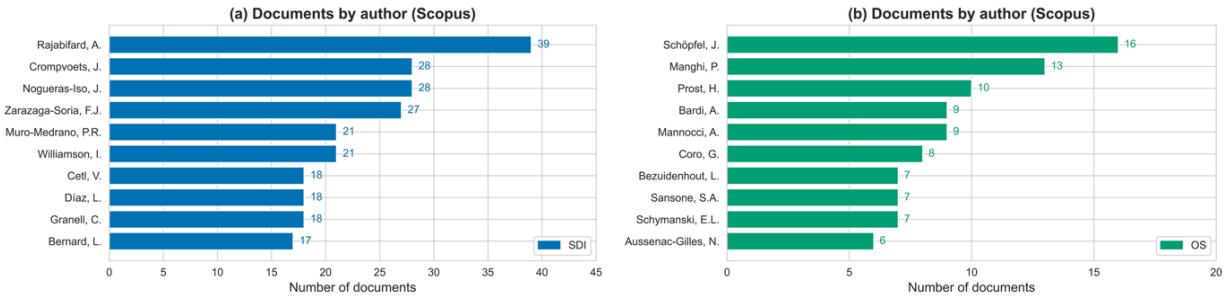
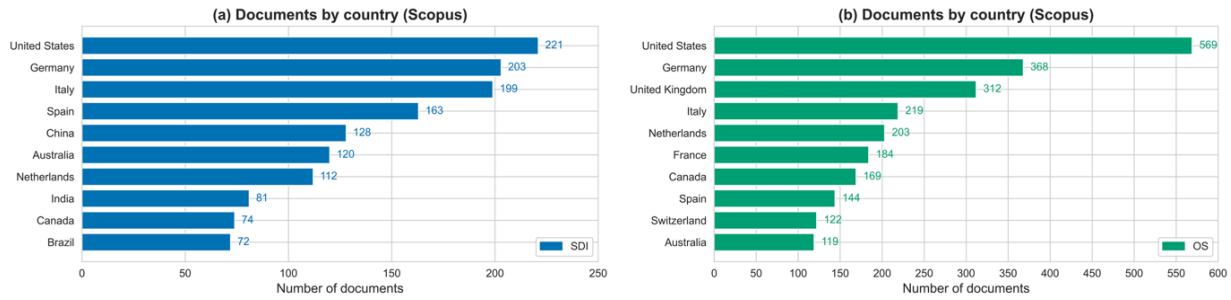
Supplementary Information

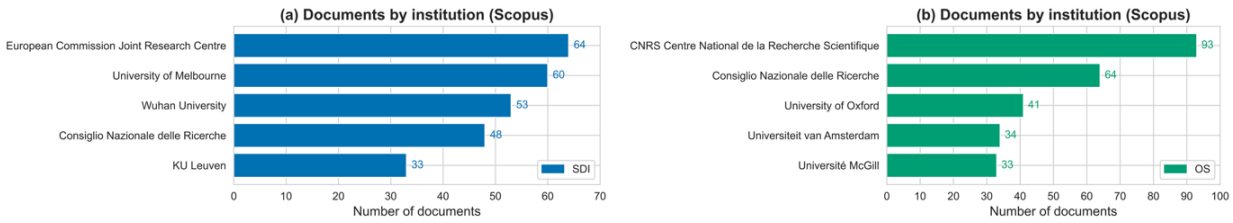
Supplementary Figures



Supplementary Figure S1|Top 10 authors ranked by number of publications for each theme: (a) SDI and (b) OS.

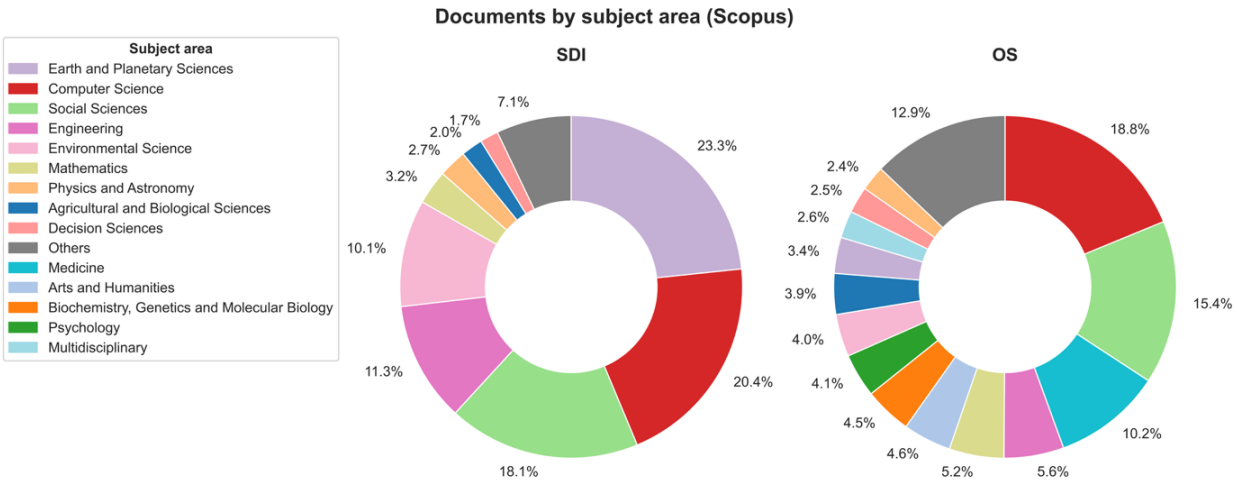


Supplementary Figure S2|Top 10 countries ranked by number of publications for each theme: (a) SDI and (b) OS.



Supplementary Figure S3|Top five institutions ranked by number of publications for each theme: (a) SDI and (b) OS.

Despite the publication leadership of the US at country level, no US institution appears among the top five for either topic. Instead, the SDI ranking is led by the European Commission Joint Research Centre, while the OS ranking is led by France's Centre National de la Recherche Scientifique. Italy's Consiglio Nazionale delle Ricerche is the only institution appearing in the top five for both fields, while Belgium's KU Leuven is included in the SDI ranking despite Belgium falling outside the top 10 countries by publication volume.



Supplementary Figure S4 | Percentage distribution of documents across major disciplines for SDI and OS.

The plots show the disciplinary distribution of documents for both domains, with SDI concentrated across fewer fields and OS exhibiting a broader disciplinary spread. For SDI, the leading subject area is “Earth and Planetary Sciences”, reflecting its focus on geospatial data and Earth observation. “Computer Science” and “Engineering” represent technological foundations, while “Social Sciences” reflects policy and governance dimensions. For OS, the distribution is led by “Computer Science”, where foundational openness concepts, such as the open-source movement, emerged. “Social Sciences” and “Medicine” also account for significant shares, reflecting established open-access and data-sharing practices within the health research community.

Supplementary Tables

Supplementary Table S1|FAIR principles [1].

Findable	The first step in (re)using data is to find them. Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services, so this is an essential component of the FAIRification process.
Accessible	Once the user finds the required data, she/he/they need to know how they can be accessed, possibly including authentication and authorization.
Interoperable	The data usually need to be integrated with other data. In addition, the data need to interoperate with applications or workflows for analysis, storage, and processing.
Reusable	The goal of FAIR is to optimize the reuse of data. To achieve this, metadata and data should be well-described so that they can be replicated and/or combined in different settings.

Supplementary Table S2 | Top 10 countries ranked by total link strength (TLS) in the co-authorship network.

The co-authorship analysis identified 79 countries, out of 202 total, meeting the inclusion criterion of at least five documents. TLS quantifies the overall intensity of a country's collaborative ties within the network; links represent co-authorship relationships between countries, and link strength reflects collaboration frequency. Absolute metrics are contrasted with ratio-based indicators to provide additional insight into collaboration intensity and citation intensity relative to publication volume.

Country	TLS	Links	Documents	Citations	TLS-to-document ratio	Citation-to-document ratio
US	1203	70	788	23153	1.5	29.4
Germany	1193	71	568	16866	2.1	29.7
UK	1072	67	380	14273	2.8	37.6
Netherlands	872	65	313	13221	2.8	42.2
France	723	62	232	11517	3.1	49.6
Italy	680	57	416	12418	1.6	29.9
Spain	599	64	305	11734	2.0	38.5
Australia	579	61	237	12584	2.4	53.1
Canada	558	55	240	11604	2.3	48.4
Belgium	543	57	150	9662	3.6	64.4

Supplementary Table S3 | Top 20 keywords ranked by total link strength (TLS) in the co-occurrence network.

The co-occurrence analysis identified 131 author keywords, out of 8,378 total, meeting the inclusion criterion of at least 10 occurrences, following targeted semantic harmonization, such as merging variants of "Spatial Data Infrastructure(s)" and "SDI". The keyword landscape is dominated by "Open Science", "Spatial Data Infrastructure", "Interoperability", "Metadata" and "Reproducibility". "Open Science" is the most frequent and highly connected term, indicating its role as a central conceptual hub. Several related but operationally distinct terms were intentionally not consolidated, including "FAIR Principles", "FAIR" and "FAIR Data". Within this network, "Standards" shows the highest TLS-to-occurrence ratio among the top 20 keywords, indicating strong systemic connectivity, whereas "Spatial Data Infrastructure" exhibits the lowest ratio, reflecting high occurrence but comparatively fewer diverse cross-domain connections.

Keyword	TLS	Links	Occurrences	TLS-to-occurrence ratio
Open Science	1256	94	811	1.5
Spatial Data Infrastructure	885	96	728	1.2
Interoperability	350	83	158	2.2
Metadata	321	72	135	2.4
Reproducibility	311	61	131	2.4
Open Data	293	61	124	2.4
Geographic Information System	285	69	182	1.6
Data Sharing	266	62	104	2.6
FAIR Principles	231	54	109	2.1
Open Access	229	48	110	2.1
FAIR	192	49	89	2.2
FAIR Data	176	39	76	2.3
INSPIRE	162	44	90	1.8
Data Management	142	52	55	2.6
Research Data Management	142	36	62	2.3
Standards	134	54	49	2.7
Ontology	127	40	69	1.8
Artificial Intelligence	107	37	45	2.4
Transparency	102	28	43	2.4
Web Service	102	39	56	1.8

Supplementary Table S4 | FAIR coding criteria used in the qualitative synthesis.

The RDA FAIR Data Maturity Model was used as a reference framework to guide the qualitative coding of FAIR implementation. Instead of assigning formal FAIRness scores to individual sources, the analysis identified concrete mechanisms described in the literature that correspond to findability, accessibility, interoperability, and reusability. These mechanisms were then used to support the integrative synthesis.

FAIR Principle	RDA-informed coding criterion used in this review	Evidence coded under this principle
Findable	Sources were coded as Findable (**F**) when they described mechanisms that support the discovery, identification, indexing, or citation of marine data, metadata, services, software, workflows, or other research objects.	Rich metadata; persistent unique identifiers; searchable catalogues; indexed repositories; metadata harvesting; geoportals; thematic portals; data and metadata citation.
Accessible	Sources were coded as Accessible (**A**) when they described mechanisms that enable data, metadata, services, or research objects to be retrieved through clear technical protocols and documented access conditions.	Standardized access protocols; application programming interfaces (API); Open Geospatial Consortium (OGC) web services and API; Environmental Research Division Data Access Program (ERDDAP); automatic and manual download services; authentication and authorization; access policies and conditions.
Interoperable	Sources were coded as Interoperable (**I**) when they described mechanisms that allow data, metadata, services, vocabularies, or workflows to be exchanged, combined, interpreted, or processed across systems, disciplines, applications, or communities.	International and community standards, vocabularies, and best practices promoted or implemented by organizations, networks, and initiatives such as the International Council for the Exploration of the Sea (ICES), International Hydrographic Organization (IHO), International Oceanographic Data and Information Exchange (IODE), International Organization for Standardization (ISO), OGC, SeaDataNet, European Marine Observation and Data Network (EMODnet), and Infrastructure for Spatial Information in Europe (INSPIRE); controlled vocabularies; ontologies; semantic mapping; knowledge graphs; linked data; metadata brokering; harmonized data models; machine-readable data and metadata.
Reusable	Sources were coded as Reusable (**R**) when they described mechanisms that support future use, trust, reproducibility, interpretation, quality assessment, or fitness-for-purpose evaluation of data, services, workflows, applications, or research outputs.	Quality assurance (QA) and quality control (QC); validated datasets; provenance or lineage; licensing; documentation; data management plans; open code; notebooks; containers; reproducible workflows; trusted repositories; quality management systems (QMS).

Supplementary Table S5 | FAIR implementation mechanisms identified in the qualitative corpus.

The table groups the main implementation mechanisms identified during the qualitative synthesis according to the FAIR principle to which they primarily contribute. Some sources appear under more than one implementation mechanism because individual publications address multiple FAIR principles or describe mechanisms that contribute to more than one aspect of FAIR implementation. Implementation logic indicates whether each mechanism is mainly operationalized through an MSDI perspective, an OS perspective, or through both perspectives.

FAIR principle	Implementation mechanism	Operational meaning	Implementation logic	Supporting sources
Findable	Metadata catalogues, geoportals, and thematic portals.	Enable users to discover marine datasets, services, maps, and data products across distributed infrastructures.	MSDI	2–8
	Persistent identifiers, citation, and indexed repositories.	Support unambiguous identification, citation, tracking, and long-term discovery of datasets, services, workflows, and research outputs.	MSDI-OS	9–13
	Federated discovery and metadata brokering.	Allow metadata from multiple marine data infrastructures to be harvested, mapped, brokered, and exposed through common discovery interfaces.	MSDI-OS	14–18
Accessible	Standardized web services, Application Programming Interfaces (API), and download mechanisms.	Enable users and machines to retrieve marine data and metadata through documented protocols and service interfaces.	MSDI-OS	4,6,11,19–22
	Data portals, repositories, cloud services, and access to virtual research environments (VRE).	Provide user interfaces and virtual environments where data can be accessed, processed, visualized, and shared.	MSDI-OS	13,18,23–27
	Access policies, conditions, authentication, and authorization.	Define who can access data or services, under what conditions, and through which technical or institutional procedures.	MSDI-OS	9,28–31
Interoperable	Common standards, data models, and service specifications.	Enable heterogeneous marine data to be structured, exchanged, and processed according to shared technical rules.	MSDI	6–8,11,19–21,32
	Controlled vocabularies, ontologies, semantic mapping, and linked data.	Support shared interpretation of marine data across systems, disciplines, communities, and applications.	MSDI-OS	9,14,15,22,33
	Harmonized data products and cross-infrastructure integration.	Convert distributed and heterogeneous datasets into comparable products for assessment, modelling, and decision support.	MSDI-OS	4,5,17,18,34,35
Reusable	Quality assurance (QA) and quality control (QC), validation, and fitness-for-purpose information.	Support trust, interpretation, and assessment of whether data are suitable for new scientific, operational, or policy uses.	MSDI	4,11,13,29,30,34,35
	Provenance, lineage, uncertainty, and documentation.	Preserve contextual information needed to interpret data origin, processing history, transformations, and limitations.	MSDI-OS	2,7,9,11–13
	Open code, notebooks, containers, reproducible workflows, and VRE.	Enable analyses to be repeated, adapted, shared, and reused by different users and communities.	MSDI-OS	13,17,18,23,24,26,27
	Trusted stewardship, repository certification, quality management, and data culture.	Provide institutional arrangements that ensure data preservation, service reliability, and long-term reuse.	MSDI-OS	10,28–31,36
	Analysis-ready and machine-ready datasets, libraries, and decision-support tools.	Transform curated data into reusable products that support modelling, Artificial Intelligence (AI), monitoring, restoration, and policy decisions.	MSDI-OS	12,13,26,27,37,38

References

1. GO FAIR Initiative. GO FAIR initiative: Make your data & services FAIR. https://www.go-fair.org/ (2025).
2. Maguire, D. J. & Longley, P. A. The emergence of geoportals and their role in spatial data infrastructures. *Comput. Environ. Urban Syst.* **29**, 3–14 (2005).
3. Lathrop, R. G. *et al.* The role of Mid-Atlantic ocean data portal in supporting ocean planning. *Front. Mar. Sci.* **4**, (2017).
4. Giorgetti, A. *et al.* EMODnet Chemistry Spatial Data Infrastructure for marine observations and related information. *Ocean Coast. Manag.* **166**, 9–17 (2018).
5. Martín Míguez, B. *et al.* The European Marine Observation and Data Network (EMODnet): Visions and Roles of the Gateway to Marine Data in Europe. *Front. Mar. Sci.* **Volume 6-2019**, (2019).
6. Vaitis, M., Kopsachilis, V., Tataris, G., Michalakis, V. I. & Pavlogeorgatos, G. The development of a spatial data infrastructure to support marine spatial planning in Greece. *Ocean Coast. Manag.* **218**, (2022).
7. Foglini, F. & Grande, V. A Marine Spatial Data Infrastructure to manage multidisciplinary, inhomogeneous and fragmented geodata in a FAIR perspective the Adriatic Sea experience. *Oceanologia* **65**, 260–277 (2023).
8. Mohamad, Y. *et al.* Marine Geospatial Data Centre (MGDC): A Framework For Effective Marine Research Data Management In Malaysia. *An International Journal of Marine Sciences* **41**, 80 (2025).
9. Tanhua, T. *et al.* Ocean FAIR Data Services. *Front. Mar. Sci.* **Volume 6-2019**, (2019).
10. Robinson, E. *et al.* FAIR Island: real-world examples of place-based open science. *GigaScience* vol. 12 Preprint at https://doi.org/10.1093/gigascience/giad004 (2023).
11. Simoncelli, S., Reseghetti, F., Fratianni, C., Cheng, L. & Raiteri, G. Reprocessing of eXpendable BathyThermograph (XBT) profiles from the Ligurian and Tyrrhenian seas over the time period 1999–2019 with a full metadata upgrade. *Earth Syst. Sci. Data* **16**, 5531–5561 (2024).
12. Ohall, A., Bisson, K., Garaba, S. P. & Rivero-Calle, S. The MArine Debris hyperspectral reference Library collection (MADLib). *Earth Syst. Sci. Data* **17**, 7293–7311 (2025).
13. Wall, C. C. *et al.* Big data, sound science, lasting impact: A framework for passive acoustic monitoring. *Big Earth Data* https://doi.org/10.1080/20964471.2025.2583505 (2025) doi:10.1080/20964471.2025.2583505.
14. Jeffrey, K. G., Bailo, D., Paciello, R. & Vinciarelli, V. European Environmental Metadata. in *Metadata and Semantic Research - 16th Research Conference* (eds. Garoufallou, E. & Vlachidis, A.) (Springer, 2023). doi:10.1007/978-3-031-39141-5_24.
15. Boldrini, E. *et al.* Blue-cloud DAB: developing a platform to harmonize, assess and disseminate marine metadata collections. *Int. J. Data Sci. Anal.* https://doi.org/10.1007/s41060-024-00664-0 (2024) doi:10.1007/s41060-024-00664-0.

16. Larkin, K., Vera, J. & Pittonet, S. Blue-Cloud: Demonstrating the potential of Open Science for the ocean through a thematic EOSC. in *Open Science Fair 2021 (OSFair2021)* (Zenodo, 2021). doi:10.5281/zenodo.5535589.
17. Schaap, D., Assante, M., Pagano, P. & Candela, L. Blue-Cloud: Exploring and demonstrating the potential of Open Science for ocean sustainability. in *2022 IEEE International Workshop on Metrology for the Sea; Learning to Measure Sea Health Parameters, MetroSea 2022 - Proceedings* 198–202 (Institute of Electrical and Electronics Engineers Inc., 2022). doi:10.1109/MetroSea55331.2022.9950819.
18. Delvenne, C. *et al.* Blue-Cloud Virtual Labs in Support of Sustainable Development Goals. in *Oceans Conference Record (IEEE)* (Institute of Electrical and Electronics Engineers Inc., 2025). doi:10.1109/OCEANS58557.2025.11104725.
19. Abramic, A., Gonzalez, D., Bigagli, E., Che-Bohnenstengel, A. & Smits, P. INSPIRE: Support for and requirement of the Marine Strategy Framework Directive. *Mar. Policy* **92**, 86–100 (2018).
20. Abramic, A. *et al.* Maritime spatial planning supported by infrastructure for spatial information in Europe (INSPIRE). *Ocean Coast. Manag.* **152**, 23–36 (2018).
21. Wulff, E. Open Standards used in Oceanography Research Spatial Data Repositories in Spain. *DESIDOC Journal of Library & Information Technology* **40**, 306–312 (2020).
22. De Sousa, L. M., Palma, R., Janiak, B. & Van Genuchten, P. The template for a Semantic SensorThings API with the GloSIS use case. in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* vol. 48 29–34 (International Society for Photogrammetry and Remote Sensing, 2024).
23. Assante, M. *et al.* Enacting open science by D4Science. *Future Generation Computer Systems* **101**, 555–563 (2019).
24. Coro, G., Ellenbroek, A. & Pagano, P. An Open Science approach to infer fishing activity pressure on stocks and biodiversity from vessel tracking data. *Ecol. Inform.* **64**, (2021).
25. Tsontos, V. M. *et al.* CEOS Ocean Variable Enabling Research & Applications for GEO (COVERAGE): A Platform to Simplify and Expand the Accessibility and Usage of Inter-agency Satellite and in-situ Oceanographic Data. in *Oceans Conference Record (IEEE)* vols 2022-October (Institute of Electrical and Electronics Engineers Inc., 2022).
26. Azmi, E. *et al.* Best practices for AI-based image analysis applications in aquatic sciences: The iMagine case study. *Ecol. Inform.* **91**, (2025).
27. Veylit, L. *et al.* Democratizing marine restoration best practices in a digital toolbox. *Int. J. Data Sci. Anal.* **21**, (2026).
28. Bertelli, A., Acciai, M. & Rossi, G. The European Open Science Cloud as a common good Potentials and limitations of this endeavour. *Open Research Europe* **5**, (2025).

29. Flynn, S. *et al.* Establishing a Data Culture Using Frameworks to Navigate the Waves of Marine Data. *Data Sci. J.* **24**, (2025).
30. IOC. *IODE Quality Management Framework for National Oceanographic Data Centres and Associate Data Units*. (UNESCO, Paris, 2025).
31. Mesman, J. P. *et al.* Challenges of open data in aquatic sciences: issues faced by data users and data providers. *Front. Environ. Sci.* **12**, (2024).
32. Contarinis, S., Pallikaris, A. & Nakos, B. The value of marine spatial open data infrastructures-potentials of IHO S-100 standard to become the universal marine data model. *J. Mar. Sci. Eng.* **8**, (2020).
33. Al-Subhi, N. H. & Al-Suqri, M. N. Integrating Conceptual Modelling and Ontologies in Marine Spatial Data Infrastructure Design. *Int. J. Semant. Web Inf. Syst.* **21**, (2025).
34. French, M. A. & Lipizer, M. Improving data reliability to support marine pollution assessment according to MSFD Descriptor 8 in the European Seas: the contribution of EMODnet Chemistry. *Front. Mar. Sci.* **10**, (2023).
35. Molina Jack, M. E. *et al.* A Mediterranean Focus Overview of EU Marine Litter Data Management Performed in the Framework of the European Marine Observation and Data Network Chemistry. *Sustainability (Switzerland)* **15**, (2023).
36. Oestreich, W. K., Czapanskiy, M. F., Katija, K., Record, N. R. & Chapman, M. S. Collective Science to Inform Global Ocean Protections. *Ecol. Lett.* **28**, (2025).
37. Burger, E. F. *et al.* Data Processing and Management at PMEL: A 50-Year Perspective. *Oceanography* **36**, 26–31 (2023).
38. Schaap, D., Thijsse, P., Krijger, T. & Kooyman, R. Blue-Cloud 2026 project- Deploying BEACON data lakes for harmonizing ocean data access for Virtual Research Environments. in *EGU General Assembly 2025* (Vienna, 2025). doi:10.5194/egusphere-egu25-4993.

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Title page, Abstract and Methods. The paper is presented as an integrative review; the Methods section specifies the PRISMA-guided review process used for source identification, screening, eligibility and inclusion.
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Abstract. The abstract reports the review aim, methodological approach, data sources, qualitative corpus size, synthesis approach, main findings and priorities identified.
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Introduction. The rationale is reported in the paragraphs describing the importance of marine data management, MSDI, OS, FAIR principles, and the need to align MSDI and OS for sustainable ocean knowledge systems.
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Introduction. The objective and research questions are reported in the paragraph beginning “This study aims to investigate…” and in RQ1 and RQ2.
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Methods; Data collection and selection; Qualitative analysis. Eligibility criteria are reported through the description of database selection, language restriction, search fields, publication-year restriction for the qualitative corpus, marine-domain focus, screening process, full-text eligibility assessment, and exclusion of records outside the thematic scope.
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Methods; Data collection and selection; Table 2. Information sources are reported as Scopus and Web of Science, complemented by citation-based snowballing. The database search date is reported in Table 2.
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Table 2; Data collection and selection. Search strategies, search fields, search queries, filters and limits are reported in Table 2 and the accompanying Methods text.
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Methods; Data collection and selection; Figure 7. The selection process is reported through the PRISMA workflow, duplicate removal, title/abstract screening, full-text retrieval, eligibility assessment, and inclusion counts.
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Methods; Qualitative analysis; Supplementary Tables S4 and S5. The data collection and coding process is reported through the description of source-level coding, use of FAIR principles as analytical categories, grouping of implementation mechanisms, and recording of operational meaning, implementation logic and supporting sources.
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all	Not applicable in the intervention-review sense. The review did not extract clinical, experimental or effect outcomes. Instead, the review

Section and Topic	Item #	Checklist item	Location where item is reported
		measures, time points, analyses), and if not, the methods used to decide which results to collect.	sought bibliometric indicators and FAIR implementation mechanisms, as described in Methods, Quantitative analysis and Qualitative analysis.
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Methods; Quantitative analysis; Qualitative analysis; Supplementary Tables S4 and S5. Other data items are reported as publication and citation trends, citation-to-document ratios, co-authorship patterns, country-level collaboration, subject-area distribution, keyword co-occurrence, and FAIR implementation mechanisms related to findability, accessibility, interoperability and reusability.
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Not applicable. The review did not assess intervention effects, clinical outcomes or methodological risk of bias in individual studies. Instead, the review used bibliometric analysis and integrative qualitative synthesis; limitations of the review process are discussed in Discussion.
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Not applicable. No effect measures, such as risk ratios, mean differences or effect sizes, were used because no statistical meta-analysis or intervention-effect synthesis was conducted.
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Methods; Data collection and selection; Quantitative analysis; Qualitative analysis. The processes used to decide which records contributed to each synthesis are reported through the distinction between the quantitative dataset based on queries 1 and 4, the marine-domain qualitative corpus based on the final query, and additional sources identified through snowballing.
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Methods; Quantitative analysis; Qualitative analysis. Data preparation is reported through the use of Scopus records for bibliometric consistency, VOSviewer analysis, threshold selection for author keywords, targeted semantic harmonization of equivalent terms, source-level coding, and grouping of FAIR implementation mechanisms.
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Methods; Results; Figures 1–7; Tables 1–2; Supplementary Figures S1–S4; Supplementary Tables S2–S5. Methods used to tabulate and visually display results include PRISMA workflow visualization, bibliometric maps, annual publication and citation plots, country collaboration networks, keyword co-occurrence and overlay visualizations, and synthesis tables.
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Methods; Quantitative analysis; Qualitative analysis; Results. The synthesis methods are reported as bibliometric analysis using VOSviewer and integrative qualitative synthesis structured around the four FAIR principles. No meta-analysis was performed.
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Not applicable in the statistical sense. No meta-analysis, subgroup analysis or meta-regression was conducted. Conceptual and implementation differences across sources, FAIR principles and infrastructure contexts are discussed in Results and Discussion.

Section and Topic	Item #	Checklist item	Location where item is reported
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Not applicable. No sensitivity analyses were conducted because no statistical synthesis or meta-analysis was performed.
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Not applicable. No formal assessment of reporting bias due to missing results was conducted because the review did not synthesize intervention outcomes or effect estimates. Relevant limitations related to database coverage, English-language restriction, inaccessible records and possible underrepresentation of technical or policy documents are discussed in Discussion.
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Not applicable. No formal certainty or confidence assessment of an intervention-effect body of evidence was conducted. The review provides a bibliometric and integrative qualitative synthesis rather than a certainty-graded outcome synthesis.
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Methods; Data collection and selection; Figure 7. The study selection process is reported from records identified through the final query, duplicate removal, title/abstract screening, full-text retrieval, full-text eligibility assessment, database-derived inclusions, snowballing inclusions and final qualitative corpus size.
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Data collection and selection; Figure 7; Data Availability. Exclusion counts and broad reasons are reported in the PRISMA workflow and Methods text. The records considered in each review stage, including excluded records, are provided in the publicly available dataset reported in the Data Availability section.
Study characteristics	17	Cite each included study and present its characteristics.	Results; Table 1; Supplementary Table S5; References. Included sources are cited throughout the Results section, and the source-level FAIR implementation mechanisms and supporting sources are reported in Supplementary Table S5.
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Not applicable. Risk of bias assessments for individual studies were not conducted because the review did not evaluate intervention effects or clinical outcomes.
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Not applicable in the effect-estimate sense. The review did not extract summary statistics, group comparisons or effect estimates from individual studies. Source-level qualitative evidence is reported in Supplementary Table S5 and synthesized in Results and Table 1.
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Results; Table 1; Supplementary Table S5. The characteristics of contributing sources are summarized through FAIR implementation mechanisms, MSDI and OS contributions, integration value, and remaining gaps. Risk of bias was not assessed because it was not applicable to this review design.
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Not applicable. No statistical synthesis or meta-analysis was conducted. Bibliometric results are reported in Results and Figures 1–4.

Section and Topic	Item #	Checklist item	Location where item is reported
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Not applicable in the statistical sense. No statistical investigation of heterogeneity was conducted. Differences across domains, FAIR principles, infrastructures and governance contexts are interpreted qualitatively in Results and Discussion.
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	Not applicable. No sensitivity analyses were conducted because no statistical synthesis or meta-analysis was performed.
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Not applicable. No formal assessment of reporting bias due to missing results was conducted. Review-process limitations are discussed in Discussion.
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Not applicable. No formal certainty assessment was conducted because the review did not synthesize intervention-effect evidence.
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Discussion. The general interpretation of the results is provided in the opening Discussion paragraphs and in the subsequent interpretation of MSDI-OS integration, Blue-Cloud, FAIR implementation and the data-information-knowledge continuum.
	23b	Discuss any limitations of the evidence included in the review.	Discussion. Limitations of the included evidence are discussed in the paragraph beginning “Several limitations should be considered…”, including possible underrepresentation of regional, national, non-English, technical, operational and policy sources.
	23c	Discuss any limitations of the review processes used.	Discussion. Limitations of the review process are discussed in the same limitations paragraph, including English-language restriction, reliance on Scopus for quantitative analysis, database coverage, 2017-onward qualitative corpus, snowballing and Blue-Cloud as an illustrative rather than evaluative case.
	23d	Discuss implications of the results for practice, policy, and future research.	Discussion. Implications for practice, policy and future research are discussed in the integration pathways, governance, machine-actionability, sustainable ocean knowledge systems and future work paragraphs.
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Not registered. The review was not registered.
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	No separate review protocol was prepared.
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Not applicable. No registration or protocol amendments were made because the review was not registered and no separate protocol was prepared.
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Acknowledgements. Sources of financial support are reported in the Acknowledgements section.
Competing interests	26	Declare any competing interests of review authors.	Competing Interests. Competing interests are declared in the Competing Interests section.

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Data Availability; Code Availability; Supplementary Information. The bibliographic datasets and code used to process the datasets are publicly available through the Zenodo DOI reported in the Data Availability and Code Availability sections. Additional review materials are provided in the Supplementary Information.