operational scenarios. This insight prompts users to consider the relevance of those missing domains and whether they need to collect additional data or focus their attention on those particular scenarios. The relevance of domains is determined by the user, as they possess the contextual knowledge and understanding of their specific system and operational requirements. Each user or organization may have unique considerations and objectives that influence the selection and definition of relevant domains. Since users are intimately familiar with their system and its intended operating environment, they are best positioned to determine which domains are significant and impactful for their analysis. They can assess the specific operational scenarios, conditions, and constraints that are critical to their system's performance, safety, and effectiveness.

In summary, the ODD framework assists users in converting their data into a structured framework, enabling them to classify domains and analyze the distribution of their datasets within those domains. By identifying areas with limited data coverage, users can proactively address data gaps, design targeted data collection strategies, and gain a deeper understanding of the behavior of their system under various operational scenarios.

9. Discussion

In this section, the research questions that were posed earlier in the thesis are thoroughly addressed and answered. Each research question is carefully examined and discussed, providing a comprehensive analysis and explanation of the findings.

- *How can the diversity of collected data for autonomous maritime systems be analyzed in a scalable way?*

  The thesis proposes the usage of the ODD framework, adapted from the domain of autonomous cars, as a means to systematically analyze the diversity of collected data for autonomous maritime systems.

  To answer the research question, the thesis introduces an ODD taxonomy specifically designed for the maritime domain. This taxonomy allows for the construction of different ODDs that represent specific operational scenarios and conditions in which the autonomous maritime system is expected to operate. By defining these ODDs, the thesis provides a structured framework to categorize and analyze the diversity of collected data.

  The constructed ODDs are then applied to the collected data to assess the extent of its diversity. The thesis demonstrates how the data corresponds to the designed ODDs, highlighting the range of scenarios and conditions captured by the collected data. This analysis provides insights into the scalability of data analysis for autonomous maritime systems, showing that the ODD framework offers a scalable approach to handle and interpret diverse data sources.

- *Using existing solutions from the autonomous driving domain, how can the domains be designed for the autonomous maritime system to operate in?*

  To answer the research question, the thesis conducts a comprehensive analysis of related works in the autonomous driving domain. The aim is to identify relevant approaches and frameworks that can be adapted and applied to the autonomous maritime domain. The thesis explores various existing solutions and methodologies used in the autonomous driving domain to design domains where the system operates.

  Among the different approaches examined, the thesis highlights the ODD as a promising solution. The ODD provides a structured framework for defining the boundaries and characteristics of the operating environment in which an autonomous system functions. By studying the ODD standards introduced by various organizations such as BSI, WISE, AVSC, and ASAM in the autonomous driving domain, the thesis identifies the most suitable ODD standard to be applied in the maritime domain.

  By leveraging the insights gained from the analysis of related works and the selected ODD standard, the thesis proposes a methodology to design domains for the autonomous maritime system. The application of the chosen ODD standard ensures that the domains are appropriately defined, taking into account the specific requirements and characteristics of the maritime environment.

- *What challenges arise when applying domain describing techniques from the autonomous car domain in the autonomous shipping domain?*

  The thesis focuses on identifying the challenges that emerge when attempting to apply domain describing techniques from the autonomous car domain to the autonomous shipping domain.

  To answer the research question, the thesis acknowledges the similarities between the maritime and shipping domains. However, it also recognizes that there are distinct differences between the two domains that need to be considered when applying domain describing techniques. The thesis specifically examines the ODD taxonomy from the autonomous car domain and explores the challenges that arise when adapting it to the maritime domain.

  Through a thorough analysis, the thesis identifies that modifications to the ODD taxonomy are necessary to align it with the specific requirements and characteristics of the maritime domain. Some fields from the original ODD taxonomy may not be relevant or applicable to the maritime context and need to be removed. Additionally, certain new fields may need to be introduced to capture the unique aspects of the autonomous shipping domain.

> By addressing these challenges, the thesis proposes necessary modifications and adjustments to the ODD taxonomy to ensure its suitability for the autonomous shipping domain. This includes refining the taxonomy by removing irrelevant fields and incorporating new fields that address the specific needs of the maritime environment.

During the development of the ODD taxonomy for autonomous maritime vehicles, several discussions were held with the company. As a maritime industry stakeholder, the company's input was invaluable in ensuring that the taxonomy would be relevant to the industry's needs. These discussions involved sharing the initial version of the taxonomy, seeking feedback on its suitability for the maritime environment, and discussing any necessary modifications or additions. The company provided valuable insights into the unique challenges and requirements of autonomous vessels operating in the maritime environment, which helped to shape the final version of the taxonomy. In addition, the company's input was instrumental in ensuring that the resulting ODD would be effective and accurate in meeting their specific needs. This collaboration between the academic research team and the industry stakeholders demonstrates the importance of working together to create solutions that are practical, effective, and tailored to the needs of the end-users. Overall, the discussions with the company during the transition process were essential in developing a taxonomy that would be relevant, useful, and applicable to the real-world challenges of the maritime industry.

In the maritime domain, there was a lack of existing work specifically addressing the application of ODD concepts. This limited the possibility of directly comparing the proposed solution in this thesis to other existing approaches in the maritime domain.

However, it is worth noting that there have been developments in autonomous navigation systems in related domains. One example is the concept of Operational Envelope, as discussed in the work by Fjortoft et al. [36]. The concept of Operational Envelope focuses on the role of humans in the design domain process of autonomous navigation systems.

The thesis also focuses on applying the concept of ODD to real-world sensor data collected by a navigational awareness system. A use case scenario is presented, where the Groke Pro product provided by Groke Technologies is used as the navigational awareness system. The ODDs are applied to the collected sensor data from this system, showcasing how the concept of ODDs can be effectively utilized in analyzing and understanding the data. Throughout the thesis, the usability, flexibility, and effectiveness of applying ODD concepts to the collected sensor data are extensively discussed. The thesis highlights the advantages and benefits of using ODDs in terms of providing a structured and comprehensive approach to analyzing and interpreting sensor data in the context of the navigational awareness system.

10. Conclusions and future work

In conclusion, this thesis highlights the growing popularity of autonomous systems across various industries, including the maritime domain. With the increasing interest in autonomous navigation technologies in the maritime sector, numerous companies and researchers are actively working towards developing solutions to enable autonomous navigation in ships and vessels. However, the design process of these autonomous systems poses certain challenges, particularly in relation to the vast amount of data collected from multiple sensors.

One of the key challenges is the management and understanding of the collected data. As autonomous systems rely heavily on data-driven algorithms and machine learning, it becomes crucial to determine the quality, relevance, and diversity of the data used for training and testing these systems. Additionally, ensuring the readiness of these systems in different situations and environments based on the data they have been trained on is a complex task.

This thesis addresses this challenge by using the concept of ODD. Initially developed for the autonomous car industry, ODD provides a framework for describing the specific domain in which an autonomous system is intended to operate. However, it is important to note that the maritime domain presents distinct characteristics and challenges that differ from those of the autonomous car industry. As a result, the existing ODD concept is not directly applicable to the maritime environment.

To bridge this gap, the thesis takes a proactive approach and proposes a new ODD taxonomy specifically tailored for the maritime domain. The design of this new taxonomy involves carefully considering the unique aspects of maritime operations, such as different types of vessels, diverse environmental conditions, and the state of the system.

This thesis successfully demonstrates the usability and effectiveness of the ODD approach in the maritime domain through a practical use case provided by Groke Technologies. The collaboration with industry experts from Groke Technologies has been instrumental in ensuring that the designed ODD taxonomy aligns with the specific needs and requirements of the maritime industry.

The primary focus of this thesis has been on the application of the ODD concept to the Groke Pro product, which is a situational awareness system developed by Groke Technologies. By utilizing the newly designed ODD taxonomy for the maritime domain, the thesis showcases how the ODD framework can help the analysis process of the data the system has collected.

The outcome of this thesis encompasses two key contributions: the designed ODD taxonomy specifically tailored for the maritime domain and the practical application and demonstration of this taxonomy in a real-life use case. The designed ODD taxonomy for the maritime domain serves as a foundational framework that captures and describes the essential characteristics, constraints, and boundaries of maritime systems. It provides a structured and standardized approach for defining the operational environment in which these systems operate. By creating this taxonomy, the thesis offers a comprehensive and organized framework that can be utilized by researchers, companies, and industry professionals in the development of their own ODD documents.

Furthermore, the practical application of the newly designed ODD taxonomy in a real-life use case highlights its effectiveness and potential impact in real-world scenarios. During the practical application of the ODD taxonomy, several challenges have been encountered and effectively addressed. One of the challenges faced relates to describing certain attributes of the ODD taxonomy based on the available data. For example, determining whether the system is operating in a harbor or in the open sea can be a non-trivial task based solely on the sensor data. To overcome such challenges, this thesis has dedicated attention to exploring potential solutions and discussing the associated complexities.

By implementing the ODD framework within the context of the Groke Pro product, the thesis demonstrates how the taxonomy can be effectively utilized to analyze, interpret, and address the challenges and requirements specific to the maritime domain. This practical application serves as a proof-of-concept, showcasing the relevance and applicability of the designed ODD taxonomy in enhancing the understanding and management of autonomous maritime systems.

In order to demonstrate the practical application of the ODD taxonomy in real-world scenarios, sensor data from the Groke Pro product has been utilized. The use of this sensor data allows for a concrete and tangible representation of the system's operational environment.

To begin, the distribution of the sensor data across different attributes has been carefully analyzed and visualized. This analysis provides valuable insights into the characteristics and variations present in the collected data. By examining the distribution of attributes such as speed, roll, environmental conditions, road structure, and road users, patterns and trends within the data can be identified. Next, the ODDs derived from the newly designed taxonomy are employed to group and categorize these attributes. By utilizing the ODD framework, it becomes possible to define specific domains that capture distinct operational scenarios or contexts. These domains can encompass various combinations of attributes, reflecting the unique conditions and constraints of different maritime environments.

Through the application of the ODDs, the incoming sensor data can be effectively analyzed and classified based on the domain to which it belongs. This classification allows for a deeper understanding of the data and facilitates more accurate interpretations and decisions. By associating the incoming data with specific ODD-defined domains, it becomes possible to assess whether the data aligns with expected scenarios or if it deviates from the defined operational boundaries.

It is important to note that the designed ODD taxonomy presented in this thesis serves as an initial draft and provides a foundation for further iteration and refinement. Different researchers or companies can build upon this base taxonomy and customize it according to their specific needs, taking into consideration the unique characteristics and requirements of their autonomous maritime systems. The introduced taxonomy serves as a starting point and a valuable resource for future endeavors, enabling the development of more comprehensive and tailored ODD documents in the maritime domain.

10.1. Future Work

In the realm of future work, there are several avenues to explore and enhance the ODD taxonomy for the maritime domain. Firstly, the taxonomy can be refined and expanded to incorporate more detailed information for various use cases. As different maritime scenarios and operational environments present unique challenges and requirements, the taxonomy can be enriched with additional attributes, sub-domains, or specific conditions that are relevant to specific use cases. This would enable a more granular representation of the operational domains and improve the accuracy and applicability of the ODD framework.

Furthermore, the taxonomy can be extended to cover corner cases or exceptional scenarios that may occur in maritime operations. These corner cases often possess unique characteristics or conditions that differ significantly from the standard operational domains. By including these corner cases in the taxonomy, the system can be better prepared to handle and respond to such exceptional situations, ensuring enhanced safety and performance.

In addition to refining the taxonomy, it is also valuable to explore and compare alternative concepts and approaches alongside the ODD framework. This could involve investigating other domain-describing techniques or methodologies utilized in different industries or research domains. By examining and contrasting these alternative approaches, valuable insights can be gained, potentially leading to improvements or synergies with the ODD concept in the maritime domain.

Moreover, it is crucial to evaluate the performance and effectiveness of the ODD taxonomy in real-world scenarios. Conducting comprehensive validation studies and experiments can provide valuable feedback on the practical applicability and limitations of the taxonomy. This may involve collecting and analyzing additional sensor data from various maritime operations, assessing the accuracy and reliability of the domain classification, and seeking input from industry experts to validate the taxonomy's suitability for different use cases.

Lastly, future work should also focus on integrating the ODD taxonomy into existing maritime autonomous systems and evaluating its impact on system performance, decision-making processes, and overall safety. By implementing the ODD framework in real-world applications and gathering feedback from system users and stakeholders, further insights can be gained regarding its practical implications and potential improvements.

By addressing these future directions, the ODD taxonomy can be continually refined, expanded, and validated, ultimately enhancing its usefulness and applicability in autonomous maritime systems.

References

[1] L. Elkins, D. Sellers, and W. R. Monach, "The autonomous maritime navigation (amn) project: Field tests, autonomous and cooperative behaviors, data fusion, sensors, and vehicles," *Journal of Field Robotics*, vol. 27, no. 6, pp. 790–818, 2010.

[2] S. Thombre, Z. Zhao, H. Ramm-Schmidt, J. M. Vallet García, T. Malkamäki, S. Nikolskiy, T. Hammarberg, H. Nuortie, M. Z. H. Bhuiyan, S. Särkkä, and V. V. Lehtola, "Sensors and ai techniques for situational awareness in autonomous ships: A review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 1, pp. 64–83, 2022.

[3] G. Technologies. (2023) Groke technologies. Accessed on: 2023-04-18. [Online]. Available: https://www.groke-tech.com/

[4] J. Kocić, N. Jovičić, and V. Drndarević, "Sensors and sensor fusion in autonomous vehicles," in *2018 26th Telecommunications Forum (TELFOR)*, 2018, pp. 420–425.

[5] N. Reddy, H. Farah, Y. Huang, T. Dekker, and B. Van Arem, "Operational design domain requirements for improved performance of lane assistance systems: A field test study in the netherlands," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 1, pp. 237–252, 2020.

[6] E. Ziajka-Poznańska and J. Montewka, "Costs and benefits of autonomous shipping—a literature review," *Applied Sciences*, vol. 11, no. 10, 2021. [Online]. Available: https://www.mdpi.com/2076-3417/11/10/4553

[7] O. Zendel, K. Honauer, M. Murschitz, M. Humenberger, and G. F. Domínguez, "Analyzing computer vision data — the good, the bad and the ugly," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 6670–6680.

[8] S. Thombre, Z. Zhao, H. Ramm-Schmidt, J. M. Vallet García, T. Malkamäki, S. Nikolskiy, T. Hammarberg, H. Nuortie, M. Z. H. Bhuiyan, S. Särkkä, and V. V. Lehtola, "Sensors and ai techniques for situational awareness in autonomous ships: A review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 1, pp. 64–83, 2022.

[9] R. Collis, "Lidar," *Applied optics*, vol. 9, no. 8, pp. 1782–1788, 1970.

[10] M. I. Skolnik, "Introduction to radar," *Radar handbook*, vol. 2, p. 21, 1962.

[11] P. D. Groves, "Principles of gnss, inertial, and multisensor integrated navigation systems, [book review]," *IEEE Aerospace and Electronic Systems Magazine*, vol. 30, no. 2, pp. 26–27, 2015.

[12] B. J. Tetreault, "Use of the automatic identification system (ais) for maritime domain awareness (mda)," in *Proceedings of Oceans 2005 Mts/Ieee.* Ieee, 2005, pp. 1590–1594.

[13] A. Faisal, M. Kamruzzaman, T. Yigitcanlar, and G. Currie, "Understanding autonomous vehicles: A systematic literature review on capability, impact, planning and policy," *Journal of Transport and Land Use*, vol. 12, no. 1, pp. 45–72, 2019. [Online]. Available: https://www.jstor.org/stable/26911258

[14] J. M. Anderson, N. Kalra, K. D. Stanley, P. Sorensen, C. Samaras, and T. A. Oluwatola, *Autonomous Vehicle Technology: A Guide for Policymakers.* Santa Monica, CA: RAND Corporation, 2016.

[15] SAE. (2018) Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. Accessed on: 2023-07-11. [Online]. Available: https://www.sae.org/standards/content/j3016_201806

[16] ——. (2018) Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. Accessed on: 2023-03-10. [Online]. Available: https://www.sae.org/standards/content/j3016_201806/

[17] B. S. Limited. (2020) Operational design domain (odd) taxonomy for an automated driving system (ads) – specification. Accessed on: 2023-03-12. [Online]. Available: https://www.bsigroup.com/globalassets/localfiles/en-gb/cav/pas1883.pdf

[18] Asam. (2020) Asam openodd: Concept paper. Accessed on: 2023-01-21. [Online]. Available: https://code.asam.net/simulation/proposal/openodd/-/jobs/artifacts/master/raw/OpenODD_Project_Proposal.html?job=html

[19] C. Krzysztof, "Automated driving system (ads) high--level quality requirements analysis driving behavior safety," *Waterloo Intelligent Systems Engineering (WISE) Lab*, 2018.

[20] G. Technologies. (2023) Products. Accessed on: 2023-05-09. [Online]. Available: https://www.groke-tech.com/en/products

[21] B. Elasticsearch, "Elasticsearch," *software], version*, vol. 6, no. 1, 2018, accessed on: 2023-05-10.

[22] C. Gormley and Z. Tong, *Elasticsearch: the definitive guide: a distributed real-time search and analytics engine.* " O'Reilly Media, Inc.", 2015.

[23] devopsideas.com. (2022) Different elasticsearch components and what they mean in 5 mins. Accessed on: 2023-05-10. [Online]. Available: https://devopsideas.com/different-elasticsearch-components-and-what-they-mean-in-5-mins/

[24] ——. (2022) Elastic search component relation. Accessed on: 2023-05-10. [Online]. Available: https://devopsideas.com/wp-content/uploads/2018/11/Elasticsearch_component_relation.png

[25] Elastic. (2023) Query dsl. Accessed on: 2023-05-10. [Online]. Available: https://www.elastic.co/guide/en/elasticsearch/reference/current/query-dsl.html

[26] ——. (2023) Match query. Accessed on: 2023-05-10. [Online]. Available: https://www.elastic.co/guide/en/elasticsearch/reference/current/query-dsl-match-query.html

[27] M. Ito, "Odd description methods for automated driving vehicle and verifiability for safety," *JUCS - Journal of Universal Computer Science*, vol. 27, pp. 796–810, 08 2021.

[28] AVSC. (2023) Automated vehicle safety consortium™. Accessed on: 2023-04-23. [Online]. Available: https://avsc.sae-itc.org

[29] T. A. V. S. Consortium™. (2020) Avsc best practice for describing an operational design domain: Conceptual framework and lexicon. Accessed on: 2023-03-12. [Online]. Available: https://www.sae.org/standards/content/avsc00002202004/

[30] J. Erz, B. Schütt, T. Braun, H. Guissouma, and E. Sax, "Towards an ontology that reconciles the operational design domain, scenario-based testing, and automated vehicle architectures," in *2022 IEEE International Systems Conference (SysCon)*, 2022, pp. 1–8.

[31] ASAM. (2022) Asam openscenario. Accessed on: 2023-04-23. [Online]. Available: https://www.asam.net/project-detail/asam-openscenario-v20-1/

[32] Ø. J. Rødseth, L. A. Lien Wennersberg, and H. Nordahl, "Towards approval of autonomous ship systems by their operational envelope," *Journal of Marine Science and Technology*, vol. 27, no. 1, pp. 67–76, Mar 2022. [Online]. Available: https://doi.org/10.1007/s00773-021-00815-z

[33] N. A. C. I. Society. (2012) Natural earth vector. Accessed on: 2023-04-04. [Online]. Available: https://github.com/nvkelso/natural-earth-vector/blob/master/geojson/ne_10m_coastline.geojson

[34] M. Dolatshah, A. Hadian, and B. Minaei-Bidgoli, "Ball*-tree: Efficient spatial indexing for constrained nearest-neighbor search in metric spaces," *arXiv preprint arXiv:1511.00628*, 2015.

[35] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[36] K. Fjørtoft and Ø. Rødseth, “Using the operational envelope to make autonomous ships safer,” in *The 30th European Safety and Reliability Conference, Venice, Italy*, 2020.