AISHIP: An Ontology for Extended Vessel Representation and Multimodal Data Integration

Simon Burbach*, Lennart Mackert* and Maria Maleshkova*

Professorship of Data Engineering, Helmut Schmidt University, Holstenhofweg 85, 22043 Hamburg, Germany

Abstract

The maritime sector generates vast amounts of heterogeneous data due to dense global vessel traffic. This data offers significant potential for improving operational efficiency and ensuring safety at sea. However, its effective use is hindered by a lack of semantic alignment between diverse information sources. To address this challenge, we present AISHIP, an ontology designed to unify and semantically enrich maritime data. AISHIP extends the existing VesselAI ontology by incorporating enhanced vessel characteristics, additional trajectory information, multimodal vessel representations, and a detailed conceptualization of propulsion systems. The primary objective of AISHIP is to serve as a semantic interface that facilitates the integration of heterogeneous maritime datasets, enabling consistent annotation, querying, and reasoning across systems.

By providing a shared semantic foundation, AISHIP supports a range of maritime applications, including vessel behavior analysis, fleet management, and search and rescue operations. These tasks benefit from harmonized data representations, which improve analytical precision and operational decision-making. We discuss relevant use cases to illustrate the ontology's practical value and evaluate its design to demonstrate its quality. AISHIP represents an extensible and reusable resource, aligning with FAIR principles, and offers strong potential for adoption in future maritime analytics and digital twin frameworks.

Ontology: https://burbachs.github.io/AISHIPOntology/AISHIP.owl

GitHub: https://github.com/BurbachS/AISHIPOntology

License: CC BY-NC-SA 4.0

DOI: 10.5281/zenodo.16704863

Keywords

Maritime Ontology, Vessel Data Integration, Ontology Engineering, Semantic Interoperability, Multimodal Data Representation, Propulsion Systems

1. Introduction

Maritime transportation plays a crucial role in global logistics, with approximately 80% of international trade being transported by vessels navigating international waters at any given time [1]. In addition to that, the maritime domain is also used for passenger transport, military, and private traffic with different kinds of vessels. This results in a substantial number and wide variety of vessels operating simultaneously in both national and international waters [2, 3]. Nearly all sea traffic generates relevant data via onboard sensor systems, as well as communication with other ships and land stations via the automatic identification system (AIS) [2, 3, 4]. This data mostly represents information about the ship itself (e.g., name, flag, vessel type), but also includes its position, speed, course, and other details transmitted via very high frequency radio.

All collected data constitutes a heterogeneous mass of information, typically existing as an unstructured and unrelated collection. This lack of semantic alignment significantly hinders critical applications such as anomaly detection, trajectory prediction, and fleet optimization from unleashing their full potential, making a common semantic model necessary [5, 6, 7, 8, 9]. Such a model provides the shared conceptual basis required to integrate heterogeneous datasets into a coherent framework. Without

ISWC 2025 Companion Volume, November 2–6, 2025, Nara, Japan

burbachs@hsu-hh.de (S. Burbach); lennart.mackert@hsu-hh.de (L. Mackert); maleshkm@hsu-hh.de (M. Maleshkova)
https://www.hsu-hh.de/dataeng/en/team/simon-burbach/ (S. Burbach);
https://www.hsu-hh.de/dataeng/en/team/maria-maleshkova/ (M. Maleshkova)
0009-0009-8373-6622 (S. Burbach); 0000-0003-3458-4748 (M. Maleshkova)

such a model, these tasks remain fragmented and prone to errors, as they must rely on data originating from diverse and often incompatible sources to enable well-founded and informed decision-making processes, which are essential to ensure human safety.

To organize, relate, and integrate such data, ontologies are commonly employed [5, 8]. In the maritime domain, ontologies allow information about ships and vessel traffic to be used for consistent classification, grouping entities that share the same characteristics, even if expressed in different forms. For example, different representations of ship characteristics can be standardized to fit into a predefined, uniform ship description [5]. Furthermore, these also serve as the basis for semantic interoperability across information systems, facilitating standardized communication and data exchange.

The main contribution of this paper is the introduction of the AISHIP ontology that enables the unified and detailed representation of vessels. Our work builds directly on the existing VesselAI ontology [5], as it is mostly centered around trajectories and basic vessel characteristics. We expand it with additional concepts and relationships. By integrating information from heterogeneous data sources, this extended model allows for a richer and more precise description of ship-related entities. AISHIP incorporates additional features to improve vessel identification and detailed classifications of propulsion systems (PSs) and engine configurations, as well as extended trajectory data.

Through this contribution, we aim to support the development of interoperable, extensible systems for maritime analytics and to foster collaboration through the use of a shared semantic foundation. In doing so, we specifically address the overarching goal of enabling more effective integration and interpretation of maritime data. The ontology not only enables semantic alignment across heterogeneous sources but also provides the structural basis for advanced analytics, which particularly includes anomaly detection and behavior analysis. By establishing consistent representations of ship characteristics, trajectories, and contextual factors, the model lays the groundwork for identifying deviations from expected patterns and supporting data-driven maritime decision-making. In addition, the ontology enables the analysis of historical ship data, allowing researchers and practitioners to uncover long-term operational patterns and trends.

Our work specifically makes the following contributions:

- The **AISHIP ontology** that extends the VesselAI ontology with further vessel characteristics, trajectory data, options for multi-modal vessel representation, and methods for modeling ship PSs;
- **48 competency questions (CQs)** divided into four categories that reflect questions of the real world, which were used to direct the ontology development;
- An **ontology evaluation** that assesses domain coverage.

The rest of the paper is structured as follows. First, Section 2 presents data models that are related to our aim, and show potential for reuse and extension. Then, Section 3 discusses motivating use-cases that can be supported by AISHIP. Afterwards, we describe the ontology development process, as well as its core modules, together with some axiomatic restrictions (see Section 4). In Section 5, we present the evaluation process to date and outline future directions for ontology validation. We conclude the paper with a summary and present tasks for future work in Section 6.

2. Related Data Models

Several established ontologies and data models are relevant for addressing the objectives of our work and provide useful insights for the development process. One example of such an ontology is the VesselAI ontology [5]. The ontology was developed as part of the project ”VesselAI”, funded by the Horizon 2020 Framework Programme of the European Union [10]. The VesselAI ontology provides a structured semantic model for maritime knowledge associated with vessel trajectories. The authors structured the content in four main modules: 1) Trajectories of Moving Objects, 2) Vessel Characteristics, 3) Contextual Information, and 4) Events and Weather Conditions. Trajectories consist of temporally annotated trajectory parts, each representing the connection of single geographical positions taken from AIS signals. Each trajectory part is annotated with a time stamp, travel-related data (e.g., speed and course over ground), as well as events (e.g., anchoring) and weather conditions (e.g., heavy precipitation or freezing rain). Weather conditions follow the code registry of the World Meteorological Organization[1]. These annotations are important, as environmental conditions and events influence the vessel's behavior and operational patterns.

Regions of interest, like ports, fishing areas, and environmentally protected areas, are also modeled. They are essential for linking the activities of vessels to relevant zones at sea. This allows for further analysis of vessel traffic, and compliance with protected areas is possible. These spatial concepts are represented by the OGC GeoSPARQL standard[2] [11]. GeoSPARQL enables the modeling of spatial relationships (e.g., inside, intersects, within) and enables more advanced queries through SPARQL endpoints. This integration facilitates interoperability with existing spatial datasets and ensures compliance with widely accepted geospatial standards, which is critical for maritime data analytics and decision-support applications.

Another relevant data model is the Common Information Sharing Environment (CISE)[3] [12]. CISE is an interoperability framework promoted by the European Union to enable secure and standardized information exchange among maritime authorities. CISE defines a common semantic structure for sharing data about vessels, their activities, and related events across different systems and domains. CISE provides a common structure for maritime data, which simplifies the integration of heterogeneous information sources and improves interoperability. This supports better cross-border situational awareness and strengthens maritime safety.

The W3C Semantic Sensor Network (SSN) ontology and its core module Sensor, Observation, Sample, Actuator (SOSA) provide a widely adopted standard for describing sensors and their observations [13, 14]. SSN/SOSA enables the modeling of sensors, representing their outputs in a semantically consistent way. This allows the integration of heterogeneous data from different systems while considering metadata. Standardized sensor descriptions are a necessity for tasks like anomaly detection and predictive maintenance.

Based on this review of related data models, VesselAI offers the greatest potential for extension in our use cases. Its detailed trajectory modeling provides a strong foundation for analyzing vessel behavior. However, it lacks critical features that are essential for our objectives. While the VesselAI ontology provides a solid foundation for modeling vessel trajectories, its scope is largely limited to AIS-derived attributes. It does not capture the full range of ship identification features required for comprehensive integration. Furthermore, it lacks the representation of aggregated values, which are crucial for behavioral analysis.

3. Motivation and Usage Scenarios

In the maritime domain, a wide range of heterogeneous data sources, like AIS tracks[4], technical ship specifications, and registry databases, must be combined and analyzed to support decision-making processes in logistics, fleet management, monitoring, and research. However, these datasets often lack a unified semantic structure as they are available in different formats that are not compatible, which limits interoperability, reuse, and automated reasoning [5, 8, 15].

The AISHIP ontology addresses this challenge by providing a formalized and extensible conceptual model for representing ships and their properties in a machine-interpretable way. It builds directly on the VesselAI ontology, which was systematically extended to cover additional ship-related information, relationships, and data dimensions. It enables the integration of static and dynamic data across different systems and supports consistent annotation, querying, and interpretation of maritime information.

[1] http://codes.wmo.int/306/4678, accessed: 25.07.2025
[2] https://www.ogc.org/standards/geosparql/, accessed: 27.07.2025
[3] https://emsa.europa.eu/cise-documentation/cise-data-model-1.5.3/, accessed: 27.07.2025
[4] https://www.marinetraffic.com/, accessed: 01.08.2025

By providing a unified representation of vessel-related information that includes static attributes (e.g., dimensions, ownership, engine types), dynamic behavior (e.g., trajectory data), and system-level components (e.g., propulsion), the ontology offers a foundation for numerous use cases in both commercial and operational maritime contexts.

Through standardized identifiers (e.g., Maritime Mobile Service Identity – MMSI, hull identification number – HIN), descriptive properties (e.g., ship type, dimensions, flag), and diverse data modalities to represent ships (e.g., image, sonar signature, and textual description), ships can be uniquely and consistently recognized, even when source systems differ. In addition, by formally linking ships with their historical and planned trajectories, the ontology supports predictive analytics such as trajectory forecasting, behavior pattern recognition, and route anomaly detection. This capability is particularly valuable for risk management, monitoring, and real-time operational decision-making.

The AISHIP ontology offers a wide range of potential application scenarios across different areas of the maritime domain. By providing a formal and semantically rich representation of vessels, their components, and behaviors, the ontology supports tasks that rely on structured, interoperable, and machine-readable maritime knowledge:

Fleet Management and Operational Planning. In the domain of fleet management, the ontology provides standardized and machine-readable descriptions of vessel characteristics, such as the ship type, the associated PSs, the navigational status, and the past trajectories. This standardization supports better integration of heterogeneous fleet data, enabling consistent comparison and classification of vessels across different systems and operators. By relying on a common vocabulary and formal structure, the ontology reduces ambiguity, improves data interoperability, and facilitates automated reasoning about fleet composition, readiness, and compliance. Furthermore, the integration of inferred and aggregated behavioral data, like region-specific operational stays, adds a predictive and context-sensitive dimension. As a result, operators can make more informed decisions when assigning ships to tasks, planning maintenance schedules, or evaluating environmental and regulatory criteria across the fleet.

Search and Rescue (SAR) Operations. In emergency situations, such as ship distress or man-overboard events, time-critical information is essential. The ontology supports SAR missions by enabling fast semantic querying of further relevant vessel trajectory data: estimated position and predicted trajectory. It also helps to identify nearby vessels with suitable equipment or capacities for assistance by analyzing predicted trajectories, improving coordination, and response times.

Semantic Data Integration, Search and Analysis. The ontology acts as a common semantic layer across disparate data systems, such as AIS feeds, ship registries, classification society databases, and onboard monitoring platforms. This harmonization allows for federated querying, automated data linking, and cross-system analysis using shared concepts and unified, controlled vocabularies.

Digital Twins and Life Cycle Simulations. The ontology provides a structural backbone for building digital twins of vessels by modeling the types of propulsion systems and their interconnections. Formalizing the relationships between components such as engines and propellers establishes the basic foundation for representing operational states and long-term performance characteristics. By extending the modules with additional system components and state information in the future, they support virtual testing, maintenance simulations, and system-level evaluations, thereby enabling more comprehensive analyses over the vessel's entire life cycle.

4. Ontology Development

For the development of our ontology, the SABIOx methodology [16] was chosen to guide the process. The SABIOx methodology follows a five-phase life cycle comprising the *requirement phase*, *setup phase*, *capture phase*, *design phase*, and *implementation phase*. In the *requirement phase*, the ontology's purpose, scope, and requirements are defined. The *setup phase* establishes modeling decisions such as language, foundational ontology, and reuse strategy. During the *capture phase*, domain concepts and axioms are identified and organized into a conceptual model, whereas the *design phase* focuses on modeling these concepts into a coherent ontology schema. Finally, the *implementation phase* involves formalizing the ontology in OWL. Throughout the entire process, interdisciplinary collaboration with experts from the maritime domain was essential to ensure conceptual accuracy. CQs were defined at the beginning and updated frequently to set the boundaries and contents of the ontology, guiding the development process. Additionally, diverse knowledge and data sources were identified that needed to be represented in order to make the ontology compatible with multimodal data.

Based on the defined requirements, Ontology Design Patterns (ODPs) were identified and analyzed to enable the reuse and extension of existing ontologies and their content. AISHIP incorporates widely adopted ontologies and vocabularies such as QUDT, Schema.org, NEXIF, NFO, and NIE to ensure semantic richness and interoperability. Following further principles of openness and reusability promoted within the Semantic Web community, the AISHIP ontology is published under an open license, accessible through a persistent URI, and archived on Zenodo with a DOI (10.5281/zenodo.16704863) to guarantee long-term availability and proper citation.

4.1. Core Modules

ODPs serve as established best practices that address recurring issues in ontology development by providing reusable modeling solutions. Their application improves the overall coherence and interpretability of ontologies, promotes knowledge reuse, and simplifies the processes of integrating and aligning heterogeneous models [17, 18].

The main ontology that AISHIP (prefix `ais`) builds upon is the VesselAI ontology (prefix `vai`). While VesselAI primarily focuses on ship trajectories, it provides a solid foundation for further extension. In our work, this served as the starting point to integrate additional ship-related information relevant to our specific use cases. AISHIP improves VesselAI's standardized descriptions of vessels with five key modules: *Enhanced Vessel Characteristics*, *Enhanced Trajectory Data*, *Enhanced Contextual Information*, *Motorization and Propulsion*, and *Multimodal Data Representation*. They collectively add 73 classes, 38 object properties, and three data properties.

1) Enhanced Vessel Characteristics. The first module (see Figure 1) extends the ontology with ship-related information derived from AIS datasets. Based on these datasets, we identified attributes that were not yet represented in VesselAI. These include the dimension specifications for the GPS antenna placement, the ship owner, and the HIN.

To improve behavioral analysis of vessels, modeling a ship's typical location is essential. The property `ais:TypicalLocation` is represented as an OGC Geometry (e.g., polygon) calculated over a given time period. This way, AISHIP represents the expected location of a vessel. In addition, aggregated values such as `ais:hasAverageSpeedOverGround` are relevant, as they correlate with the ship type.

Another important enhancement introduced by AISHIP concerns datatype restrictions for both dynamic and static AIS data. VesselAI initially allowed only the string datatype for measurements like speed and course, as well as vessel dimensions, which is semantically inaccurate. To address this, we integrated the QUDT (prefix `qudt`) ontology. Consequently, AISHIP models data properties such as `ais:hasSpeedOverGround`, `ais:hasHeading`, and `ais:hasGrossTn` as object properties, pointing to `qudt:QuantityValue`.

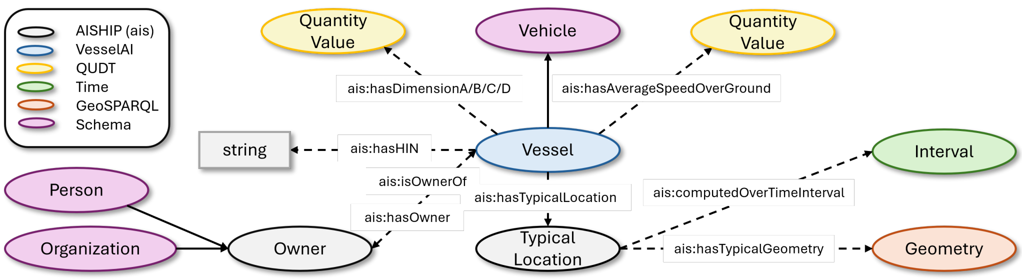

Figure 1: Further characteristics of a ship, including its HIN, owner, and dimensional information of the GPS antenna. Aggregated values like the average speed over ground and the typical location are relevant for analyzing the vessel behavior

2) Enhanced Trajectory Data. VesselAI primarily focuses on modeling vessel trajectories. To extend this capability, we introduced additional dynamic and aggregated properties, like `ais:hasDraught` and `ais:hasAverageTrajectorySpeedOverGround`. These attributes provide further insights into vessel behavior, for example, by indicating activities like loading traded goods. In the context of behavioral analysis, we also added a new subclass, `ais:PredictedTrajectory`, which specializes in representing predicted vessel movements. This class is defined as a subclass of `vai:SyntheticTrajectory`. Figure 2 shows the relationships:

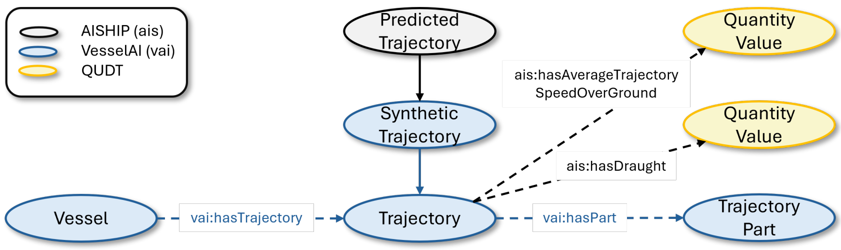

Figure 2: A predicted trajectory is a synthetic trajectory, calculated based on historical vessel behavior and contextual factors such as weather, traffic density, and navigational constraints

3) Enhanced Contextual Information. The extension of contextual information adds another subclass `ais:RestrictedArea` to `vai:Place`, since VesselAI does not provide the ability to model restricted areas specifically. Restricted areas, according to our definition, can be regions where access or navigation is limited due to security, safety, or legal regulations. It is distinguished from `vai:ProtectedArea`, which represents areas of environmental constraints.

4) Motorization and Propulsion. A major contribution of our work is the introduction of a module that models ship PSs, which can be complex and composed of multiple main components (see Figure 3). In this model, a ship is associated with a PS, where each PS can be classified according to its energy source. Vessels most commonly feature a motorized PS (MPS). Others may rely solely on muscle power (Muscle-powered propulsion system, MPPS), exclusively use wind for thrust (Pure wind propulsion system, PWPS), or employ wind as assisted propulsion (Wind-assisted propulsion system, WAPS) to support motorized propulsion. The latter is increasingly used in large cargo and tanker ships to reduce fuel consumption and minimize environmental impact by carbonization [19].

However, `ais:PropulsionSystem` only describes the concept of a PS, not the specific way in which propulsion is generated. To capture this, we define an inverse property-pair that links a PS to the class `ais:PropulsionTechnology`. Propulsion Technologies (PTs) actively or passively generate thrust for the vessel. They are classified as well into the four main types: `ais:MusclePoweredPropulsionTechnology`, `ais:WindAssistedPropulsionTechnology`, `ais:PureWindPropulsionTechnology`, and `ais:MotorizedPropulsionTechnology`. The categorization of PTs utilizing wind follows the International Maritime Organization (IMO) classification [19]. Our ontology therefore distinguishes between systems that are exclusively used in WAPS (e.g., Suction Wings, Rotor Sails, Wind Assisting Soft Sails) and those designed for pure wind propulsion (e.g., Pure Wind Soft Sail). In contrast, muscle-powered PTs (MPPTs) may include technologies such as paddles.

Motorized PTs (MPTs) represent the more technically complex PTs and are further divided into two subclasses: `ais:JetPropulsor` and `ais:BladedPropulsor`. Jet propulsors generate thrust through high-pressure water or gas jets, a principle widely used in high-speed vessels [20, 21].

MPTs are frequently exposed to cavitation. Cavitation is a phenomenon where vapor bubbles form and collapse on propeller surfaces. They can cause erosion, which reduces efficiency. To mitigate this, MPTs incorporate anti-cavitation mechanisms, which are currently modeled in a generic manner without specialized subtypes.

Among all PTs, bladed propulsors represent the most widely used systems in modern shipping. For these systems, characteristics such as the number of blades and propeller diameter play a crucial role, depending on the vessel type. Both of which are captured in our ontology.

Bladed propulsors also vary in design and are adapted to specific ship types. Our model differentiates between older designs like the paddle wheel, specialized configurations such as the Voith Schneider Propeller, and conventional screw propellers. The class `ais:ScrewPropeller` includes the most common variants (e.g., fixed pitch propeller, controllable pitch propeller, and counter-rotating propeller) [20, 21].

To represent how MPTs are powered, we modeled `vai:Vessel` as a subclass of `schema:Vehicle`, to enable motorization details through `schema:EngineSpecification`. We introduced several subclasses of engines. They are distinguished whether they have pistons or not, or if they represent a combination of different engine types [22]. Examples include standard engines such as `ais:DieselEngine` and `ais:SteamTurbineEngine`, as well as combined engine systems like CODOG (combination of diesel-engine or gas-turbine) and CONAS (combination of nuclear power and steam-turbine) [22].

Finally, engines themselves are connected to an `ais:PowerGenerationUnit`, which provides the necessary energy for propulsion (e.g., nuclear reactor, diesel combustion unit, solar array).

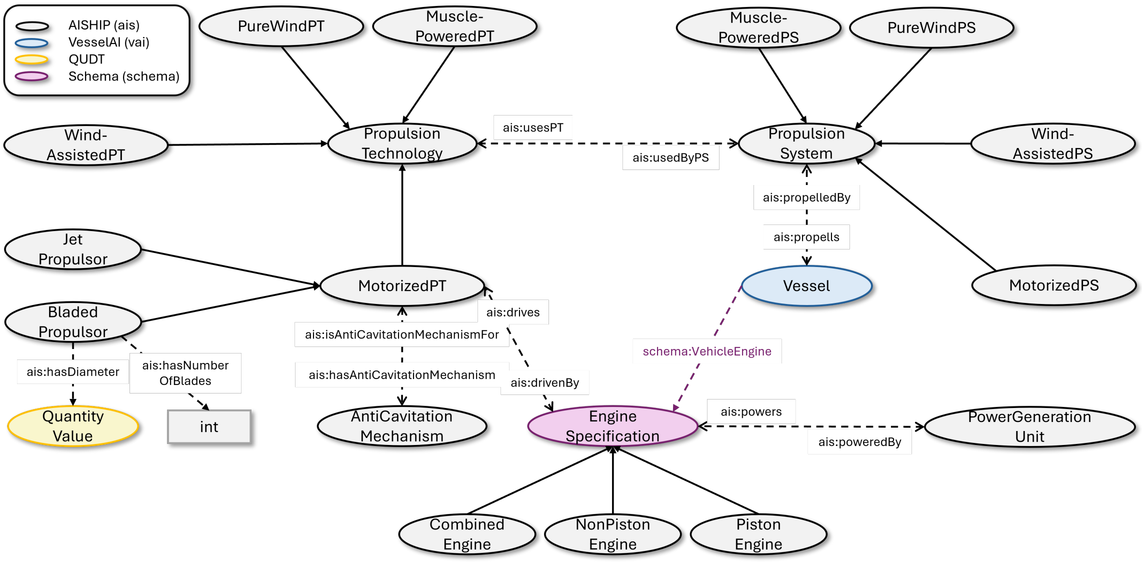

Figure 3: The main classes of the module *Motorization and Propulsion*

5) Multimodal Data Representation. Ships can be described not only through AIS-related data but also by leveraging additional sources of information. Common examples include visual representations such as images, as well as textual descriptions that provide further context. Furthermore, ships exhibit distinctive sonar signatures, which can be used to infer the vessel type (see Figure 4).

To ensure accuracy, these multimodal representations must include metadata such as the capture date and modification date, since ships may undergo structural or visual changes over time. For instance, a hull repaint can make visual recognition unreliable. Similarly, sonar characteristics can change due to modifications in the PS, hull shape, cargo load, or the application of insulation materials. To model this metadata in a standardized way, we integrated concepts from the NEXIF, NFO, and NIE ontologies, which provide established vocabularies for describing media properties, file information, and related entities.

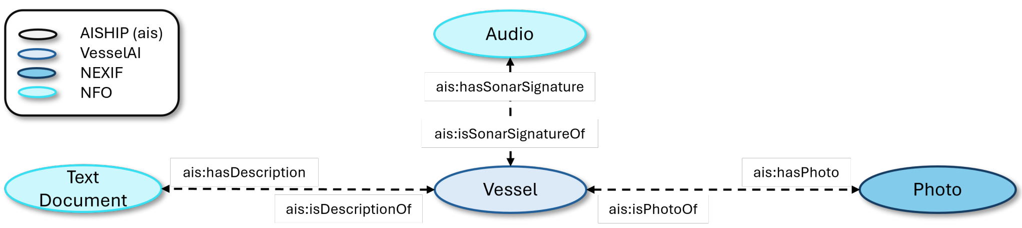

Figure 4: A vessel can be characterized by a textual description, a photo, and a sonar signature

4.2. Restricting Data Types

As already mentioned in the context of the module *Enhanced Vessel Characteristics*, we extended the scope of how certain numerical values can be expressed. To achieve this, we specified the datatypes of several data properties by creating new object properties following the same naming convention, linking them to `qudt:QuantityValue`. This design allows QUDT's mechanisms for unit annotation to be applied, which significantly improves semantic interpretability. The newly introduced object properties include `ais:hasSpeedOverGround`, `ais:hasCourseOverGround`, `ais:hasHeading`, `ais:hasGrossTn`, `ais:hasDeadweight`, `ais:hasLength`, `ais:hasWidth`, `ais:hasHeight`, and `ais:hasActualHeight`. The same modeling approach applies to all other properties we added that reference static or dynamic numerical values that have units.

This modeling strategy offers several advantages. First, it ensures semantic correctness, as numerical values are represented in a logically precise and standardized way. Furthermore, by using QUDT's unit vocabulary, ambiguity regarding measurement units is avoided, which facilitates consistent interpretation of data across different sources. In addition, this approach enables reasoning over numerical values, such as comparisons, and allows validation of value ranges through SHACL, for instance, when checking if the course over ground lies within the range from 0° to 360°. Without this approach, only string-based comparisons would have been possible, which are semantically weak and prone to errors.

However, this decision also introduces a significant drawback, as the existing published datasets of VesselAI need to be adapted accordingly to comply with the new object property definitions and QUDT-based structure.

4.3. T-Box Restrictions

To ensure semantic consistency and avoid modeling redundancies, we introduce a set of T-Box restrictions that govern class disjointness, property constraints, and domain-specific logical dependencies between propulsion-related classes. These axioms provide a foundation for reasoning and validation in our ontology.

Class Disjointness. All subclasses of PS and all subclasses of PT are declared as pairwise disjoint. This ensures that no individual can simultaneously belong to multiple sibling classes. This restriction maintains a clear and unambiguous classification.

Property Range Restrictions. Several properties are restricted to appropriate datatypes and units. Since the properties refer to `qudt:QuantityValue`, the associated datatype and unit constraints are applied through QUDT's structure. For example: `ais:hasDraught` (type: decimal, unit: meters), `ais:hasSpeedOverGround` (type: decimal, unit: knots), and `ais:hasDiameter` (type: decimal, unit: centimeters). For navigational attributes such as `ais:hasHeading` and `ais:hasCourseOverGround`, we enforce realistic value ranges from 0° to 360°, as values outside this range are not physically meaningful.

Access Control Between Propulsion Systems and Propulsion Technologies. To prevent inconsistencies and redundant object property definitions, we constrain the relationships between PSs and PTs:

- Muscle-powered propulsion systems can only utilize muscle-powered propulsion technology.
- Motorized propulsion systems can only utilize motorized propulsion technology.
- Pure wind propulsion systems can only utilize pure wind propulsion technology.
- Wind-assisted propulsion systems can only utilize wind-assisted propulsion technology.

These constraints eliminate the need to redundantly model a separate direct object property for each pair of subclasses, reducing complexity while preserving semantic precision.

Exclusivity of Pure Wind Propulsion. Any vessel that employs a PWPS is restricted from having any additional PS. This reflects the domain rule that vessels relying solely on wind propulsion cannot combine it with other system types. The restriction represented as a logical expression:

$$\forall x\Big(\big(x : \text{Vessel} \land \exists y\big(\text{propelledBy}(x, y) \land y : \text{PWPS}\big)\big) \rightarrow \neg\exists z\big(\text{propelledBy}(x, z) \land z \not: \text{PWPS}\big)\Big)$$

Incompatibility Between PWPS and Motorized or Wind-Assisted Systems. A stricter consistency rule specifies that vessels equipped with an MPS or a WAPS must not include any PWPS. This enforces the interpretation that wind assistance only supplements motorized propulsion, never replacing it entirely. The restriction represented as a logical expression:

$$\forall x\Big(\big(\exists y(\text{propelledBy}(x, y) \land (y : \text{MPS} \lor y : \text{WAPS}))\big) \rightarrow \neg\big(\exists z(\text{propelledBy}(x, z) \land z : \text{PWPS})\big)\Big)$$

Wind-Assisted Propulsion Requires Motorization. A final consistency constraint ensures that vessels equipped with WAPS must also include an MPS. This requirement reflects the semantic meaning of “wind-assisted”: wind propulsion serves as a supplementary mechanism and not as the sole means of propulsion. Without an MPS, a vessel would rely entirely on wind power, which contradicts the intended definition of wind assistance. The restriction represented as a logical expression:

$$\forall x\Big(x : \text{VesselWithWAPS} \leftrightarrow \big(x : \text{Vessel} \land \exists y\big(\text{propelledBy}(x, y) \land y : \text{WAPS}\big)\big)\Big)$$

$$\forall x\Big(x : \text{VesselWithWAPS} \rightarrow \exists y\big(\text{propelledBy}(x, y) \land y : \text{MPS}\big)\Big)$$

5. Evaluation

The evaluation of the presented ontology was carried out in alignment with the SABIOx methodology, applying a twofold strategy that combines verification and validation [16]. Verification aims to check the internal consistency and correctness of the ontology, ensuring that its structure, classes, and relationships have been implemented according to the specified design criteria and formal requirements. In contrast, validation is concerned with assessing the usefulness of the ontology and its adequacy in addressing real-world scenarios. It examines whether the ontology provides sufficient power and semantic clarity to support the intended application tasks effectively.

5.1. Ontology Verification

The requirements for the ontology are specified through a structured collection of CQs, which serve as a foundational instrument for guiding and evaluating ontology development. These questions define the types of information the ontology should be able to address and thereby help determine the essential domain knowledge that must be explicitly represented within the model. The formulation of the CQs was informed by three primary sources: the analysis of stakeholder requirements, the examination of specific use cases relevant to maritime operations, and the structure and content of a representative dataset on maritime vessels (e.g., AIS). This approach, out of multiple perspectives, ensures that both theoretical coverage and practical applicability are adequately reflected in the ontology design. To facilitate analysis and validation, the CQs were assigned to two categories. The first category comprises general CQs, which aim to delineate the overall conceptual scope and boundaries of the ontology. These questions help ensure that the model provides a coherent and comprehensive semantic framework for representing vessel-related knowledge at a high level of abstraction. The second category includes use case-specific CQs, which address concrete analytical challenges, such as predicting ship trajectories, determining ownership distribution, identifying propulsion characteristics, or integrating multimodal data sources for vessel representation.

To systematically evaluate the final ontology with the defined CQs, we aligned each CQ to the associated classes and properties it is modeled by. This way, it can be ensured that each CQ is represented in the ontology. Furthermore, it provides a structured approach to support design decisions and allows the integration of future extensions. Organized into general and use case-specific questions, Table 1 presents an excerpt of eight CQs, further divided according to the module to which they belong.

Table 1
CQs are opposed to the entities they are answered by. We categorized the CQs according to the subject they address: identification (C01), trajectory (C02), propulsion (C03), and multimodal representation (C04). They are divided into general (g) and use case-specific (u) questions. We only list entities that we added to the VesselAI ontology as part of AISHIP.

Subject	CQ	Question	Entities
Identification	C01.1g	How are ships clearly identified?	(Vessel - hasIMO - integer)
	C01.3u	Which organization owns the most ships?	(Vessel - hasOwner - Organization123)
Trajectory	C02.2g	What information is known about a single trajectory part?	(TrajectoryPart - hasDraught - QuantityValue)
	C02.1u	What unit does the draught of a ship have?	(TrajectoryPart - hasDraught - QuantityValue) (QuantityValue - unit - Meters)
Propulsion	C03.4g	What types of engines are used in ships?	(PistonEngine - subClassOf - EngineSpecification) (NonPistonEngine - subClassOf - EngineSpecification) (CombinedEngine - subClassOf - EngineSpecification)
	C03.2u	What is the most common type of propeller?	(ScrewPropeller - subClassOf - BladedPropulsor) (FixedPitchPropeller - subClassOf - ScrewPropeller)
Multimodality	C04.2g	How is sonar data of ships stored?	(Vessel - hasSonarSignature - Audio)
	C04.1u	What does ship X look like?	(ShipX - hasPhoto - ImageShipX)

After carefully mapping all CQs to the model, we can confidently confirm that the model achieves a domain coverage of 100%. It demonstrates that the model is conceptually complete with respect to its intended scope and capable of supporting all identified use cases.

5.2. Ontology Validation

To fully assess the ontology's applicability and adequacy in real-world scenarios, it is essential to test its behavior with representative data. The validation process of SABIOx involves creating a dataset of individuals that covers the ontology and represents realistic entities. CQs are translated into SPARQL queries and executed against this dataset. If the query results align with the expected answers, the ontology can be considered to fulfill its intended purpose.

Currently, such a dataset has not yet been created, but we outline the approach for its development. We plan to adapt and extend the dataset published by the VesselAI authors [23]. This adaptation will require datatype adjustments (as we restricted them) and a strict alignment with the modeling approach based on QUDT.

In addition, SHACL shapes will be defined based on the ontology's axioms to validate whether the dataset complies with the defined restrictions. For example, a shape will ensure that a vessel using rotor sails for wind-assisted propulsion must also include an engine as the primary MPS. To complement synthetic data, we plan to consult the maritime experts to incorporate real-world domain data into the dataset. By following this strategy, we aim to ensure the ontology's robustness and achieve a recall of 1.0 during query-based validation.

6. Conclusion and Future Work

In this paper, we introduced AISHIP, an extension of the VesselAI ontology designed to enable a more comprehensive and semantically consistent representation of vessels. Our work addresses key limitations of VesselAI by incorporating additional vessel characteristics, enhanced trajectory modeling, multimodal vessel descriptions, and a detailed conceptualization of PSs. These extensions close the gap in semantic alignment between heterogeneous maritime data sources and resolve inaccuracies related to the representation of numerical measurements and their units. By integrating recognized standards such as QUDT, GeoSPARQL, and OWL-Time, AISHIP adheres to FAIR principles, ensuring interoperability, reusability, and semantic precision across maritime information systems.

Future work will focus on creating an extended dataset to validate the ontology through SPARQL-based CQ testing and SHACL shape evaluation. Furthermore, we plan to assess the ontology in practical scenarios, including trajectory prediction and SAR operations, to demonstrate its applicability in real-world contexts. Another major direction involves linking AISHIP with SSN/SOSA to incorporate sensor data, enabling advanced use cases such as estimating fuel consumption along specific routes. These additions will broaden the applicability for maritime analytics, decision support, and operational optimization use-cases.

Declaration on Generative AI

During the preparation of this work, the author used ChatGPT and Grammarly in order to: Grammar and spelling check, translate, paraphrase, and reword. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the publication's content.

Resource Availability Statement

The AISHIP ontology[5], the competency questions[6] are available on GitHub [24] and Zenodo (DOI: 10.5281/zenodo.16704863). All resources are licensed under Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International.

References

[1] United Nations Conference on Trade and Development (UNCTAD), Review of Maritime Transport 2024: Navigating Maritime Chokepoints, 2024. URL: https://unctad.org/publication/review-maritime-transport-2024, accessed: 01.08.2025.

[2] Z. Wang, C. Claramunt, Y. Wang, Extracting Global Shipping Networks from Massive Historical Automatic Identification System Sensor Data: A Bottom-Up Approach, Sensors 19 (2019) 3363. URL: http://dx.doi.org/10.3390/s19153363. doi:10.3390/s19153363.

[3] Y. Yang, Y. Liu, G. Li, Z. Zhang, Y. Liu, Harnessing the power of Machine learning for AIS Data-Driven maritime Research: A comprehensive review, Transportation Research Part E: Logistics and Transportation Review 183 (2024) 103426. URL: http://dx.doi.org/10.1016/j.tre.2024.103426. doi:10.1016/j.tre.2024.103426.

[4] International Maritime Organization (IMO), AIS Transponders, n.d. URL: https://www.imo.org/en/OurWork/Safety/Pages/AIS.aspx, accessed: 01.08.2025.

[5] G. M. Santipantakis, C. Doulkeridis, G. A. Vouros, An Ontology for Representing and Querying Semantic Trajectories in the Maritime Domain, Springer Nature Switzerland, 2023, p. 224–237. URL: http://dx.doi.org/10.1007/978-3-031-42914-9_16. doi:10.1007/978-3-031-42914-9_16.

[6] C. Claramunt, C. Ray, E. Camossi, A.-L. Jousselme, M. Hadzagic, A. Gennady, N. Andrienko, Y. Theodoridis, G. Vouros, L. Salmon, Maritime data integration and analysis: recent progress

[5] https://burbachs.github.io/AISHIPOntology/AISHIP.owl
[6] https://github.com/BurbachS/AISHIPOntology/blob/main/CQs/CompetencyQuestions.txt

and research challenges, 2017. URL: https://openproceedings.org/2017/conf/edbt/paper-224.pdf. doi:10.5441/002/EDBT.2017.18.
[7] M. Mirović, M. Miličević, I. Obradović, Big Data in the Maritime Industry, Naše more 65 (2018) 56–62. URL: http://dx.doi.org/10.17818/NM/2018/1.8. doi:10.17818/nm/2018/1.8.
[8] S. Brüggemann, K. Bereta, G. Xiao, M. Koubarakis, Ontology-Based Data Access for Maritime Security, Springer International Publishing, 2016, p. 741–757. URL: http://dx.doi.org/10.1007/978-3-319-34129-3_45. doi:10.1007/978-3-319-34129-3_45.
[9] C. Ray, Data Variety and Integrity Assessment for Maritime Anomaly Detection, in: M.-R. Hojeij, B. Finance, Y. Taher, K. Zeitouni, R. Haque, M. Dbouk (Eds.), Proceedings of the 1st International Conference on Big Data and Cyber-Security Intelligence (BDCSIntell 2018), volume 2343 of *CEUR Workshop Proceedings*, CEUR-WS.org, Hadath, Lebanon, 2018, pp. 4–7. URL: https://ceur-ws.org/Vol-2343/paper3.pdf, accessed: 01.08.2025.
[10] European Commission, ENABLING MARITIME DIGITALIZATION BY EXTREME-SCALE ANALYTICS, AI AND DIGITAL TWINS, CORDIS Project Fact Sheet, 2023. URL: https://doi.org/10.3030/957237, Grant agreement ID 957237, coordinated by ETHNICON METSOVION POLYTECHNION, Greece.
[11] R. Battle, D. Kolas, GeoSPARQL: Enabling a Geospatial Semantic Web, Knowledge Engineering Group, Raytheon BBN Technologies (2011) 1–17. URL: http://parliament.semwebcentral.org, accessed: 25.07.2025.
[12] European Maritime Safety Agency (EMSA), Common Information Sharing Environment (CISE), 2025. URL: https://www.emsa.europa.eu/cise.html, accessed: 27.07.2025.
[13] A. Haller, K. Janowicz, S. Cox, D. L. Phuoc, K. Taylor, M. Lefrançois, Semantic Sensor Network Ontology, https://www.w3.org/TR/vocab-ssn/, 2017. W3C Recommendation, World Wide Web Consortium (W3C), accessed: 02.05.2025.
[14] K. Janowicz, A. Haller, S. J. Cox, D. Le Phuoc, M. Lefrançois, SOSA: A lightweight ontology for sensors, observations, samples, and actuators, Journal of Web Semantics 56 (2019) 1–10. URL: http://dx.doi.org/10.1016/j.websem.2018.06.003. doi:10.1016/j.websem.2018.06.003.
[15] G. Santipantakis, K. I. Kotis, G. A. Vouros, Ontology-Based Data Integration for Event Recognition in the Maritime Domain, in: Proceedings of the 5th International Conference on Web Intelligence, Mining and Semantics, WIMS '15, ACM, 2015, p. 1–11. URL: http://dx.doi.org/10.1145/2797115.2797133. doi:10.1145/2797115.2797133.
[16] C. Z. Aguiar, V. E. S. Souza, SABiOx: the extended systematic approach for building ontologies, in: Proceedings of the 17th Seminar on Ontology Research in Brazil (ONTOBRAS 2024) and 8th Doctoral and Masters Consortium on Ontologies (WTDO 2024), CEUR Workshop Proceedings, Vitória, Brazil, 2024. URL: https://nemo.inf.ufes.br/projetos/sabiox, Available under Creative Commons License Attribution 4.0 International (CC BY 4.0).
[17] A. Gangemi, V. Presutti, Ontology Design Patterns, Springer Berlin Heidelberg, 2009, p. 221–243. URL: http://dx.doi.org/10.1007/978-3-540-92673-3_10. doi:10.1007/978-3-540-92673-3_10.
[18] E. Blomqvist, A. Gangemi, V. Presutti, Experiments on pattern-based ontology design, in: Proceedings of the fifth international conference on Knowledge capture, K-CAP '09, ACM, 2009, p. 41–48. URL: http://dx.doi.org/10.1145/1597735.1597743. doi:10.1145/1597735.1597743.
[19] International Maritime Organization (IMO), Reduction of GHG Emissions from Ships: Wind Propulsion, Technical Report MEPC 79/INF.21, Marine Environment Protection Committee, 79th session, 2022. URL: https://www.wind-ship.org/archived-site/wp-content/uploads/2022/10/MEPC-79-INF.21-Wind-Propulsion-Finland-France-Saudi-Ar....pdf, accessed: 01.08.2025.
[20] N. Kornev, Propellertheorie, Universität Rostock, Fakultät für Maschinenbau und Schiffstechnik, Rostock, Germany, 2009. URL: https://www.yumpu.com/de/document/read/8864664/prof-dr-ing-nikolai-kornev-lehrstuhl-fur-modellierung-und-, accessed: 01.08.2025.
[21] S. Jörke, Konzeptentwicklung für den Entwurf von Schiffspropellern, Diplomarbeit D 2057, Technische Universität Dresden, Fakultät Maschinenwesen, Institut für Strömungsmechanik, Professur für Turbomaschinen und Flugantriebe, Dresden, Germany, 2017. URL: https://cfturbo.com/fileadmin/content/down/publications/students/2017-07-Soeren-Konzept-Entwurf-Schiffspropeller.pdf, ac-

cessed: 01.08.2025.

[22] I. Krmek, V. Mrzljak, I. Poljak, Analysis and Comparison of Ship Propulsion Systems, Pomorski zbornik 62 (2022) 75–95. URL: https://hrcak.srce.hr/file/408027, Review article, received: 9 May 2022, accessed: 01.08.2025.

[23] S. Giorgos, D. Christos, V. George, Linked open data for the maritime domain, 2022. URL: https://zenodo.org/record/6608427. doi:10.5281/ZENODO.6608427.

[24] S. Burbach, L. Mackert, M. Maleshkova, BurbachS/AISHIPOntology: AISHIP v1.0.0: Pre-publication release, 2025. URL: https://github.com/BurbachS/AISHIPOntology. doi:10.5281/zenodo.16704863.