Open Access Original Article

Corrected: Documenting FAIR and Dublin Core for Maritime Legal Evaluation

Enrique Wulff [1]

1. The Library, Institución Milá y Fontanals de investigación en Humanidades (IMF-CSIC), Barcelona, ESP

Corresponding author: Enrique Wulff, enrique.wulff@imf.csic.es

Received 06/25/2025
Review began 07/30/2025
Review ended 09/08/2025
Published 09/12/2025
Corrected 12/04/2025

DOI:
https://doi.org/10.7759/s44389-025-08177-y

This article has been corrected.

Correction date: December 04, 2025. Cite this correction as Wulff E (December 04, 2025) Correction: Documenting FAIR and Dublin Core for Maritime Legal Evaluation . Cureus J Comput Sci 2 : es44389-025-00010-w. doi:https://dx.doi.org/10.7759/s44389-025-00010-w.

The original version of this article contained inaccuracies and omissions that have now been corrected. The following changes were made:

Terminology Adjustments

1. ‘FAIR standards’ replaced with ‘FAIR principles’ throughout the text.

2. In Table 1, ‘discoverable’ changed to ‘findable.’

3. In Table 3, ‘EDMET’ corrected to ‘EDMED’ and ‘using metadata’ changed to ‘usage metadata.’

4. ‘Linking Open Data’ updated to ‘Linked Open Data (LOD).’

Content Enhancements

1. Additional references incorporated, increasing the total from 12 to 21.

2. Expanded introduction with clarifications on EOSC objectives and metadata standards (schema.org, DataCite).

3. Acronyms fully spelled out at first mention (LRIT, OGC, IoT, FAO, MSP, ODIS, GeoJSON, LegalRuleML, STAC, OGC CSW, REST/OpenAPI, WMS).

4. Eight new explanatory phrases added in Materials & Methods section.

5. Inserted new phrase after UNCLOS mention and EOSC-related sentence in Results & Discussion.

6. Two concluding sentences added to emphasize ontology-based reasoning and EOSC integration.

Abstract

The EU's primary legal frameworks enable coordinated access to spatial and in-situ data about ocean activities, promoting the discovery and use of scientific data and metadata, including documentation on satellite and in-situ observation systems. This analysis explores the application of FAIR principles to the software codes, data, and tools used by the Space Situational Awareness (SSA) program, focusing on the legal evaluation of maritime spatial uses of Dublin Core. It suggests actions to support the European Open Science Cloud (EOSC) to engage data citation from an ontology-based framework.

Categories: AI/ML-based decision support systems, Cloud Management and Monitoring, Data Structures
Keywords: legal-data ontologies, maritime court cases, eosc (european open science cloud), dublin core, fair principles

Introduction

The first introduction of FAIR principles to the software codes, data, and interpretation tools used by ocean data focused on fisheries and meteorology, and helped to enhance their quality control practices to promote data reuse [1]. Wilkinson et al. (2016) estimates that the principles were designed to be evolved into rules and standards categorized into specific disciplines [2]. This initiative aligned with the Enabling FAIR Data Project, which is part of the Coalition on Publishing Data in Earth and Space Sciences

(COPDESS) and includes the European Geosciences Union (EGU) in its collaboration with the International Union of Pure and Applied Chemistry (IUPAC) [3]. A clear role for data management librarians is established based on this fact [4]. This proposal engaged a multi-disciplinary group ofArtificial Intelligence(AI) and Machine Learning (ML) stakeholders [5]. When the guiding principle is free and open access to ocean data content for the research and educational communities in non-commercial activities, the Dublin Core standard is one of the formats that ensures easy and widespread use and encourages citation of the dataset. All metadata fields required for citation are part of Dublin Core, as well as alternative approaches like schema.org and DataCite. These three metadata standards are, by their very nature generic; however, Dublin Core is preferable because reference managers use it to extract the persistent identifier or full citation metadata from landing pages, facilitating data discovery in data repositories. Dublin Core is the preferred standard for exchanging metadata via OAI-PMH, the web-based protocol that serves metadata in XML format [6].

The provisions concerning scientific data that support marine research include the Space Surveillance and Tracking support framework, which was launched in 2014; this program incorporates an specific computational ontology [7-8]. SpaceSurveillanceandTracking data can be linked to the Long-Range Identification and Tracking (LRIT) of ships. The organization and accessibility of data using best practices requires significant efforts in high-level access and preservation within digital libraries. This highlights that marine knowledge activities extend beyond just research. As such, there is a pressing need to establish standards for both data and metadata in the management of maritime spatial information, which is critical. Taking into account the European Open Science Cloud (EOSC) goalsto address usage beyond openly available FAIR data and engage a wider user base includes the public and the private maritime sector. However, organizations in the sector have not yet mastered the development of their own AI algorithms and techniques to connect hardware and software in the cloud. This need arises from the increasing availability and reuse of spatial data by European public entities, as well as the legal evaluation of spatial data infrastructures (SDI). In terms of their size, according to the 2022 report from the European Maritime Safety Agency (EMSA), approximately 25,000-30,000 vessels (including cargo, tankers, and passenger ships) navigate European waters daily, exchanging positional data to avoid collisions. This results in about 250-300 million Automatic Identification System (AIS) positional messages each month, received at some 1,200 AIS receiving coastal and satellite-based stations across Europe [9].

Our work introduces the FAIR Principles framework (Table *1*) that leverages prompts derived from data description and retrieval, data authentication and data service authorization, and the adoption of community-accepted standards and open licenses.

To be findable: - (meta)data are assigned a globally unique and eternally persistent identifier - (meta)data are described with rich metadata - (meta)data are recorded or indexed in a searchable resource - (meta)data specify the data identifier
To be accessible: - (meta)data are retrievable by means of their identifiers, using a standardized communication protocol - (meta)data are open, free, and universally implementable - (meta)data allow for authentication and authorization procedures when necessary - (meta)data are accessible even when the data are no longer available
To be interoperable: - (meta)data use a formal, accessible, interchangeable, and widely applicable language to represent knowledge - (meta)data use vocabularies that follow the FAIR principles - (meta)data include references Qualified to other (meta)data
To be reusable: - (meta)data have a variety of accurate and relevant attributes - (meta)data are published with a clear and accessible data usage license - (meta)data are associated with their provenance - (meta)data meet important standards within a community

TABLE 1: FAIR facets (Source: FORCE11)

Referred from [10]

As maritime data is used to monitor and protect EU naval space, the development of strong institutional arrangement for harmonization and coordination of national data-sharing platforms with open innovation and features for cross-border e-commerce is a very valuable proposal. Furthermore, extensive use of official data sources ensures that real-world maritime data is used in the spirit of the European maritime community's data-sharing efforts. For instance, EOSC is working to support real-time ingestion of maritime data into EU Digital Twins of the Ocean (DTOs). Community data with official data sources is being used to assess the FAIRness of citizen science data in the context of the Green Deal Data Space. Additionally, the high cost of collecting data across the Economic Exclusive Zones is primarily based on nationalofficialdatasources. It is proposed in this contribution to bolster reasoning capabilities over ontologies to serve as the legal knowledge basis that enhances the interpretability and credibility of the judgment prediction concerning the maritime semantic space. Some observers doubt that making decisions based on ontologies regarding whether the vessel is ariskor behaving normally is without risks. Nevertheless, a knowledge-based representation of the border surveillance domain establishes potential visual associations that help infer complex scenarios. Legal authorities are already implementing these semantic descriptions as they are technology independent, support portability and make a first step towards knowledge-driven vessel activity recognition. That is a way to extend the existing tools for maritime surveillance across the maritime regions of Europe. The multiple sources available for such an interpretable legal reasoning framework are provided in Fig. *1*.

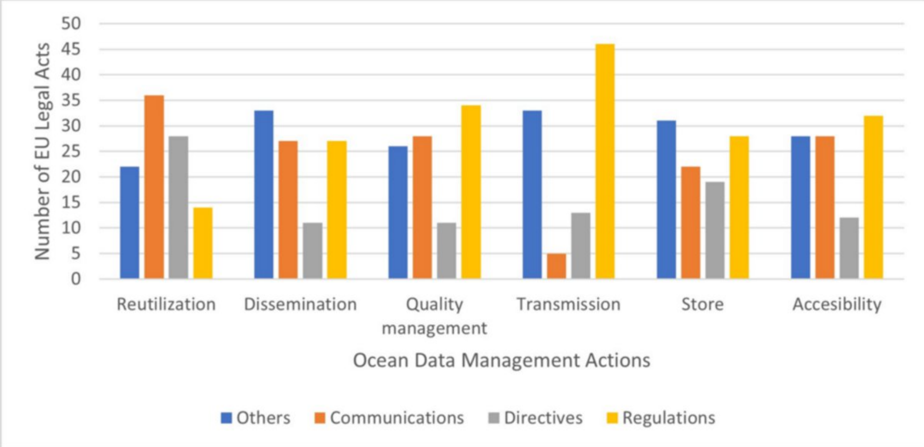

FIGURE 1: Legal basis at European level to support oceanographic data management actions

Some crucial recent literature on ontology architecture in the maritime domain presents the methodology in automatic reasoning mechanism to deal with uncertainty in quasi-empirical models. Aussenac-Gilles et al. (2022) search to provide semantic interoperability, ontology-based data access, andreasoningcapabilities to the different systems relying on the Common Information Sharing Scheme for the Maritime Domain (CISE) [11]. Yahya et al. (2023) assert that using ontologies for reasoning enables us to determine logical entailment and reuse object properties from the Dublin Core ontology in order to follow the best practices of Linked Open Data (LOD) [12]; furthermore, they evaluate their ontology for FAIR principles, using OOPs! Tool [13]. The Ontology Pitfall Scanner (OOPS!) evaluates ontologies during their creation by identifying design imperfections. For example, to strengthen the results of the Fish Ontology (FO) evaluation, the online ontology evaluation tool OOPs! has been used [14-15]. As is well-known publishing rich ontology metadata under the w3id.org domain (w3id innovatively generates secure digital object identifiers (DOIs) suitable for the decentralized Web 3.0 ecosystem) [16], ensures compliance with the FAIR principles [17]. Camossi & De Rosa (2020) advance that context reasoning has a specific focus in the real world maritime security use case and that because the semantics of sensors leverages the specification of maritime data and events [18]. Building on this knowledge, our approach takes into account the critical role of legal knowledge in making judgments, and searches to improve the classification of maritime data that are either easily confused or rare.

Our main contribution is a set of indicators to make explicit the legal basis of judgments concerning maritime data. This framework offers tailored explanations for

a. Architecture, norms, and standards for network services.

b. Catalogs and metadata.

c. Norms and standards that ensure the interoperability of datasets and services (discovery, visualization, download, transformation, invocation).

We first explain the mainstream methods and existing issues in the current analysis of European policies concerning ocean data, and then formally introduce the proposed framework.

Materials And Methods

The focus of this study is on maritime data as referred to in European legislation. The main EU legal tools concerning the discovery, access, and use of maritime data understand the term 'maritime' in a cross-sectoral sense: any data linked to things associated with the sea. The evaluation data set was gathered by searching the EUR-Lex database [19]. A first inventory identified 439 documents, of which 228 were discarded because they did not correspond to the data authentication and data services authorization aspects of the FAIR principles (see Table *1*). MS Excel was used to assist with analysis and classification, creating a workbook divided into 12 fields to describe the current EU legal status on ocean data. This file secured a tool to ensure the interpretability of judgments.

A content analysis of the legal texts has been carried out based on a set of 12 descriptive and content indicators. To collect information, the database was divided into 12 fields, one for each of the established indicators (title, document type, organization, description, date, location, URL, geographical scope, data type, citations in legislation, comments and actions (oceanographic data management)) (Table *2*). It includes presenting the current status of data management portals, data policies, open access, licensing, typology, formats, Open Geospatial Consortium (OGC) web services, metadata, data services, interoperability, data management, and compliance with legislation.

Field	Description	Possible values	Comments
Title	Title of the legal document		Name of the regulation
Document Type	Type of EU legal source identified	Regulation, directive, communication, decision, recommendation, international agreement, convention, program, resolution, declaration, code, law, supplementary protocol, order, strategic plan, reports, background papers, proposed regulation, statistical information	
Promoting Organization	Institutional responsibility for the development of legislation	European Parliament, Council of the European Union, European Commission, Food and Agricultural Organization of the United Nations (FAO), etc.	Organizations that legislate about European seas and oceans
Description	Brief summary of the content		
Date	Original date of approval of the legislation		
City	City where the legislation was originally signed		
URL	Electronic access to the legislation		Internet page address
Geographic Scope	Sea, ocean, or territorial area whose legal status is discussed	Europe, coastal states of the Baltic Sea, Black Sea, Northeastern Atlantic, Mediterranean Sea, and others	
Data Type	Mention of the “type of data” in the analyzed legislation	Fisheries, marine, maritime, water, spatial, other data	The indicator observes and classifies legislation and data policy as relevant to improving comparability and accessibility of data
Citations in Legislation	References to oceanographic data management actions	Legal text sections referring to documentary action	Transcribes the content of the published legislation
Comments	Commentary on the objectives of the legal text		
Actions	Data management actions for oceanographic data	Access, storage, transmission, quality management, dissemination, reuse	

TABLE 2: Indicators for analyzing EU ocean data legislation (Source: the author)

Consistent with a reading of the legislation that takes into account the context in which the legal guarantees were issued to predict the structure and reasoning of court rulings concerning European maritime data, Table 3 formalizes the task of classifying the metadata used in this domain.

Metadata	Description
Dublin Core	A basic, domain-agnostic standard that is easy to understand and implement; it is one of the best-known and most widely used metadata standards. It was published as ISO 15836 in February 2009.
EDIOS	Defined metadata format, based on the ISO 19115 content model, for the SeaDataNet EDIOS (ocean observing systems) directory.
EDMED	Defined metadata format, based on the ISO 19115 content model, for the SeaDataNet EDMED (marine environmental data) directory.
EDMERP	Defined metadata format, based on the ISO 19115 content model, for the SeaDataNet EDMERP directory (marine environmental research projects).
EML	Ecological Metadata Language is a specification developed for the discipline of ecology.
RDF	A general method for conceptual description or information modeling that is implemented in web resources, using various syntax formats.
SensorML	OGC standard as a sensor markup language, which provides a rich collection of metadata that can be mined and used to detect sensor systems and observation processes.
DIF	Directory Interchange Format, a descriptive and standardized format for exchanging information about scientific data sets.
CDI	Common Data Index, a metadata format that is a marine profile of the ISO 19115 standard for metadata content.
CF	Climate and Forecast Metadata Conventions is a standard for "usage metadata" that seeks to both distinguish quantities (such as physical description, units, and preprocessing) and locate data in space and time; sponsored by netCDF Climate and Forecast Metadata Convention.
FGDC-CSDGM	Federal Geographic Data Committee Content Standard for Digital Geospatial Metadata, a widely used standard for defining the content of a digital geospatial dataset; in September 2010, the Federal Geographic Data Committee adopted ISO 19115.
GCMD	The Global Change Master Directory uses DIF metadata to promote the discovery, access, and use of Earth science data and data-related global services, with a particular focus on NASA data.
ISO 11179	A standard that describes the metadata and activities necessary to manage data elements in a registry and create a common understanding of the data across organizational elements and between organizations.
ISO 14224	A standard for the collection and exchange of reliability and maintenance data for equipment in the petroleum, petrochemical, and natural gas industries.
ISO 17369	It facilitates the exchange of statistical data and metadata (SDMX), allowing for interoperable implementation within and between systems that deal with them.
ISO 19101	Defines the reference model for standardization in the field of geographic information.
ISO 19115	Defines the schema required to describe geographic information and services through metadata.
ISO 19119	A geographic information service standard that allows it to be specified independently of one or more underlying distributed computing platforms.
ISO 19139	Defines geographic metadata for images and grid-encoded data.
O&M (ISO 19156)	It provides models for information exchange that describe observational activities and their results, both within and between different scientific and technical communities.

TABLE 3: Metadata used in the 116 European oceanographic centers analyzed (Source: RDA | Metadata Directory)

For the critical improvement of these methods and to verify the performance of the proposed framework we have approached this database with the DeepSeek Large Language Model (LLM). Large language model-based ontology construction for resolving maritime traffic conflicts involves to access and filter structured data such as ontologies and databases. One of the key steps is to automaticaly or semi-automatically extract knowledge units such as entities, relationships and attributes from text. Examples of extractions can be seen in Table *3*. Knowledge fusion into an unified framework to support a well-defined hierarchy is the main effect awaited. Two stages of reasoning have been employed with the DeepSeek machine-learning model. The production of maritime indicators based on prompts for better legal judgments' predictability. The verdicts prediction prompts serve as the final component. An additional reasoning stage search to provide context for a maritime ontology using Dublin Core. DeepSeek has been used as it is exceptionally accurate in question-asking tasks. What makes DeepSeek special is that achieves the same or better results than other players in the AI industry (eg, OpenAI's ChatGPT) with only a fraction of the training costs [20].

Running it locally Table *3* has been processed to extract where each piece of legal knowledge contains the concept and characteristics of a certain metadata. Then, we analyze the topic distribution of each metadata with a learning strategy that encodes the set of legal knowledge obtained from the database. LLM is employed to obtain the reduced amount of information where key perspectives are likely to emphasize the optimization objective of the training: to provide information that help distinguish cases to predict maritime veredicts at the court. For this purpose, we designed the following instruction (prompt) for LLM reasoning: "Define 'maritime situational indicators' by discerning the views in mobility data following the perspective from the maritime domain as expressed in the uploaded Table *3*". After that, an inference from the original metadata description have been obtained. The generated results may have an impact on the legal judgment as their textual description is clarified with explicit and structured elements (Table *4*).

In a second stage of reasoning, we aim to produce context for an ontology intended to enhance citation rate to the data utilized. Specifically, based on the texts obtained in the first hop, we designed prompt templates for two different subtasks, respectively: explore specific examples (e.g., piracy detection, emissions monitoring) to produce an ontology intended to enhance citation rates to the maritime mobility data (e.g., AIS, Internet of Things (IoT) systems); it should be highlight that the metadata is in Dublin Core format.

By linking these two separate bots to work together, we obtain improved text input for the corresponding case samples. This provides the model with clearer information and more appropriate task content.

Specifically, in the task studied in this contribution we introduced knowledge from multiple sources, including European maritime ontologies, LLM prior knowledge, and task-oriented knowledge to improve court verdicts prediction handling cases were maritime metadata had been employed. In this process, we hope to provide example applications were Dublin Core had been prioritized. The simplicity and broad applicability of Dublin Core make it valuable in situations where standardized, interoperable metadata is required. In terms of predicting court decisions regarding maritime metadata, we present three examples from court cases where the Dublin Core has played a key role. This is for three reasons: First, data disclosure in legal disputes. When data is held by multiple institutions, the simplicity of the Dublin Core allows metadata to be aggregated from different sources, which can speed up evidence gathering; second, interoperability. Courts may need to access data from different systems. The standardization of the Dublin Core ensures interoperability, making it easier to present data as evidence; third, provenance tracking. In cases where the authenticity of data is in doubt, Dublin Core elements such as creator, date, and publisher can establish the origin and history of the data.

As part of the European Commission's Maritime Spatial Planning (MSP) roadmap for 2022-2027, metadata standards for MSP has been advanced by the Technical Expert Group (TEG) on Data [21]. The group's objectives are to make maritime spatial plans comply with good data management and the FAIR Data Principles. In conformity with these criteria we consider three example applications:

- Marine Pollution Dispute. Many organisations (research institutes, NGOs, government agencies) provide environmental data as evidence. Dublin Core ensures that metadata fields (e.g. name, creator, date, subject) are consistent across datasets. This allows courts to quickly verify the provenance and relevance of data (e.g. by linking a pollution report to a specific vessel using creator and date). Integration with ontologies such as the Marine Data and Information Network (MEDIN) is also facilitated by linking the Dublin Core "subject" field to legal terms such as "dangerous spill". One example is a 2018 European case where Dublin Core metadata from 15 agencies helped establish a timeline of the effects of an oil spill, which directly influenced the court's decision on liability.

- Cargo Damage Claim. A shipping company disputes cargo damage claims. Gauge logs, maintenance logs and contracts are marked in accordance with Dublin Core to standardize descriptions. Important fields such as description (e.g. "pre-shipment inspection report") and format (e.g. PDF/A to meet archival requirements) ensure that metadata meets acceptability standards. LLMs trained in Dublin Core-marked arbitrations identify trends (e.g. cases without date metadata are 30% more likely to be rejected). An example is the 2021 Rotterdam arbitration, which favoured maintenance logs in Dublin Core format, resulting in a quick decision in favour of the insurer due to clear retention periods for the metadata.

- Maritime Boundary Conflict. Historical records of fishing and shipping are being digitized using Dublin Core to resolve sovereignty claims. Fields such as Coverage (geographic coordinates) and Rights (permissions to use) comply with the European INSPIRE directive on geospatial data management. Courts use the Dublin Core Language field to filter evidence (for example, excluding documents not written in English unless they are translated). Take a case brought before the International Court of Justice in 2020:

Dublin Core archives provided unambiguous historical evidence of fishing rights, supporting a landmark territorial demarcation decision.

Results And Discussion

In this section, we present the experimental setup components, describe what we are trying to test, the datasets used, the software employed, and the procedure outlined to run the experiment. Following this, we present and analyze the experimental results.

Experimental setup

What we are trying to illustrate is the judgment prediction capability of ontologies concerning the maritime policy in Europe. We used an integrated dataset extracted from EUR-Lex, the official portal for EU legislation, to assess the relevance of our hypothesis. This research compiles 211 European legal texts on oceanographic data management, selected from official documents of the Commission, the Council, and the Parliament, as well as legal studies on data discovery, access, and policies in the European Union. EUR-Lex includes legal texts from EU institutions published in the Official Journal and in reports from the Court of Justice of the European Union in all official EU languages. EUR-Lex was consulted using its advanced search module. This search was complemented by a keyword formulation in Google with the search equation used in “EUR-Lex” to obtain detailed pertinent information for the constructed dataset. The original dataset was cleaned by filtering out the European Commission Staff Working Documents and Impact Assessment reports, given that these are assessment documents addressing specific issues and are not considered legal in nature. The cleaned EUR-Lex dataset version was used for analysis.

The experimental outline utilizes two stages of reasoning. Elements for supporting maritime indicators are formed by designing prompts to generate better legal judgments' predictability. Context for an ontology is then sought, focusing on provenance and citation tracking starting from the 15 core elements that integrate the Dublin Core metadata structure.

Result

In true judgment prediction, judges make the final decision based on the key facts of the case and relevant laws. Inspired by this process, we propose the following framework which explores the application of FAIR principles and Dublin Core format to the codes, data, metadata and tools used by the Space Situational Awareness (SSA) program, focusing on the legal evaluation of maritime spatial with multi-knowledge sources prompted.

Key points are data sources (mobility data), purposes (detecting illegal activities, monitoring environmental risks, improving security), analytical structures (transforming data into insights such as density maps, anomaly detection), and technology integration (big data, real-time processing). The goal is to map these elements into a legal judgment framework. So that it shows the flow from data collection to legal decisions, incorporating the analytical steps and how they inform judgments. The visual operational flow of this framework is shown in Figure 2.

[Data Acquisition] → [Analytical Processing] → [Legal Criteria Mapping] → [Evidence Synthesis] → [Judicial Decision]

FIGURE 2: Legal judgment reasoning framework for maritime spatial governance

The Figure 2 shows the flow from data collection to legal decisions, incorporating the analytical steps and how they inform judgments. For example, mobility data processed using machine learning leads to anomaly detection, which is related to legal standards such as United Nations Convention on the Law of the Sea (UNCLOS) articles. For intelligent situational awareness using AIS data marine anomaly detection is mainly connected to unsupervised learning techniques.

Consistent use of FAIR data would transform maritime situational indicators fromdescriptive(what happened) toprescriptive(what to do next). The original indicators remain valid but are augmented by AI’s predictive power, IoT’s real-time granularity, and fused data’s holistic context (Table 4). Stakeholders must balance technological gains with governance frameworks to avoid problems such as misuse or over-automation.

Original indicators	With FAIR data
(e.g., vessel density maps, anomaly detection) relied on traditional AIS and historical data.	**Enhanced Accuracy and Granularity:** IoT sensors (e.g., buoys, drones, satellite networks) and fused data streams (e.g., weather, cargo manifests, port logs) reduce gaps in coverage.
Historical or delayed insights (e.g., post-incident analysis of oil spills).	**Real-Time Responsiveness:** Real-time IoT feeds (e.g., oil spill sensors, engine-status telemetry) enable proactive indicators (e.g., predicting mechanical failures or environmental hazards before they escalate).
Anomalies flagged by AIS data alone (e.g., a vessel stopping) could be misinterpreted (e.g., fishing vs. piracy).	**Reduction of False Positives:** AI integration contextualizes behavior using fused data (e.g., correlating stops with fishing licenses, weather conditions, or port activity), refining indicators to focus on true risks.

TABLE 4: Maritime situational indicators (Source: Author)

AIS, Automatic Identification System

In practice to ensure that this framework makes maritime legal evaluations technically robust, legally defensible, and interoperable across jurisdictions and stakeholders, data catalogs and metadata should be available from a centralized repository by using Dublin Core and ISO 19115 formats. Nevertheless, by following the schema.org approach and the principles of Linked Data (LD), no highly centralised system should be needed. The IOC Ocean Data and Information System (ODIS) (Ocean Data Information System) is such an independent, voluntary data-sharing platform. Otherwise solving interoperability issues should be focused on datasets codified in Geographic JavaScript Object Notation (GeoJSON), Legal Rule Markup Language (LegalRuleML), Spatio Temporal Asset Catalog (STAC) and purposely Open Geospatial Consortium Catalog Service for the Web (OGC CSW), Representational State Transfer/Open Application Programming Interface (REST/OpenAPI), and Web Map Services (WMS) should be offered. This requires a scalable, secure, and maritime and cybersecurity compliant architecture; the use of FAIR compliant metadata to provide accurate discovery catalogues; and standards-based pipelines for cross-system data/service integration.

Experimental analysis

We illustrate the interpretability evaluation of our proposed framework by training a dataset to improve distinguishing smuggling from legitimate deviations. A varied collection of EU legal texts was used to train diverse iterations of DeepSeek, the LLM for the chain prompt reasoning module. Each AI model was assessed on both the design of prompts for legal reasoning and the exploration of how to build an ontology using provenance and citation tracking.

Combining these AI reasoning techniques with metadata structuring would improve both the quality of legal predictions and the traceability of their informational sources. The high accuracy of space situational awareness makes it ideal for treating maritime zones like "spatial domains", incentivizing academic recognition for data generators and making maritime spatial indicators traceable and citable in scientific and legal contexts. This study supports the process of judgment prediction by using AI models, particularly LLM, trained on meticulously annotated legal datasets. The experimental analysis demonstrates how anchoring legal and maritime datasets in Dublin Core metadata and FAIR principles enables structured, interoperable, and citable data ecosystems, in particular within European initiatives like the EOSC. Interaction with EOSC's services make FAIR digital objects an emerging concept. Their evaluation within the maritime domain presents a growing interest for alignment strategies such as semantic mapping or external ontology queries.

Conclusions

To improve the accuracy, transparency, and fairness of court verdict predictions in maritime cases where metadata (e.g., vessel trajectories, AIS logs, environmental sensors) are used, the main proposed findings are: adopt FAIR-principles and develop legal-data ontologies. Dublin Core's role is crucial for legal-technical interoperability. By anchoring an ontology in Dublin Core, the data citation tracking and impact metrics could feed into EOSC's monitoring of data usage and impact assessment.

Large-language model ontology construction would allow maritime authorities to achieve better predictability in their legal judgments. As in practice most institutions ended up making their own ontologies, cross-sectoral interactions with the EOSC in the traffic data ecosystem and themaritimedataspace initiatives would lead to beneficial synthesis.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Enrique Wulff

Acquisition, analysis, or interpretation of data: Enrique Wulff

Drafting of the manuscript: Enrique Wulff

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Tanhua T, Pouliquen S, Hausman J, et al.: Ocean FAIR Data Services. Frontiers in Marine Science. 2019, 6:440. 10.3389/fmars.2019.00440
2. Wilkinson M D, Dumontier M, Aalbersberg I J, et al.: The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data. 2016, 3:160018. 10.1038/sdata.2016.18
3. Cannon M, Kelly A, Freeman C: Implementing an Open & FAIR data sharing policy—A case study in the earth and environmental sciences. Learned Publishing. 2022, 35:56-66. 10.1002/leap.1442
4. Alves T: FAIR data: what it is and how we can support its principles. Science Editor. 2024, 47:110-12. 10.36591/se-4703-04
5. Kinkade D, Shepherd A: Geoscience data publication: Practices and perspectives on enabling the FAIR guiding principles. Geoscience Data Journal. 2022, 9:177-86. 10.1002/gdj3.120
6. Parsons MA, Fox PA: Is data publication the right metaphor?. Data Science Journal. 2013, 12:WDS32-WDS46. 10.2481/dsj.wds-042
7. Rovetto RJ: An ontology for satellite databases. Earth Science Informatics. 2017, 10:417-27. 10.1007/s12145-017-0290-x
8. Riga M, Kontopoulos E, Ioannidis K, Kintzios S, Vrochidis S, Kompatsiaris I: EUCISE-OWL: An ontology-based representation of the Common Information Sharing Environment (CISE) for the maritime domain. Semantic Web. 2021, 12:603-15. 10.3233/sw-200403
9. European Maritime Safety Agency (EMSA). Annual Report 2022. Lisbon, Portugal: European Maritime Safety Agency, 2023.. (2023). Accessed: May 27, 2025: https://www.emsa.europa.eu/publications/anual-reports.html.
10. FORCE 11. The FAIR Data Principles. Davis, CA: FORCE11. (2021). Accessed: May 27, 2025: https://force11.org/info/the-fair-data-principles/.
11. Aussenac-Gilles N, Comparot C, Dupuy A, El Malki N, Tournier R, Tran BH, dos Santos CT: eCISE-OWL: Représentation OWL du Schéma de Données de l'Environnement Commun d'Échange d'Information pour le Domaine Maritime. 33èmes Journées Francophones d'Ingénierie des Connaissances (IC 2022). AFIA - Association Française pour l'Intelligence Artificielle, Saint-Etienne, France; 2022. 82-91.
12. Berners-Lee T: Linked Data-Design Issues.. (2006). Accessed: August 25, 2025: https://www.w3.org/DesignIssues/LinkedData.
13. Yahya M, Zhou B, Breslin JG, Ali MI, Kharlamov E: Semantic modeling, development and evaluation for the resistance spot welding industry. IEEE Access. 2023, 11:37360-7377. 10.1109/access.2023.3267000
14. Poveda-Villalón M, Gómez-Pérez A, Suárez-Figueroa MC: OOPS! (Ontology pitfall scanner!): An on-line tool for ontology evaluation.. International Journal on Semantic Web & Information Systems.. 2014, 10:7-34. 10.4018/ijswis.2014040102
15. Kalita D, Deka D: Searching the great metadata timeline: A review of library metadata standards from linear cataloguing rules to ontology inspired metadata standards. Library Hi Tech. 2021, 39:190-204. 10.1108/lht-08-2019-0168
16. Yun J, Lifton E, Lee E, et al.: W3ID: A Quantum Computing-Secure Digital Identity System Redefining Standards for Web3 and Digital Twins. Preprint: Computer Science, Cryptography and Security. 2025, 10.48550/arXiv.2501.09802
17. Wilkinson SR, Aloqalaa M, Belhajjame K, et al.: Applying the FAIR Principles to computational workflows. Scientific Data. 2025, 12:328. 10.1038/s41597-025-04451-9
18. Camossi E, de Rosa F: Semantic-driven modelling of context and entity of interest profiles for maritime situation awareness. Proceedings of the Joint Ontology Workshops co-located with the Bolzano Summer of Knowledge (BOSK 2020). 2020.
19. EUR-Lex database. Luxembourg: Publications Office of the European Union. (2025). Accessed: May 27, 2025: https://eur-lex.europa.eu/.
20. Gao T, Jin J, Ke ZT, Moryoussef G: A comparison of deepseek and other LLMs. Preprint: Computer Science, Computation and Language. 2025, 10.48550/arXiv.2502.03688
21. Leadbetter A, Conway A, Attard A, et al.: Metadata standard for maritime spatial plans: A common structure to describe data associated with maritime spatial plans. Technical Expert Group (TEG) on Data for MSP.

Supported by CINEA and DG MARE (EC) (ed): Publications Office of the European Union, 2023. 10.2926/751919