



# Establishing a Data Culture Using Frameworks to Navigate the Waves of Marine Data

RESEARCH PAPER

SARAH FLYNN 

TARA KEENA

YVONNE BOGAN

LAURA BROPHY

DAVID CURRIE

ADAM LEADBETTER

MARTINA MALONEY

KEITH MANSON

COLIN MELVILLE

EOIN O'GRADY

ROB THOMAS

BRENDAN WHITTLE

ANDREW CONWAY

 ubiquity press

[\\*Author affiliations can be found in the back matter of this article](#)

## ABSTRACT

The Marine Institute has worked to establish itself as a trusted source of FAIR (Findable, Accessible, Interoperable, Reusable) data through internationally recognised frameworks such as the Data Management Quality Management Framework and CoreTrustSeal. These have supported the effective handling of complex, multi-stakeholder marine data central to Ireland's Marine Spatial Planning. This paper presents a summary of these frameworks, and how they have evolved to meet ongoing needs, an analysis of how they fulfil international standards and principles, and lessons learned during their implementation.

Preserving marine data is essential to maximise the long-term value of oceanographic research and environmental monitoring. Marine datasets, ranging from biodiversity and ocean chemistry to climate trends, require extensive resources to gather. Ensuring their preservation allows future researchers to analyse historical changes, improve predictive models, and collaborate across disciplines and regions. Standardisation, secure storage, and proper documentation are vital for keeping data accessible, interpretable, and trustworthy.

The Marine Institute's approach highlights that frameworks and certifications, while critical, are not enough. Success depends on a data culture that supports skilled and empowered people, clear communication, and continuous learning. Technical systems must be matched by human engagement and adequate resourcing. Marine data management is a collaborative discipline, requiring integration of people, technologies, and methodologies to ensure data quality and reuse. Through continued leadership support, staff engagement and a willingness to embrace change across all levels, the Marine Institute's marine data management is well positioned for the future.

## CORRESPONDING AUTHOR:

**Sarah Flynn**

Marine Institute, Oranmore,  
Galway, Ireland

[sarah.flynn@marine.ie](mailto:sarah.flynn@marine.ie)

## KEYWORDS:

Data Management; Marine Data Governance & Reuse; FAIR Principles; Digital Preservation; CoreTrustSeal; Expectation Management; People

## TO CITE THIS ARTICLE:

Flynn, S., Keena, T., Bogan, Y., Brophy, L., Currie, D., Leadbetter, A., Maloney, M., Manson, K., Melville, C., O'Grady, E., Thomas, R., Whittle, B. and Conway, A. 2025 Establishing a Data Culture Using Frameworks to Navigate the Waves of Marine Data. *Data Science Journal*, 24: 24, pp. 1–19. DOI: <https://doi.org/10.5334/dsj-2025-024>

## 1. INTRODUCTION

### IMPORTANCE OF DATA MANAGEMENT IN MARINE SCIENCE

Data management in marine science is critical for ensuring the integrity, accessibility, and usability (Flynn et al., 2023; Owens et al., 2022) of the vast amounts of data collected from marine research, including oceanographic observations and model results, socio-economic data and data concerning other human activities in the marine environment. With the complexity and scale of marine ecosystems (Fogarty et al., 2016), scientists, policy-makers, and researchers rely on well-organised data to track environmental changes, study marine species, and understand ocean processes. Effective data management enables the sharing and integration of datasets across institutions and disciplines, facilitating collaboration and ensuring that data remains available for future research. It also enhances the accuracy of predictive models (Ali et al., 2021) and decision-making for conservation and sustainable use of marine resources. Inadequate data management can lead to data loss, misinterpretation, or inefficiencies, hindering scientific progress and the development of marine policies (Pasquetto et al., 2024).

Preserving marine data for reuse is essential for maximising the value of oceanographic research and ensuring long-term environmental monitoring (Carolyn, 2023; Richards et al., 2013). Marine data, whether it pertains to ocean chemistry, biodiversity, or climate impacts, often requires significant resources to collect (Shucksmith and Kelly, 2014). By preserving this rich data, researchers enable future generations to analyse historical trends, validate new findings, and refine predictive models with richer datasets (Zilioli et al., 2021). Reusable data also fosters collaboration across disciplines, regions, and timeframes, creating opportunities for discoveries that were not possible during initial research. Ensuring proper documentation, standardised formats, and secure storage systems is key to making this data accessible (Kapiszweski and Karcher, 2020), interpretable (Coussement and Benoit, 2021), and trustworthy (Koedel et al., 2022). Without preservation efforts, critical information could be lost, weakening our ability to monitor and protect marine ecosystems in the face of ongoing environmental changes (Tray et al., 2020).

Data Governance is key to reuse (Cerrillo-Mártinez and Casadesús-de-Mingo, 2021) without this clarity there would be very little reuse as a precautionary approach would likely be adopted (Marcucci et al., 2023). Data reuse in marine science saves time and money by eliminating the need to repeatedly collect similar datasets, which can be costly and logistically challenging in demanding ocean environments. By making previously collected data accessible, researchers can build on existing knowledge, avoiding duplication of effort, and resource expenditure. This allows them to focus on new analyses, advanced modelling, or filling knowledge gaps rather than repeating fieldwork or experiments. Reusing well-preserved data also accelerates scientific discovery (Cousijn et al., 2022), as researchers can quickly compare historical and current data, leading to more efficient and cost-effective research outcomes. Moreover, it supports collaborative efforts, where institutions can share and benefit from each other's datasets, maximising the return on investment in marine research.

This paper asks if data management frameworks can be effective in helping to establish policies and processes to govern the collection, management, and long-term preservation of marine data. It describes the methodological approaches implemented by a multidisciplinary team at the Marine Institute to achieve accreditation and certification for data management practices. It discusses how engaged, motivated and supported people are integral to the process. Whilst the paper specifically focuses on the Marine Institute many of its findings will apply to other organisations working in similar data management contexts.

## 2. BACKGROUND AND CONTEXT

### OVERVIEW OF THE MARINE INSTITUTE

The Marine Institute (MI) is the State agency responsible for marine research, technology development, and innovation in Ireland, providing scientific and technical advice to Government to help inform marine policy and to support the sustainable development of Ireland's marine resource. The MI was set up under the Marine Institute Act 1991: 'to undertake, to coordinate, to promote and to assist in marine research and development and to provide such services

related to research and development that in the opinion of the Institute, will promote economic development and create employment and protect the marine environment.’ Clause 4.2(e) of the Marine Institute Act (1991) further states that a function of the Institute is ‘to collect, maintain and disseminate information relating to marine matters’ (eISB, 1991).

As a dynamic and client-focused national agency, the MI provides cross-government supports for national marine programmes such as for fisheries and aquaculture management, marine planning, marine environment, offshore renewable energy policy, transport, and marine research across many areas (Flynn et al., 2024).

## **NATIONAL OCEANOGRAPHIC DATA CENTRE AND DIGITAL ECOSYSTEM**

The MI plays a central role in Ireland’s national and international marine research efforts, acting as a key node within a broad network of programmes that advance ocean knowledge, sustainable development, and climate resilience. The MI actively contributes to the sustainable management of oceans, coasts, and marine ecosystems as part of a global network of Member States within the Intergovernmental Oceanographic Commission (IOC) of UNESCO, which promotes international collaboration in marine science. This role is underpinned by the operation of Ireland’s National Oceanographic Data Centre (NODC), which provides a foundational platform for the acquisition, management, and dissemination of marine data from Ireland. Through enhanced digital connectivity, strategic partnerships, and participation in initiatives such as EMODnet (EMODnet, 2009) and Copernicus Marine Services (CMEMS, 2015), the MI ensures that data flows efficiently from acquisition to application. Data governance is anchored by robust policies and practices promoting Open Data and adherence to the FAIR principles (Wilkinson et al., 2016) ensuring transparency, reusability, and broad access.

As the NODC for Ireland, under its parent department, the Department of Agriculture, Food and the Marine (DAFM), the MI has played a leading role in marine data management and availability for over 30 years, operating a number of data acquisition platforms and programmes which underpin national marine related services (ERDDAP, 2015; MI, 2010). In addition to operational activities, the MI is a research-performing organisation and a marine research funder in Ireland (eISB, 1991).

Demonstrating the Marine Institute’s commitment to a multi-stakeholder digital ecosystem, users of Marine Institute services, including scientific advisory services and access to marine data and information are typically in academia and education, the European Commission, local authority organisations, government departments (e.g., Department of Agriculture, Food and Marine (DAFM), Department of Housing, Local Government and Heritage (DHLGH), Department of the Environment, Climate and Communications (DECC), and agencies (e.g., Environmental Protection Agency (EPA), Sustainable Energy Authority Ireland (SEAI), Geological Survey Ireland (GSI)) industry, and environmental consultants. Data from the MI is also used by international organisations such as the International Council for the Exploration of the Sea (ICES) and the International Commission for the Conservation of Atlantic Tunas (ICCAT). Users also include interested members of the public. The UN Decade of the Ocean, through its Challenge 8, aims to create a trusted, inclusive, and interconnected digital representation of the ocean. While the global model can support regional programmes, it will feed and rely on national and sub-national programmes, data, and services to meet the relevant Ocean Decade outcome – an accessible ocean with open and equitable access to data, information and technology and innovation. These will provide the data inputs to the global models and the granularity required for national and local decision making. This representation, through multi-stakeholder collaboration, will provide a dynamic map of the ocean and free, open access to ocean and marine data (Calewaert et al., 2024). In the context of such international collaboration, a value chain from observations or collection to end-user services has been proposed to facilitate cooperation in the creation of an ocean commons (Alvarez Fanjul et al., 2022; Bahrel et al., 2010).

## **CHALLENGES OF MARINE DATA MANAGEMENT WITHIN AN IRISH LANDSCAPE**

Despite the MI’s leadership in marine data management, several challenges persist in fully realising the value of marine data in Ireland. Some of these challenges are similar to other environmental data management organisations (Biber, 2013) whilst some are specific to the Irish context:

- Data collection can't be reproduced
  - Marine data is collected at a specific time and place. Whilst data is typically collected multiple times, often over many years, the data collected at a specific time cannot be directly reproduced since the environment that is being measured has changed.
- Timeseries
  - Marine data will typically need to be collected over many years, and there will often be changes of personnel, equipment, and methods during this period. The data must be managed during this extended period, and the changes correctly documented. It might also be the case that data cannot be used until the time-series has reached a minimum length.
- Cost
  - Some types of marine data collected are expensive, in particular, surveys on dedicated research vessels. These data collection programmes must be funded on an ongoing basis, often over many years. The management of the data collected should be considered a core cost and included within the budget, but this has not always happened.
- Culture
  - Ongoing marine data collection might not yield as many scientific publications as novel experimentation, making it less attractive to scientists. Data is often collected for a specific purpose – once this initial purpose is completed the collectors might not have the time or interest to document and publish the data for use outside of their community. Data stewards might have received little training in how to make their data FAIR.
- Interoperability
  - Achieving syntactic interoperability (common formats) is only part of the solution; semantic interoperability, using shared vocabularies to align meaning, is essential to ensure datasets can be reliably discovered, understood, and integrated. This includes generalised vocabularies as well as domain-specific ontologies for marine and environmental sciences ([Schaap and Lowry, 2010](#)). When appropriately implemented, these vocabularies allow machine readability and enable the construction of knowledge graphs, which are increasingly used in advanced data discovery interfaces and Artificial Intelligence applications ([Han et al., 2025](#); [Leadbetter et al., 2014](#); [Liu and Cheng, 2024](#)). Such technologies hold particular promise in integrating diverse marine datasets to generate insights that would otherwise remain hidden.
- Technical limitations
  - The growing influx of real-time, high-resolution data from remote sensing, buoys, and autonomous systems is often underutilised due to processing limitations and funding constraints.
- Organisational and legislative complexity
  - Platforms like Ireland's Marine Atlas and ERDDAP have improved access to environmental and oceanographic datasets, yet data fragmentation remains an issue, with datasets held across multiple agencies in varying formats, with inconsistent metadata standards and access restrictions. This complicates the creation of integrated views necessary for critical applications such as Marine Spatial Planning and Climate Resilience Modelling. While the Marine Institute contributes to European platforms like EMODnet and Copernicus Marine Services, aligning national outputs with international standards demands ongoing, resource-intensive harmonisation efforts.
- Confidential and personal data
  - Ethical, legal, and governance concerns, particularly around sensitive data such as fisheries activity and protected habitats, further complicate the implementation of open and FAIR data principles.

- Diverse data collection communities
  - Implementing standards in marine data management is not uniform across disciplines. Physical oceanography, with its long-established involvement in international observing programmes (e.g., through EMODnet, Copernicus Marine Services), benefits from well-developed data standards, shared vocabularies, and common data formats. For instance, time-series data from ocean buoys or satellite observations are often standardised using protocols such as NetCDF, with metadata aligned to standards like ISO 19115 or SeaDataNet. In contrast, other domains' data management can lag behind, often relying on legacy systems, varied data collection methods, or restricted access due to commercial sensitivities. This leads to inconsistent data quality, incomplete metadata, or limited interoperability, which poses challenges for national-level integration and reuse.

Many of these challenges reflect wider global gaps in marine data curation and stewardship, including limited interoperability, sparse biological and biodiversity data, especially in underexplored regions and insufficient long-term preservation infrastructure.

Addressing the persistent gaps requires sustained investment, improved coordination, and robust national policies to ensure Ireland's marine data ecosystem can meet both domestic and global needs.

The diversity of marine data types ranging from high-frequency sensor data and remote sensing outputs to biological survey results, socio-economic datasets, and fisheries catch reports requires adaptable data governance. A one-size-fits-all approach does not suffice. Instead, the MI's strategy has been to apply and adapt data standards contextually, recognising where established frameworks exist and where further development or harmonisation is needed ([MI, 2023](#)).

These challenges are especially significant as emerging technologies, such as Digital Twins of the Ocean, place new demands on data quality, consistency, and timeliness. Effective marine data management is a critical enabler for Digital Twins and decision support tools, which reuse marine data to simulate real-world conditions and model complex scenarios ([Lv et al., 2023](#); [McFerren et al., 2018](#); [Miedtank et al., 2024](#)). These models enhance ocean management by supporting informed decisions on fisheries, conservation, and coastal development ([Coro et al., 2020](#)). Digital Twins reduce costs and risks by testing strategies in virtual environments, modelling elements such as ocean currents ([Schneider et al., 2023](#)), ecosystem dynamics ([Chust et al., 2022](#)), and climate scenarios ([Brönnner et al., 2023](#)). Well-curated marine data also underpins machine learning, synthetic data generation, and AI-driven innovation, further strengthening sustainable ocean governance.

## QUALITY AND STANDARDS FRAMEWORKS IN MARINE DATA MANAGEMENT

Since 2017, the MI has strategically prioritised the adoption of quality processes and international standards to establish itself as a trusted source of marine data within Ireland and globally. This approach supports the effective management of complex, multi-disciplinary marine datasets, ensuring they are FAIR, and aligned with national and international best practices.

In practice, the tools put in place have supported the delivery of repeatable, transparent, and robust data processes, as demonstrated through Ireland's Marine Spatial Planning (MSP) processes. They have enabled the integration of stakeholder consultation feedback while maintaining the integrity of marine datasets used for policy, environmental protection, and resource planning ([Flynn et al., 2021](#); [Flynn et al., 2023](#)).

Internationally recognised certifications and quality frameworks build trust in marine data repositories by promoting consistency in data handling, from acquisition through to dissemination. They foster collaboration across sectors, including government, academia, and industry, and help to streamline workflows and improve decision-making based on reliable data. In the marine context, where decisions affect biodiversity, coastal communities, and national economic activities, such reliability is not just beneficial but essential.

It is important to acknowledge that implementing these standards is not a linear or finite task. Each adoption involves customisation to fit internal institutional processes, while still aligning

with externally mandated requirements (Bosch, 2016). At the MI, staff expertise has played a central role in tailoring frameworks to suit Ireland’s national marine data priorities, from deep-sea monitoring and oceanographic research to fisheries management and marine spatial planning. Moreover, the Institute recognises that framework implementation is a continuous process, requiring ongoing investment, evaluation, and flexibility to adapt to new technologies, shifting policy landscapes, and evolving international expectations.

### 3. METHODOLOGY

#### EVOLUTION OF FRAMEWORKS

The Marine Institute’s Data Strategy (2017) describes the commitment of the Marine Institute ‘To become an international centre of excellence in the provision of open access, high-quality data and information services that meet the needs of our clients and stakeholders, advance ocean knowledge, provide integrated advice and support blue growth.’ One of the objectives of the Data Strategy is to ensure that high quality marine data underpins the development of Ireland’s marine sectors with the strategic objective that the MI ‘...is a trusted source of high-quality data with well-defined reproducible processes’.

To address and achieve these ambitious goals the MI set about establishing a Data & IT Roadmap along with an Implementation Plan. Figure 1 demonstrates the timeline of events. The first step involved the establishment of a Data Management Framework: DM-QMF.

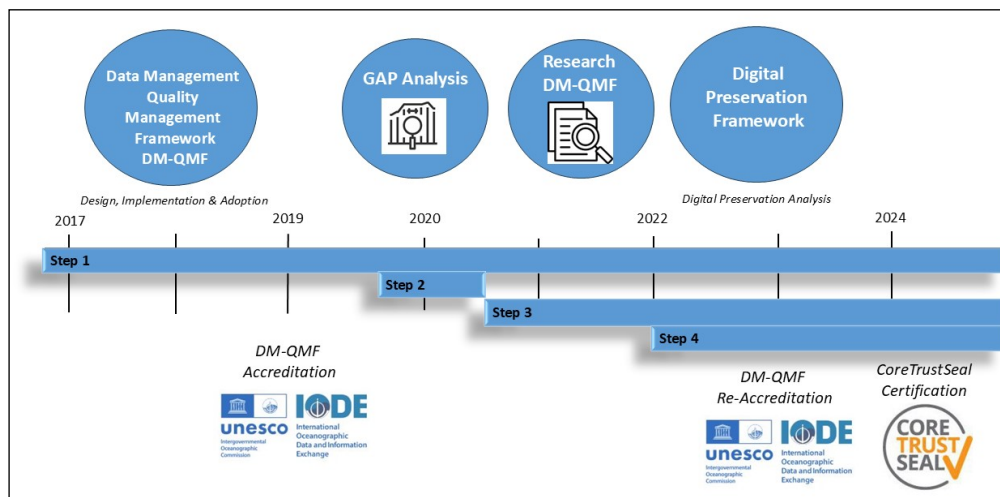


Figure 1 Methods and timelines of frameworks.

#### STEP 1 – DM-QMF – ACCREDITATION AND IMPLEMENTATION

In 2019, the Marine Institute in its remit as Ireland’s National Oceanographic Data Centre, received international accreditation of its Data Management Quality Management Framework by the Intergovernmental Oceanographic Commission of UNESCO’s International Oceanographic Data and Information Exchange programme (IOC-IODE) (IOC-IODE, 2019; Leadbetter et al., 2020a; Reed, 2023). The Data Management Quality Management Framework (DM-QMF) (Leadbetter et al., 2020a), aligned to ISO 9001, was developed to guide data management activities from collection, storage, quality control, and analysis to data outputs from providing persistent identifiers for datasets to legislative reporting. Since the initial accreditation, maintaining this recognition from the international community has been a priority, and there has been a focus on the further development of the Institute’s DM-QMF in its organisational strategy ‘Ocean Knowledge that Informs and Inspires’ (MI, 2023). The evolved DM-QMF has been identified as a key support for the wider digitalisation of scientific workflows and adoption and use of advanced data processes and analytical tools. Re-accreditation was awarded by IOC-IODE in 2023.

In the submission for IOC-IODE accreditation, a Data Management Quality Management model was developed along with a manual detailing the model and its implementation across the data producing areas of the Marine Institute (Leadbetter et al., 2020a). At a more practical level, an implementation pack (internally referred to as a ‘DM-QMF Pack’) consisting of a number of templates to assist in the compilation of the documentation required by the model and the manual was developed (Leadbetter et al., 2020a). Leadbetter et al. also reported that

establishing a network of Data Owners, Data Stewards and Data Coordinators was instrumental for the establishment and successful implementation; and was honest in that a high degree of coordination is in fact required including the freeing up of such resources from their daily tasks.

## STEP 2 – DM-QMF VS. ISO 9001 GAP ANALYSIS

In 2020 the MI conducted a thorough gap analysis between the full ISO 9001 standard and the MI’s DM-QMF (*internal document*) to ascertain the effectiveness of the newly established DM-QMF. The fundamental difference between the MI’s DM-QMF Accreditation and the ISO 9001:2015 Certification process comes down to full implementation and auditing. Whilst the MI DM-QMF is aligned to ISO 9001 there is no external auditing involved, this is limited to an internal ‘performance evaluation’ (Leadbetter et al., 2020a) and not all data processes in the Marine Institute are yet fully managed under the framework. This is encapsulated by the scope of the initial DM-QMF submission being in the initial phase limited to the organisation’s core operational work with research projects considered out of scope. With the DM-QMF the Marine Institute are making a statement that they are committed to working towards all processes under the framework over time.

The main distinction between accreditation and certification is around who and how it is done; certification is a third party ‘certifying’ a completion to an acceptable level; whereas accreditation is the assessment body first assessing and then allocating the accreditation if deemed appropriate. In this case, the assessment body for the DM-QMF was (UNESCO) International Oceanographic Commission’s IODE programme.

To conduct a thorough gap analysis successfully it was first important to understand the structure of the full ISO standard. For the purposes of the analysis, it was identified that there are 131 ‘shall’ statements written in the ISO 9001:2015 standard;

- ‘Shall’ indicates a requirement
- ‘Should’ indicates a recommendation
- ‘May’ indicates a permission
- ‘Can’ indicates a possibility or capability.

These ‘Shall’ statements expand to 238 line items, e.g. ‘The organisation shall do: a, b, c, d, f, and g’, meaning that throughout the ISO 9001:2015 standard there are 368 individual requirements. Each of these requirements can be audited and failure to demonstrate evidence sufficient to show how each are being delivered has the potential to result in a non-conformance. These non-conformances have limited timeframes in which they must be addressed and actioned upon to maintain accreditation.

The ISO standard employs a process approach which incorporates the Plan-Do-Check-Act (PCDA) cycle and risk-based thinking. Risk has always been part of ISO standards, however ISO 9001:2015 now looks to apply risk-based thinking to a variety of processes across planning, operations, and performance evaluation. Risks do not necessarily have to be negative, they can be seen as opportunities for increasing the effectiveness of individual processes, as well as the quality management system itself, by achieving improved results and preventing undesirable results.

To ascertain if a process is functioning under ISO 9001:2015 an Auditor looks for 3 key elements:

1. Is the process compliant?
2. Is the process effective?
3. Is the process continuously improving?

Therefore, this can be represented as over 1,000 (~1,104) different ways in which a process can fail over during an audit under ISO 9001:2015. Table 1 gives an indication of the scale of the requirements to achieve full accreditation.

CHAPTERS OF ISO 9001:2015	‘SHALL’ (LINE STATEMENTS)	COMPLIANCE	EFFECTIVENESS	RISK-BASED THINKING
Total	131 (368)	368	736	1,104

**Table 1** ISO requirement statements.

It was agreed that in order to ascertain how closely aligned the MI's DM QMF is to the ISO 9001:2015 was to conduct a thorough analysis of the framework against these 3 considerations: compliance, effectiveness, and risk-based thinking (also referred to as 'continuous improvements').

### **STEP 3 – RESEARCH DM-QMF**

From the learnings obtained both during the establishment and implementation of the DM-QMF and in the gap analysis, it was recommended to expand the scope of the datasets and processes covered by the DM-QMF. As noted above, the initial scope was limited to the organisation's core operational work and research projects were treated as out of scope. In the context of the DM-QMF 'operational' is deemed to refer to those processes that form part of the core and ongoing work that the Marine Institute is mandated and funded to deliver (e.g., long-term environmental monitoring, provision of data services).

As the Marine Institute is also a research performer and is actively involved in research projects in addition to delivering its core operational services ([eISB, 1991](#)). A research project has been defined in this context as a discrete scientific endeavour to answer a research question or a set of research questions with a defined time-period and resources. Involvement in research projects by the organisation is primarily targeted at enhancing or supporting the Marine Institute's operational capacity.

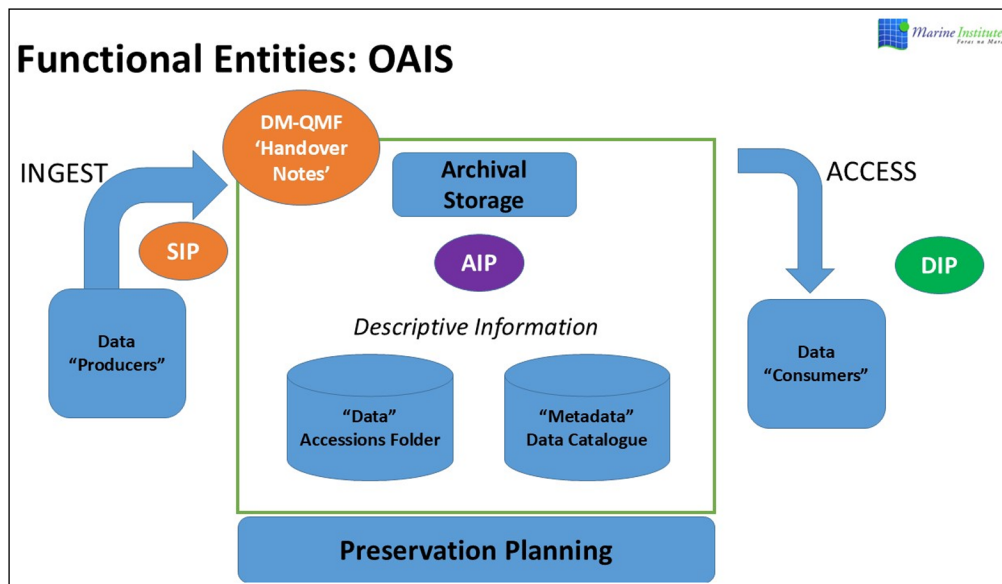
The Research DM-QMF was developed and is intended to document research project deliverables involving Marine Institute researchers from all funding sources, not limited to research funding provided through the Marine Institute itself.

Without managing research data outputs, there is a risk of reputational damage as data hungry initiatives like Marine Strategy Framework Directive (MSFD), the Marine Spatial Planning (MSP), and Climate Change Mitigation require publicly funded data to fulfil their goals. Having data stored 'somewhere' in an email or on a local, or even shared drive is unsuitable; similarly, not knowing which the definitive version is to retain is not sustainable. It is essential that there is sufficient supporting descriptive metadata and that the data are stored in open, optimised formats for long-term preservation. This becomes even more important when committing to ingesting datasets from an external organisation, as is the function of a NODC ([IODE, 2020b](#)). In the Marine Institutes remit as coordinating the NODC, it is committed to becoming a competent and confident data custodian, which led the Institute to consider complementary accreditations and/or certifications to demonstrate this capability.

### **STEP 4 – DIGITAL PRESERVATION FRAMEWORK**

In extending the DM-QMF remit to discrete datasets and data originating from third parties, such as research project data, the Marine Institute explored complementary frameworks to support and accredit or certify data management activities. Some of the datasets in question comprise data which has been previously collected and is being repurposed ([Gov.ie, 2021](#)) or data which are the result of a discrete scientific activity to answer research questions with a defined time period and resources. When managing discrete data, particularly from an external organisation, there is a greater requirement for appropriate documentation to ensure the data are available for reuse. The ability to recover undocumented knowledge and data outputs becomes increasingly difficult with time.

Following a full analysis of internal digital preservation processes, and an assessment of the suitability of the existing DM-QMF process, the outcome was to develop a Digital Preservation Framework and later to submit an application for CoreTrustSeal Certification (October 2022) in support of the Marine Institutes Digital Preservation work. CoreTrustSeal ([CoreTrustSeal Standards and Certification Board, 2020](#); [Dillo and Leeuw, 2018](#); [L'Hours et al., 2019](#)) was chosen to be a suitable framework complementary to the existing DM-QMF to manage and preserve data over time. All data archival processes follow the Open Archival Information System (OAIS) model (aligned to ISO 16363) ([The Consultative Committee for Space Data Systems \(CCSDS\), 2012](#)).



**Figure 2** Open Archive Information System (OAIS) model depicting how DM-QMF can become ingestion suite.

Figure 2 illustrates how documentation created as part of a Research DM-QMF can be further utilised as part of the CoreTrustSeal Ingestion Process.

The OAIS model (The Consultative Committee for Space Data Systems (CCSDS), 2012) consists of three information packages:

1. The **Submission Information Package** – SIP is the version of the information package that is transferred from the Producer to the NODC following OAIS and what is ingested into the Archive.
2. The **Archival Information Package** – AIP, the version of the information package stored and preserved, commences once the ingestion process has been completed.
3. The **Dissemination Information Package** – DIP is the version of the information package delivered to the Consumer in response to an access request.

The Marine Institute, in its remit as the National Oceanographic Data Centre (NODC) were Certified as a Trustworthy Data Repository by the CoreTrustSeal Standards and Certification Board until 19 September 2027 (CoreTrustSeal-AMT, 2024).

## ADDITIONAL FRAMEWORKS

There are a number of other frameworks in use at the MI including, ISO 27001, used to safeguard marine data through a structured Information Security Policy, ensuring data confidentiality, integrity, and availability. An ICT Risk Management framework, aligned with ISO 27005:2022, supports this by addressing legal, regulatory, and business requirements. Cyber threats are growing, with education and research now the second most targeted sector by nation-state actors (Microsoft, 2024). ISO 27001 enables MI to build strong cybersecurity systems covering strategic and operational needs.

The MI supports seafood industry testing and complies with EU legislation, maintaining ISO 17025 accreditation since 2002 to ensure technically valid results via independent audits.

## 4. RESULTS

### BENEFITS AND LEARNINGS FROM DM-QMF

The implementation within the operational activities of the Marine Institute have resulted in (at the time of writing) 135 datasets/processes now managed through the DM-QMF and a cumulative total of 304 formal meetings known as ‘performance evaluations’ to identify opportunities for improvement. The quality assurance of these data processes provides a solid foundation for the evidence base that underpins national marine decision making. While being very active in marine data acquisition, including through the national offshore research vessels, the Marine Institute has long acted as a broker for a range of data from partner organisations including research data. These data sets are important contributions to national marine programmes,

such as the National Marine Planning Framework, the Ocean Renewable Energy Development Plan, and for marine licensing services. [Table 2](#) lists the benefits, both formal and informal that have been provided from the implementation of the DM-QMF at the Marine Institute.

COMMUNICATION CHANNEL	BENEFIT
Formal*	Improves customer satisfaction: Delivering a more consistent service/product using well documented processes and procedures.
	Increases staff Engagement.
	Improves risk management.
	Aids reduction of waste.
	Provides direction for defining, improving, and controlling processes.
	Facilitates and helps identify training opportunities.
	Aids with setting organisation-wide direction.
	Communicates an ability and commitment to produce consistent high-quality results.
Informal^	Greater consistency in data products and services produced by the MI.
	Increased efficiency by improving time and resources, improved customer satisfaction.
	Consistency with all processes across the service areas.
	Continuous assessment and opportunities for improvement.
	Training materials for all staff – particular benefit to new starters.
	Reduced risk of undocumented processes and procedures.
	The identification of any sensitives associated with General Data Protection Regulations (GDPR) as well as the more recent Artificial Intelligence Act.
	Ensuring the quality of the management of data for use in Marine Spatial Planning, Marine Strategy Framework Directive, European Maritime Fisheries Fund, and other data-demanding legislative drivers into the future.
The use of data visualisation dashboards is gaining traction which offers further insights into the underlying data; from a Quality Control perspective but also added value opportunities.	

**Table 2** Benefits of the DM-QMF.

\*Formal benefits were captured as part of an OceanTeacher Global Academy Training Course that was developed and delivered by originators of the DM-QMF at the Marine Institute October 2020 ([UNESCO, 2002](#)).

^Informal communications with Data Stewards and Data Coordinators within the Marine Institute (*pers. comm.*).

## LEARNINGS FROM GAP ANALYSIS

Results demonstrated that from a ‘Compliance’ perspective the DM QMF Manual was clearly very well written covering all the necessary elements of the ISO 9001:2015 standards 10 chapters together with all mandatory records ‘documented information’ or ‘documented Procedures/Records’ that must be kept.

In relation to ‘effectiveness’, in 2020, it was shown that there was not sufficient documentation yet available to demonstrate the effectiveness of the processes beyond the apparent successful completion of the process outputs. Based on a test case examined it was not possible for experienced data professionals to step in and follow the processes without substantial time-consuming ‘piecing together’ of the individual components, accompanied by significant support from numerous MI staff involved (~10 hours of additional support were required). This meant that more detailed documentation in relation to the processes were necessary; this has been tested significantly more recently as staff changes have seen the documentation under the DM-QMF used as a suite for both starter notes and handover notes during staff changes.

Finally regarding ‘continuous improvement’ or ‘risk-based thinking’, in 2020, it was not evident that the test case had ever been analysed in detail; by identifying the potential to remove manual steps, implement standardised technologies or identify and/or remove any redundant components by streamlining and refining the processes. In fact, what it did show was despite the process being run very effectively each year, it did take approximately four weeks, involving 172 steps: 74 of which were manual and the use of 17 technologies (see Section In-house Test Case).

## BENEFITS OF OPTIMISED DIGITAL PRESERVATION

The MI’s Digital Preservation Framework provides a structure for policies and practices related to digital preservation. By developing this and achieving CTS certification, the benefits include [\(Table 3\)](#).

COMMUNICATION CHANNEL	BENEFIT
Formal*	Validates the quality and transparency of internal processes
	Increases awareness of and compliance with established standards
	Builds stakeholder confidence
	Enhances the reputation of the NODC
	Demonstrates that the NODC is following best practices
Informal^	Is a benchmark for comparison and helps to determine the strengths and weaknesses of the NODC
	Provides a framework for the policies and practices related to scientific data curation and preservation
	Provides assurance that data are preserved and can be reused in the future
	Helps plan data preservation sustainably
The IODE has confirmed reaccreditation of the MI's existing Data Management Quality Management framework (Reed, 2023, p. 10)^	

**Table 3** Benefits of a Digital Preservation Framework (CTS).

\*Formal benefits as depicted by CTS themselves.

^Informal communications (pers, comms).

^^ 'Any NODC or ADU that has been certified by CoreTrustSeal will be awarded the status of Accredited IODE National Oceanographic Data Centre or Accredited IODE Associate Data Unit provided they can show evidence of (i) providing national reports to the IODE Committee, and (ii) adherence to IODE Standards and Best Practice.'

## FAIR PRINCIPLES – WHERE ARE WE NOW?

The FAIR Principles (Wilkinson et al., 2016) were developed without explicit guidance on how to achieve the level of compliance or to ascertain a measurement of FAIRness. However, it is reported that if a repository is successful in achieving CoreTrustSeal Certification, it indicates the data preserved can be said to meet with elements of the FAIR Principles (Mokrane and Recker, 2019). A mapping between FAIR and CoreTrustSeal (Mokrane and Recker, 2019) was used as a guide by the NODC during the development of its Digital Preservation Framework and subsequent successful application for CoreTrustSeal certification (Table 4).

FAIR PRINCIPLES	NODC 'APPROACH'	CHECKS IN PLACE
<b>FINDABLE</b>	<ul style="list-style-type: none"> <li>The Data Catalogue publishes metadata for datasets and services in ISO 19115/19139 based XML format (in compliance with the INSPIRE directive) and provides comprehensive metadata for each dataset.</li> <li>Datasets are indexed with consistent metadata that makes them easily searchable.</li> <li>A dissemination checklist examines the process by which each step of a Consumer request for data is managed within the NODC. The Data Catalogue contains a record of all datasets contained within the NODC along with the relevant information as captured within the Service Agreement, including citation details.</li> <li>The Data Catalogue assigns a unique identifier to each dataset of the form 'ie.marine.data:dataset.{dataset_id}' so the dataset can be tracked in downstream aggregators.</li> <li>A digital object identifier through a contract with DataCite may be applied to datasets via the Data Catalogue. These unique IDs are used to provide cross-linkages in the metadata where appropriate.</li> <li>Metadata records can also be harvested from the Data Catalogue to other metadata catalogues, portals and aggregators, facilitating better discoverability.</li> </ul>	<p>The DM-QMF (aligned to <b>ISO 9001</b>) requires the population of a Data Catalogue record as part of the implementation pack.</p> <p>The NODC follows the best practice guidance as described in the Draft Data Preservation Policy. This includes the <b>ISO 14721 (OAIS Model)</b> for storage as well as the other preservation functions.</p>
<b>ACCESSIBLE</b>	<ul style="list-style-type: none"> <li>Where available, the metadata provided by the Marine Institute data catalogue includes access links to the datasets and data services. The catalogue provides datasets in machine-readable formats (e.g., CSV, NetCDF, or GeoTIFF).</li> <li>The MI Data Catalogue meets domain-specific, community-defined standards, and legislative requirements placed on data publishers. Schemas and standards include ISO 19115/19139 based XML, Datacite and Observations and Measurements (O&amp;M).</li> <li>The XML is harvested into the MI's Data Catalogue, which provides an OGC Catalogue Service for the Web endpoint, allows machine harvesting to other portals including the Irish Spatial Data Exchange (ISDE), and to the Irish Open Data Portal.</li> <li>Data is accessed via defined protocols to ensure data is accessed reliably and securely. Many datasets contain URLs for dataset downloads and other RESTful URLs.</li> </ul>	<p>A Data Catalogue has been developed to export metadata for datasets and services in <b>ISO 19115/19139</b> based XML format in compliance with the INSPIRE implementing rules for metadata (Leadbetter et al., 2020b).</p>

**Table 4** Mapping between FAIR, CTS requirements and how the MI has met these requirements.

FAIR PRINCIPLES	NODC 'APPROACH'	CHECKS IN PLACE
<b>INTEROPERABLE</b>	<ul style="list-style-type: none"> <li>The use of standardised formats, and vocabularies support data integration and interoperability. The Data Catalogue uses well-managed, community governed controlled vocabularies, including the SeaDataNet Vocabularies and British Oceanographic Data Centre (BODC) Vocabularies. Detailed description of parameters within datasets enables mapping to BODC Parameter Usage Vocabulary (PUV) and its semantic model. The BODC PUV is a controlled vocabulary for labelling variables in databases and data files in oceanography and related domains; a collection of unique and persistent identifiers attached to structurally logical labels and textual definitions.</li> <li>These controlled vocabularies are hosted on vocabulary servers which also enable FAIR principles using web semantic description, and are reachable through SPARQL endpoints</li> <li>Data citations are provided for datasets.</li> </ul>	<p>OGC CSW (Catalog Service for the Web) which aligns with <b>ISO 19128</b>: supports metadata retrieval using <b>ISO 19115/ISO 19139</b> standards, ensuring compatibility with standardised metadata schemas</p>
<b>REUSABLE</b>	<ul style="list-style-type: none"> <li>Records in the Data Catalogue provide sufficient detail and metadata to ensure reproducibility and future use</li> <li>Lineage Statements: Information and references for sources and processes that were used to create the dataset and an audit trail for modifications to the original data. ISO 19115:2014, meeting the FAIR principles on re-use's requirement around dataset provenance</li> <li>Licence Conditions: providing official permission to do, use or own something. Data licenses exist on a spectrum from entirely open to very restricted. As per MI Data Policy (Workflow) e.g., CC-BY 4.0, meeting the FAIR principle on re-use's requirement around advertising a data license and where possible using an open license</li> <li>Documentation can be included to provide the context of how data was generated and should be interpreted using a Supplementary Information field and linking with the Marine Institute's Open Access Repository (OAR) where applicable. The OAR collects, preserves and provides open access to the publications of the Marine Institute</li> <li>Under the Digital Preservation Framework, certified under CoreTrustSeal, the criteria for data integrity and requirements for a preservation plan ensures that data is curated with long-term reuse in mind.</li> </ul>	<p>A Data Catalogue has been developed to export metadata for datasets and services in <b>ISO 19115/19139</b> based XML format in compliance with the INSPIRE implementing rules for metadata (<a href="#">Leadbetter et al., 2020b</a>).</p> <p>The Digital Preservation Framework, Certified under CoreTrustSeal, is aligned to <b>ISO 16363</b> which is a standard for Trust Digital Repositories.</p>

The experience of the MI as the NODC in addressing some of the CoreTrustSeal requirements has shown the importance of managing the reuse context over time. It is well understood that data reuse must be in accordance with the context in which data was originally collected.

### IN-HOUSE TEST CASE

Since the gap analysis learnings in 2020, the PDCA process of improvement has resulted in several efficiencies for a scientific operational process when processing one years' worth of data. Historically it took ~four weeks – involving 172 steps – 74 of which were manual – using 17 technologies to achieve. This now takes ~three days where the Team are fully trained to competently run the process themselves. This was achieved largely down to a more simplified process involving a Jupyter Notebook that embraces the ethos of a 'human in the loop'; meaning staff are more confident that they understand and are in control as the process progresses. This would not have been possible if the process was not fully documented (using the DM-QMF), a performance evaluation conducted, the actions presented clearly to senior leadership and a coordinated and vested effort to bring in the appropriate skillset to invest in the efficiencies. More importantly, none of these efficiencies would have been possible if it was not for the commitment to improvement from the scientists themselves – to invest in the time and effort to challenge their existing practices and work collectively – learning new skills and approaches to arrive at the outcome achieved.

As the recently updated Digital Preservation processes, now certified under CoreTrustSeal, are still in its infancy, it is not yet possible to generate similar statistics in relation to the tangible impacts, but it is expected to yield similar efficiencies.

Reporting on progress of the uptake of the DM-QMF is important in terms of re-accreditation of the DM-QMF by UNESCO IOC-IODE but also more locally for reporting to the Marine Institute's senior leadership and board (Leadbetter et al., 2020a). Key Performance Indicators (KPIs) are presented every quarter, which facilitates the board of directors to stay informed about the functioning of the DM-QMF, and to provide guidance and feedback as necessary, to ensure the achievement of long-term goals.

As the DM-QMF has been in operation at the MI for some time numerous supporting processes have evolved. One example includes the establishment of 'inventories' from each service area ( $n = 6$ ) pertaining to the individual processes being actively managed under the DM-QMF. These inventories were initially developed in Excel capturing information such as status (complete, in progress or not started), target dates for completion, and information pertaining to the individual status of DM-QMF components (e.g. Data Management Plans). They also recorded the number of performance evaluations conducted per process. Over time divergence between spreadsheets occurred resulting in it becoming time-consuming to collate statistics for a reporting cycle with many manual steps in the maintenance of the inventory and validating the statistics drawn out of the inventories. It subsequently resulted in repetition to generate information for a range of audiences (e.g. KPIs, service area meetings, DM-QMF quarterly newsletter). A structured database now provides a consistent approach to managing the DM-QMF pack inventory and in conjunction with a MS PowerBI report gives a mechanism to interrogate the database for KPIs and other measures of progress. This has resulted in optimised reporting of performance indicator statistics regarding the uptake of the DM-QMF in real time to support the operationalisation of reporting.

## 5. DISCUSSION

### WHAT HAS BEEN IMPROVED?

#### Coordinated management 'PDCA'

Since the outcome of the gap analysis in 2020, further effort was placed upon addressing the two elements that required more attention. As a direct result of a fuller implementation of the 'performance evaluation' component of the DM-QMF (Leadbetter et al., 2020a) a live register known as an 'Actions Log' has been developed to capture opportunities for improvement. This collective effort across the Marine Institute has seen the development of a 'ticketing system' being developed and a series of questions being asked to ensure each opportunity is clearly captured and evaluated, such as timelines, budget and resource availability. These Logs are now reviewed monthly by senior management to facilitate optimised coordination and planning of resources to coincide with the most pragmatic opportunities for improvements.

Having the DM-QMF has clearly demonstrated a sharing of expertise so that a process is always run sequentially, and several capable individuals are able to use the DM-QMF documentation to ensure no process remains incomplete. Previously this was not always obvious until someone left the Institute.

#### Management of updated or new data requirements/legislation

As well as enabling each process to capture categorically any association with GDPR, the DM-QMF has meant highlighting any changes to existing or new relevant legislations e.g. the Artificial Intelligence Act 2024 (EU Artificial Intelligence (AI) Act, 2024); new categories can be added very quickly highlighting very swiftly to senior management any action that needs to be taken. It is worth noting that this does not push the responsibility onto the DM-QMF framework to mitigate these elements; but merely to highlight the need for any new mitigation measures that might need to be put in place.

#### Relative rigour of the standards

One of the key differences between the quality frameworks in place at the Marine Institute and full certification to ISO 9001:2015 standards is the level of external scrutiny that is placed on the quality system. For example, the National Standards Authority of Ireland (NSAI) is accredited to certify that organisations meet the ISO 9001:2015 standard (NSAI, 2025). They do this through

a multi-stage process compared to the IODE accreditation process (IODE, 2020a), which is relatively short.

Both the IODE and CoreTrustSeal processes rely on self-assessment and expert review of documentation provided by the data centre, whilst the ISO 9001:2015 certification process involves on-site visits and assessments by external auditors. These on-site assessments are likely to be more rigorous than self-assessment and can highlight any differences between what should be done and what actually is done. For example, whilst a policy might say that a particular process must be followed for each data set submitted it might be the case that in reality this does not always happen. This non-conformity might go un-detected if only the policy is reviewed. Whilst self-assessment and internal audits are part of the ISO 9001:2015 standard the extra level of scrutiny provided during the external audit process helps to ensure the quality management system is rigorously evaluated. The consequences of failing an audit and potentially losing certification might also enable more resources to be devoted to ensuring the quality management system is running effectively.

That said, whilst alignment often means there are no formal certification audits it does not preclude internal gap analysis, assessments or audits using tools which are publicly available.

### Level of effort

The best way to evaluate the benefits of implementing a quality management system is to define key performance indicators based on desired outcomes and ensure measurement is started before the implementation process. This will provide a baseline to measure any improvements against. This is defined in the standard under section 9.1 'Monitoring, measurement, analysis, and evaluation'.

The direct costs of implementing a quality management system should also be recorded – these could include items such as staff time, and external consultancy fees.

During critical analysis of the most pragmatic way forward, adapting the data frameworks to suit the specific infrastructure within the Marine Institute, was also felt to be very challenging. A similar summary was felt with the ISO 17025 accreditation for the laboratories (*pers. comms*). Feedback suggested the level of effort required is continuous. Required activities include annual auditing of all requirements of the standard; additional risk-based audits; non-conformance resolution; continual improvements; change control; extending scope; and the addition of new tests. There is also the need for continuing professional development for all staff since ongoing competency is required and quality systems demand dedicated personnel working within them.

Improving data quality through metadata harmonisation and incorporating digital preservation best practices is essential for long-term data usability. Metadata harmonisation ensures that data from different sources are described in a standardised way, facilitating interoperability, integration, and cross-disciplinary analysis. By aligning metadata fields – such as data collection methods, formats, units, and spatial or temporal references – data consumers can more easily compare and merge datasets, enhancing the accuracy and scope of insights. Additionally, adopting digital preservation best practices, for example, using open, non-proprietary formats, regularly migrating data to updated platforms, and maintaining secure, accessible archives, protects data from obsolescence and loss. These practices ensure that data remains usable and trustworthy over time, making it accessible for future research, analysis, and decision-making. Metadata harmonisation and digital preservation create a foundation for robust data management, extending the lifespan and reusability of valuable datasets.

Leadbetter (Leadbetter et al., 2020a) wrote that it was important to understand that whilst the establishment of the DM-QMF assures the quality of the data management process, it does not address the quality of the underlying data itself. It was hoped that by demonstrating the quality assurance of the data management process that inevitably the data quality would also be improved over time. This continues to be a work in progress through the Plan-Do-Check-Act (PDCA) cycle (Sari et al., 2017) and processes are becoming more efficient and effective. One consequence of these process optimisations is giving the data producers time back to invest in scientific quality control and incorporate this into their process.

It has been well documented that optimised data management processes are multifaceted and complex (Azami et al., 2023); this has indeed been the experience of the Marine Institute

that it requires a systematic approach encompassing a wide variety of stakeholders. Expansion of the initial Operational DM-QMF to a Research DM-QMF has indeed addressed a gap and given the Institute confidence in its data management capabilities. Furthermore, establishing a Digital Preservation Framework and obtaining CoreTrustSeal Certification demonstrates the Institute's ability to manage third party data preservation in a transparent manner.

However, while these frameworks offer structured guidance, the true driving force behind success lies in the people involved. Having a team of properly skilled, motivated, and engaged staff is a key factor in elevating an organisation's data management practices. Those working in this space need not only technical expertise but also a deep understanding of both the regulatory environment and the specific challenges inherent in marine data. This extends to the need for respect for domain expertise from data producers. It is critical that these professionals are actively supported by senior management, ensuring they have the resources, training, and backing necessary to excel in their roles. Furthermore, a willingness to adhere to relevant legislation and meet the requirements of funding bodies is essential for maintaining operational integrity. Organisations must be vigilant in staying up to date with new and evolving laws and guidelines, always ensuring compliance.

Equally important is an organisational culture that embraces innovation; constantly scanning the horizon for new and emerging technologies, methodologies, and best practices that can transform basic data management practices into highly effective, resilient, and truly robust ones. By fostering a proactive mindset toward change and continuous improvement, marine data management can not only meet but exceed industry expectations, ensuring long-term sustainability and success.

In the context of marine data management, Social Exchange Theory ([Homans, 1958](#); [Wilms et al., 2020](#)), Prospect Theory ([Kahneman and Tversky, 1979](#); [Wilms et al., 2020](#)), and Nudge Theory ([Leonard, 2008](#)) can potentially offer acuity into improving collaboration, decision-making, and compliance. After all one of the most exciting prospects in the reuse of scientific data is in the new insights it offers. The need to highlight the importance of fostering reciprocal relationships, where organisations share data by recognising the long-term benefits, such as access to larger or more diverse datasets. Using frameworks to guide decision-making by offering new technologies or compliance with data standards (like ISO 27001) as potential gains or as a means to avoid significant losses, such as data breaches or missed opportunities. Encouraging compliance and enhancing data quality without imposing strict mandates by simply making the process possible; fostering a culture of continual improvement with respect to marine data management. By the utilisation of these data frameworks, it is possible to create an environment that promotes cooperation, informed decision-making, and streamlined adherence to regulations, ultimately driving better marine data management outcomes.

## 6. CONCLUSION

Marine data management is a specialised discipline that integrates people, processes, technologies, and methodologies to effectively collect, store, analyse, and disseminate oceanographic and marine-related data. It involves scientists, data managers, and IT specialists who work together to ensure data accuracy, consistency, and accessibility. Processes like data collection, validation, and standardisation are critical, while technologies such as sensors, remote sensing tools, and data storage platforms play a key role. Additionally, methodologies for data sharing, curation, and long-term preservation ensure that marine data can be reliably used for research, environmental monitoring, and policy-making, supporting sustainable management of marine ecosystems. The use of data frameworks goes a long way to demonstrate a competency in marine data management; continuously exploring opportunities for improvements, horizon scanning and investing in people are crucial to navigate all that must be achieved.

The Marine Institute has laid the groundwork for a strong data culture and remains committed to its ongoing evolution. With continued leadership support, staff engagement, and a willingness to embrace change across all levels, marine data management is well positioned for the future. By sharing our experience, we hope to inspire others to contribute their own insights.

In essence, whilst data frameworks offer structure, it is a culture of skilled, empowered, supported, and future-focused people – underpinned by clear communication, visible value, and continuous learning – that elevates marine data management from merely good to great.

The authors wish to thank the staff at the Marine Institute for their ongoing commitment to this work.



## FUNDING INFORMATION

This article is supported by projects funded under The Marine Knowledge Schemes which is established under Priority 4 (Strengthen Ocean Governance) of Ireland's Operational Programme (OP) and is co-funded by the Irish Government and EMFAF 2021-2027.

## COMPETING INTERESTS

The authors have no competing interests to declare.

## AUTHOR AFFILIATIONS

**Sarah Flynn**  [orcid.org/0000-0002-9333-8187](https://orcid.org/0000-0002-9333-8187)

Marine Institute, Oranmore, Galway, Ireland

**Tara Keena**

Marine Institute, Oranmore, Galway, Ireland

**Yvonne Bogan**

Marine Institute, Oranmore, Galway, Ireland

**Laura Brophy**

Marine Institute, Oranmore, Galway, Ireland

**David Currie**

Marine Institute, Oranmore, Galway, Ireland

**Adam Leadbetter**

Decade Coordination Office for Ocean Data Sharing of the Intergovernmental Oceanographic Commission of the United Nations Educational, Scientific and Cultural Organization (IOC-UNESCO), Seascope Belgium Bvba, Belgium

**Martina Maloney**

Marine Institute, Oranmore, Galway, Ireland

**Keith Manson**

Marine Institute, Oranmore, Galway, Ireland

**Colin Melville**

Marine Institute, Oranmore, Galway, Ireland

**Eoin O'Grady**

Marine Institute, Oranmore, Galway, Ireland

**Rob Thomas**

Marine Institute, Oranmore, Galway, Ireland

**Brendan Whittle**

Marine Institute, Oranmore, Galway, Ireland

**Andrew Conway**

Marine Institute, Oranmore, Galway, Ireland

## REFERENCES

- Ali, A., Fathalla, A., Salah, A., Bekhit, M. and Eldesouky, E.** (2021) 'Marine data prediction: An evaluation of machine learning, deep learning, and statistical predictive models', *Computational Intelligence and Neuroscience*, p. 8551167. Available at: <https://doi.org/10.1155/2021/8551167>
- Alvarez Fanjul, E., Ciliberti, S. and Bauharel, P.** (2022) *Implementing operational ocean monitoring and forecasting systems (GOOS-275)*. 1st ed, IOC-UNESCO publication.
- Azami, M., Sadatmoosavi, A. and Chashmyazdan, M.** (2023) 'Research data management frameworks: A systematic literature review', *International Journal of Information Science and Management*, 21, pp. 1–18. Available at: <https://doi.org/10.22034/ijism.2023.1977759.0>

- Bahurel, P., Adragna, F., Bell, M.J., Jacq, F., Johannessen, J.A., Traon, P.-Y.L., Pinardi, N. and She, J.** (2010) 'Ocean Monitoring and Forecasting Core Services, the European MyOcean Example', in: *A Framework for Ocean Observing*. Presented at the OceanObs'09, Venice Convention Centre, Venice-Lido, Italy. Available at: <https://api.semanticscholar.org/CorpusID:131175322>
- Biber, E.** (2013) 'The challenge of collecting and using environmental monitoring data', *Ecology and Society*, 18(4). Available at: <https://doi.org/10.5751/ES-06117-180468>
- Bosch, J.** (2016) 'Speed, data, and ecosystems: excelling in a software-driven world', 1st ed. in *Chapman & Hall/CRC Innovations in Software Engineering and Software Development Series*. Routledge.
- Brönnner, U., Sonnewald, M. and Visbeck, M.** (2023) 'Digital twins of the ocean can foster a sustainable blue economy in a protected marine environment', *IHR*, 29, pp. 26–40. Available at: <https://doi.org/10.58440/ihr-29-a04>
- Calewaert, J.B., Sierra-Correa, P.C., de Boer, G., Crosman, K., Kågesten, G., Taconet, M., Busumprah, P.T., Alvarez-Fanjul, E., Shepherd, I., Schaap, D., Hall, S., Zinkann, A.C., Haddad, T., Jegat, V., Smit, M., Sun, M., Suzuki, T., Visbeck, M., Xingyuan, R. and McMeel, O.** (2024) *Ocean Decade Vision 2030 White Papers: Challenge 8: Create a digital representation of the ocean* [Ocean Decade Series, 51.8 White Paper], Paris.
- Carolyn, S.** (2023) *Preliminary survey of the information preservation needs of the marine conservation community*. Available at: <https://shareok.org/handle/11244/337462>
- Cerrillo-Márquez, A. and Casadesús-de-Mingo, A.** (2021) 'Data governance for public transparency. Profesional de la información', *Profesional de la información*, 30(4), p. E300402. Available at: <https://doi.org/10.3145/epi.2021.jul.02>
- Chust, G., Corrales, X., Gonzáles, F., Villarino, E., Chifflet, M., Fernandes, J.A., Ibaibarriaga, L., Borja, Á., Arandu, M., Andonegui, E., González, M., Irigoien, X. and García, D.** (2022) *Marine biodiversity modelling study* (Report). Publications Office of the European Union. Available at: <https://doi.org/10.2777/213731>
- Copernicus Marine Service (CMEMS)** (2015) *Home*. Available at: <https://marine.copernicus.eu/> (Accessed: 1 May 25).
- CoreTrustSeal-AMT** [WWW Document] (2024) *Certified Data Repositories*. Available at: <https://amt.coretrustseal.org/certificates> (Accessed: 28 April 25).
- CoreTrustSeal Standards and Certification Board** (2020) Change file: CoreTrustSeal trustworthy data repositories requirements: Extended guidance 2017–2019 to 2020–2022. *Zenodo*. Available at: <https://doi.org/10.5281/zenodo.3828636>
- Coro, G., Pagano, P. and Ellenbroek, A.** (2020) 'Detecting patterns of climate change in long-term forecasts of marine environmental parameters', *International Journal of Digital Earth*, 13, pp. 567–585. Available at: <https://doi.org/10.1080/17538947.2018.1543365>
- Cousijn, H., Habermann, T., Krzmarich, E. and Meadows, A.** (2022) 'Beyond data: Sharing related research outputs to make data reusable', *Learned Publishing*, 35(1), pp. 1–87. Available at: <https://doi.org/10.1002/leap.1429>
- Coussement, K. and Benoit, D.F.** (2021) 'Interpretable data science for decision making. Decision Support Systems', *Interpretable Data Science for Decision Making*, 150, p. 113664. Available at: <https://doi.org/10.1016/j.dss.2021.113664>
- Dillo, I. and de Leeuw, L.** (2018) 'Lisa de CoreTrustSeal', *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen & Bibliothekare*, 71(1), pp. 162–170. <http://eprints.rclis.org/34431/> (Accessed: 10 March 24).
- Electronic Irish Statute Book (eISB)** (1991) *Marine Institute Act, 1991*. Government of Ireland. <https://www.irishstatutebook.ie/eli/1991/act/2/enacted/en/html> (Accessed: 10 July 24).
- ERDDAP** (2015) *Home Page*. <https://erddap.marine.ie/erddap/index.html> (Accessed: 14 April 25).
- EU Artificial Intelligence (AI) Act** (2024) <https://enterprise.gov.ie/en/what-we-do/innovation-research-development/artificial-intelligence/eu-ai-act/eu-ai-act.html> (Accessed 22 November 24).
- European Marine Observation and Data Network (EMODnet)** (2009) *Home*. <https://emodnet.ec.europa.eu/en> (Accessed: 1 May 25).
- Flynn, N., Frost, D., O'Leary, J. and Hynes, S.** (2024) *Ireland's Ocean Economy, 2024*. Marine Institute.
- Flynn, S., Meaney, W., Leadbetter, A.M., Fisher, J.P. and Aonghusa, C.N.** (2021) 'Lessons from a Marine Spatial Planning data management process for Ireland', *International Journal of Digital Earth*, 14, pp. 139–157. Available at: <https://doi.org/10.1080/17538947.2020.1808720>
- Flynn, S., Tray, E., Woolley, T., Leadbetter, A., Heney, K., O'Driscoll, D., Nic Aonghusa, C. and Conway, A.** (2023) 'Management of spatial data integrity including stakeholder feedback in Maritime Spatial Planning', *Marine Policy*, 156, p. 105799. Available at: <https://doi.org/10.1016/j.marpol.2023.105799>
- Fogarty, M.J., Gamble, R. and Perretti, C.T.** (2016) 'Dynamic complexity in exploited marine ecosystems', *Front. Ecol. Evol.*, 4. Available at: <https://doi.org/10.3389/fevo.2016.00068>
- Gov.ie** (2021) *National Marine Planning Framework*. <https://www.gov.ie/en/publication/60e57-national-marine-planning-framework/> (Accessed: 22 November 24).

- Han, Y., Yang, T., Yuan, M., Hu, P. and Li, C.** (2025) 'Construction of a maritime knowledge graph using GraphRAG for entity and relationship extraction from maritime documents', *Journal of Computer and Communications*, 13, pp. 68–93. Available at: <https://doi.org/10.4236/jcc.2025.132006>
- Homans, G.C.** (1958) 'Social behavior as exchange', *American Journal of Sociology*, 63, pp. 597–606. Available at: <https://doi.org/10.1086/222355>
- IOC-IODE** (2019) *IODE quality management framework for national oceanographic data centres and associate data units*. <https://unesdoc.unesco.org/ark:/48223/pf0000371181> (Accessed: 8 October 24).
- IODE** (2020a) *How to get accreditation for your NODC?* <https://iode.org/faq-items/how-to-get-accreditation-for-your-nodc/> (Accessed: 28 April 25).
- IODE** (2020b) *The mission of a National Oceanographic Data Centre (NODC)*. <https://iode.org/faq-items/the-mission-of-a-national-oceanographic-data-centre/> (Accessed: 22 November 24).
- Kahneman, D. and Tversky, A.** (1979) 'Prospect theory: An analysis of decision under risk', *Econometrica*, 47, pp. 263–291. Available at: <https://doi.org/10.2307/1914185>
- Kapiszweski, D. and Karcher, S.** (2020) *Making research data accessible*. Moynihan Institute of Global Affairs. Available at: <https://doi.org/10.31235/osf.io/9kgj4>
- Koedel, U., Schuetze, C., Fischer, P., Bussmann, I., Sauer, P.K., Nixdorf, E., Kalbacher, T., Wichert, V., Rechid, D., Bouwer, L.M. and Dietrich, P.** (2022) 'Challenges in the evaluation of observational data trustworthiness from a data producers viewpoint (FAIR+)', *Front. Environ. Sci.*, 9. Available at: <https://doi.org/10.3389/fenvs.2021.772666>
- Leadbetter, A., Carr, R., Flynn, S., Meaney, W., Moran, S., Bogan, Y., Brophy, L., Lyons, K., Stokes, D. and Thomas, R.** (2020a) 'Implementation of a data management quality management framework at the Marine Institute, Ireland', *Earth Science Informatics*, 13, pp. 509–521. Available at: <https://doi.org/10.1007/s12145-019-00432-w>
- Leadbetter, A., Meaney, W., Tray, E., Conway, A., Flynn, S., Keena, T., Kelly, C. and Thomas, R.** (2020b) 'A modular approach to cataloguing marine science data', *Earth Science Informatics*, 13, pp. 537–553. Available at: <https://doi.org/10.1007/s12145-020-00445-w>
- Leadbetter, A.M., Lowry, R.K. and Clements, D.O.** (2014) 'Putting meaning into NETMAR – The open service network for marine environmental data', *International Journal of Digital Earth*, 7, pp. 811–828. Available at: <https://doi.org/10.1080/17538947.2013.781243>
- Leonard, T.C., Thaler, R.H. and Sunstein, C.R.** (2008) 'Nudge: Improving decisions about health, wealth, and happiness', *Const Polit Econ*, 19, pp. 356–360. Available at: <https://doi.org/10.1007/s10602-008-9056-2>
- L'Hours, H., Kleemola, M. and de Leeuw, L.** (2019) 'CoreTrustSeal: From academic collaboration to sustainable services', *IASSIST Quarterly*, 43, pp. 1–17. Available at: <https://doi.org/10.29173/iq936>
- Liu, D. and Cheng, L.** (2024) 'MAKG: A maritime accident knowledge graph for intelligent accident analysis and management', *Ocean Engineering*, 312, p. 119280. Available at: <https://doi.org/10.1016/j.oceaneng.2024.119280>
- Lv, Z., Lv, H. and Fridenfalk, M.** (2023) 'Digital twins in the marine industry', *Electronics*, 12, p. 2025. Available at: <https://doi.org/10.3390/electronics12092025>
- Marcucci, S., Alarcón, N.G., Verhulst, S.G. and Wüllhorst, E.** (2023) 'Informing the global data future: Benchmarking data governance frameworks', *Data & Policy*, 5, p. e30. Available at: <https://doi.org/10.1017/dap.2023.24>
- Marine Institute** (2010) *Ireland's Marine Atlas*. <https://atlas.marine.ie/#?c=53.9043;-15.8862;6> (Accessed: 14 April 25).
- Marine Institute** (2023) *Ocean knowledge that informs and inspires, corporate strategy 2023–2027* (Technical Report). Marine Institute.
- McFerren, G., Molapo, R. and McAlister, B.** (2018) 'Repeatable deployment of an open standards, open source and open data stack for building a federated marine data management and decision support system for South Africa', in *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLII-4-W8, pp. 139–145. Available at: <https://doi.org/10.5194/isprs-archives-XLII-4-W8-139-2018>
- Microsoft** (2024) *Microsoft Digital Defense Report 2024*. <https://www.microsoft.com/en-us/security/security-insider/intelligence-reports/microsoft-digital-defense-report-2024> (Accessed: 28 April 25).
- Miedtank, A., Schneider, J., Manss, C. and Zielinski, O.** (2024) 'Marine digital twins for enhanced ocean understanding', *Remote Sensing Applications: Society and Environment*, 36, p. 101268. Available at: <https://doi.org/10.1016/j.rsase.2024.101268>
- Mokrane, M. and Recker, J.** (2019) Enabling Findable, Accessible, Interoperable, and Reusable (FAIR) Data. Presented at the *iPRES 2019 – 16th International Conference on Digital Preservation*, Amsterdam, The Netherlands.
- NSAI** (2025) *ISO 9001 Quality Management*. <https://www.n sai.ie/certification/management-systems/iso-9001-quality-management/> (Accessed: 28 April 25).
- Owens, D., Abeyirigunawardena, D., Biffard, B., Chen, Y., Conley, P., Jenkyns, R., Kerschtiens, S., Lavallee, T., MacArthur, M., Mousseau, J., Old, K., Paulson, M., Pirenne, B., Scherwath, M. and**

- Thorne, M.** (2022) 'The oceans 2.0/3.0 data management and archival system', *Front. Mar. Sci.*, 9. Available at: <https://doi.org/10.3389/fmars.2022.806452>
- Pasquetto, I.V., Cullen, Z., Thomer, A. and Wofford, M.** (2024) 'What is research data "misuse"? And how can it be prevented or mitigated?' *Journal of the Association for Information Science and Technology*. Available at: <https://doi.org/10.1002/asi.24944>
- Reed, G.** (2023) *IODE quality management framework for National Oceanographic Data Centres and associate data units* (2nd Revised edition) (Reference Document). Available at: <https://oceanexpert.org/document/33321>
- Richards, J.D., Niven, K. and Jeffrey, S.** (2013) 'Preserving our digital heritage: information systems for data management and preservation', in: E. Ch'ng, V. Goffney, H. Chapman (Eds.). *Visual Heritage in the Digital Age*. Springer, London, pp. 311–326. Available at: [https://doi.org/10.1007/978-1-4471-5535-5\\_16](https://doi.org/10.1007/978-1-4471-5535-5_16)
- Sari, Y., Wibisono, E., Wahyudi, R.D. and Lio, Y.** (2017) 'From ISO 9001:2008 to ISO 9001:2015: Significant changes and their impacts to aspiring organizations', *IOP Conference Series: Materials, Science and Engineering 2007*, 273, p. 012021. Available at: <https://doi.org/10.1088/1757-899X/245/1/012021>
- Schaap, D.M.A. and Lowry, R.K.** (2010) 'SeaDataNet – Pan-European infrastructure for marine and ocean data management: Unified access to distributed data sets', *International Journal of Digital Earth*, 3, pp. 50–69. Available at: <https://doi.org/10.1080/17538941003660974>
- Schneider, J., Klüner, A. and Zielinski, O.** (2023) 'Towards digital twins of the oceans: The potential of machine learning for monitoring the impacts of offshore wind farms on marine environments', *Sensors*, 23, p. 4581. Available at: <https://doi.org/10.3390/s23104581>
- Shucksmith, R.J. and Kelly, C.** (2014) 'Data collection and mapping – Principles, processes and application in marine spatial planning', *Marine Policy*, 50, pp. 27–33. Available at: <https://doi.org/10.1016/j.marpol.2014.05.006>
- The Consultative Committee for Space Data Systems (CCSDS)** (2012) *Reference Model for an open archival information system* (OAIS). (No. Magenta Book). Available at: <https://public.ccsds.org/pubs/650x0m2.pdf>
- Tray, E., Leadbetter, A., Meaney, W., Conway, A., Kelly, C., Maoláidigh, N.Ó., de Eyto, E., Moran, S. and Brophy, D.** (2020) 'An open-source database model and collections management system for fish scale and otolith archives', *Ecological Informatics*, 59, p. 101115. Available at: <https://doi.org/10.1016/j.ecoinf.2020.101115>
- UNESCO** (2002) *OceanTeacher Global Academy*. <https://classroom.oceanteacher.org/> (Accessed: 1 May 25).
- Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Santos, L.B. da S., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., Hoen, P.A.C. 't., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., Schaik, R. van, Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., Lei, J. van der, Mulligen, E. van, Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. and Mons, B.** (2016) 'The FAIR Guiding Principles for scientific data management and stewardship', *Scientific Data*, 3, p. 160018. Available at: <https://doi.org/10.1038/sdata.2016.18>
- Wilms, K.L., Stieglitz, S., Ross, B. and Meske, C.** (2020) A value-based perspective on supporting and hindering factors for research data management. *International Journal of Information Management* 54, 102174. Available at: <https://doi.org/10.1016/j.ijinfomgt.2020.102174>
- Zilioli, M., Bergami, C., Carrara, P., Fugazza, C., Oggioni, A., Pugnetti, A. and Tagliolato Acquaviva d'Aragona, P.** (2021) 'Enabling the reuse of long-term marine biological observations in essential variables frameworks through a practical approach', *Front. Mar. Sci.*, 8. Available at: <https://doi.org/10.3389/fmars.2021.645997>

#### TO CITE THIS ARTICLE:

Flynn, S., Keena, T., Bogan, Y., Brophy, L., Currie, D., Leadbetter, A., Maloney, M., Manson, K., Melville, C., O'Grady, E., Thomas, R., Whittle, B. and Conway, A. 2025 Establishing a Data Culture Using Frameworks to Navigate the Waves of Marine Data. *Data Science Journal*, 24: 24, pp. 1–19. DOI: <https://doi.org/10.5334/dsj-2025-024>

**Submitted:** 06 May 2025  
**Accepted:** 27 August 2025  
**Published:** 08 September 2025

#### COPYRIGHT:

© 2025 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

*Data Science Journal* is a peer-reviewed open access journal published by Ubiquity Press.