

Data Harmonisation and Interoperability

Helps practitioners harmonise and integrate ocean accounting data from diverse sources.

Version 0.7.1 **Status** Draft **Stability** Applied

Last updated September 2026 **Responsible officer** Dr Jordan Gacutan

Drafting AI-assisted drafting from GOAP Secretariat source materials; source list to be published at the close of consultation.

DRAFT

1. Outcome	3
2. Requirements	3
3. Guidance Material	4
3.1 Data Standards	5
3.2 Exchange Formats	7
3.3 Interoperability Protocols	8
3.4 Classification Concordances	9
3.5 Data Harmonisation Workflow	10
3.6 Modular Data Architecture Patterns	14
3.7 Science-to-Accounts Translation Protocols	17
3.8 Time-Series Data Investment Guidance	20
4. Summary	22
Implementation Considerations	23
5. Acknowledgements	23
6. References	23

ABOUT THIS CIRCULAR

RECOMMENDED CITATION

Data Harmonisation and Interoperability (2026) (GOAP Technical Guidance on Ocean Accounting, IG-4.0, V0.7.1). GOAP Secretariat.
<https://guidance.oceanaccounts.org/circulars/section/4-0-data-harmonisation/>

© 2026 Global Ocean Accounts Partnership. Licensed under CC BY-NC 4.0.

International comparability: Countries implementing ocean accounts require comparable indicators for benchmarking, knowledge exchange, and coordination of regional ocean governance. The SDMX framework provides the international standard for exchanging statistical data, enabling ocean accounting indicators to be disseminated through the same infrastructure as national accounts, environmental-economic accounts, and SDG indicators. Alignment with geospatial standards from the Open Geospatial Consortium and International Hydrographic Organization ensures that spatial data components of ocean accounts can similarly be exchanged across jurisdictions.

Standardised dissemination: Publication of ocean accounting data through SDMX-compliant data repositories enables automated discovery and retrieval by policy analysts, researchers, and decision-support systems. Classification concordances ensure that data collected under national frameworks can be mapped to international reference classifications, so that national detail is preserved and the data remain reusable.

This Circular covers the principal international standards for statistical data exchange, including (1) the Statistical Data and Metadata Exchange (SDMX) framework, (2) geospatial interoperability standards from the Open Geospatial Consortium (OGC) and International Hydrographic Organization (IHO), and (3) classification concordance mechanisms that enable mapping between different classification systems. By applying the principles and protocols described here, compilers can establish data integration workflows that raise data quality and reduce manual processing.

1. Outcome

¶1 After completing this Circular, practitioners will be able to harmonise and integrate ocean accounting data from diverse sources, formats, and domains, applying interoperability standards that address the technical and methodological challenges of multi-source data integration. Data harmonisation underpins three decision contexts for ocean accounting: cross-agency data sharing within national statistical systems, international comparability of ocean accounts across jurisdictions, and dissemination of ocean accounting data to external users through standardised exchange platforms.

¶2 **Cross-agency data sharing:** Ocean accounts draw upon data from national statistical offices (economic statistics), environmental agencies (ecosystem monitoring), hydrographic organisations (bathymetry), fisheries management authorities (catch and effort data), and maritime administrations (vessel tracking). Without harmonised classifications, coordinate reference systems, and exchange formats, each bilateral data transfer requires bespoke transformation procedures. The standards described in this Circular enable agencies to establish common data integration protocols that reduce transaction costs and improve data quality through systematic quality checks.

¶3 **International comparability:** Countries implementing ocean accounts require comparable indicators for benchmarking, knowledge exchange, and coordination of regional ocean governance. The SDMX framework provides the international standard for exchanging statistical data, enabling ocean accounting indicators to be disseminated through the same infrastructure as national accounts, environmental-economic accounts, and SDG indicators. Alignment with geospatial standards from the Open Geospatial Consortium and International Hydrographic Organization ensures that spatial data components of ocean accounts can similarly be exchanged across jurisdictions.

¶4 **Standardised dissemination:** Publication of ocean accounting data through SDMX-compliant data repositories enables automated discovery and retrieval by policy analysts, researchers, and decision-support systems. Classification concordances ensure that data collected under national frameworks can be mapped to international reference classifications, so that national detail is preserved and the data remain reusable.

¶5 This Circular covers the principal international standards for statistical data exchange, including (1) the Statistical Data and Metadata Exchange (SDMX) framework, (2) geospatial interoperability standards from the Open Geospatial Consortium (OGC) and International Hydrographic Organization (IHO), and (3) classification concordance mechanisms that enable mapping between different classification systems. By applying the principles and protocols described here, compilers can establish data integration workflows that raise data quality and reduce manual processing.

2. Requirements

¶1 **Essential prerequisites:**

- ¶2 **TG-0.1 General Introduction to Ocean Accounts:** conceptual framework and key components of Ocean Accounts.

¶3 **Helpful background:**

- ¶4 **TG-0.7 Quality Assurance Principles:** quality dimensions and assessment frameworks, particularly coherence and comparability.

¶5 ⓘ Framework Position

¶6 This Circular addresses data harmonisation and interoperability across the accounting domains. Harmonisation ensures consistent measurement across edges linking economic flows to environmental stocks and ecosystem services to economic activity. In the Ocean Accounts Framework (TG-0.1 Figure 0.1.2):

Edge	Direction	Description
E3	FG1 ↔ SG1	Monetary flows between assets and economic sectors
E9	SG3 → FG1	Ecosystem services to economy (SEEA)

3. Guidance Material

¶1 Ocean accounts draw upon a diverse range of sources: fisheries catch statistics from national statistical offices, bathymetric surveys from hydrographic agencies, satellite observations from earth observation programmes, ecosystem condition assessments from environmental monitoring networks, and economic statistics from national accounts. Each of these data streams has evolved within its own institutional context, employing different standards, classifications, formats, and temporal and spatial reference systems. Ocean accounting must harmonise these disparate data streams into a coherent accounting framework whilst preserving data quality and provenance. For guidance on spatial delineation of accounting units, see **TG-2.12 Spatial Data Frameworks** Sections 3.2 and 3.3. For temporal considerations, see **TG-2.12 Spatial Data Frameworks** Section 3.4.

¶2 This section examines four interrelated dimensions of data harmonisation and interoperability: data standards that define common structures and semantics, exchange formats that enable machine-readable data transmission, interoperability protocols that support automated data discovery and retrieval, and classification concordances that enable translation between different coding systems. Together, these elements form the technical infrastructure for integrated ocean accounting.

¶3 Figure 4.6.1 shows how economic data (SDMX), geospatial and marine data (OGC/IHO S-100), and classification concordances feed an integration layer that produces coherent ocean accounts¹.

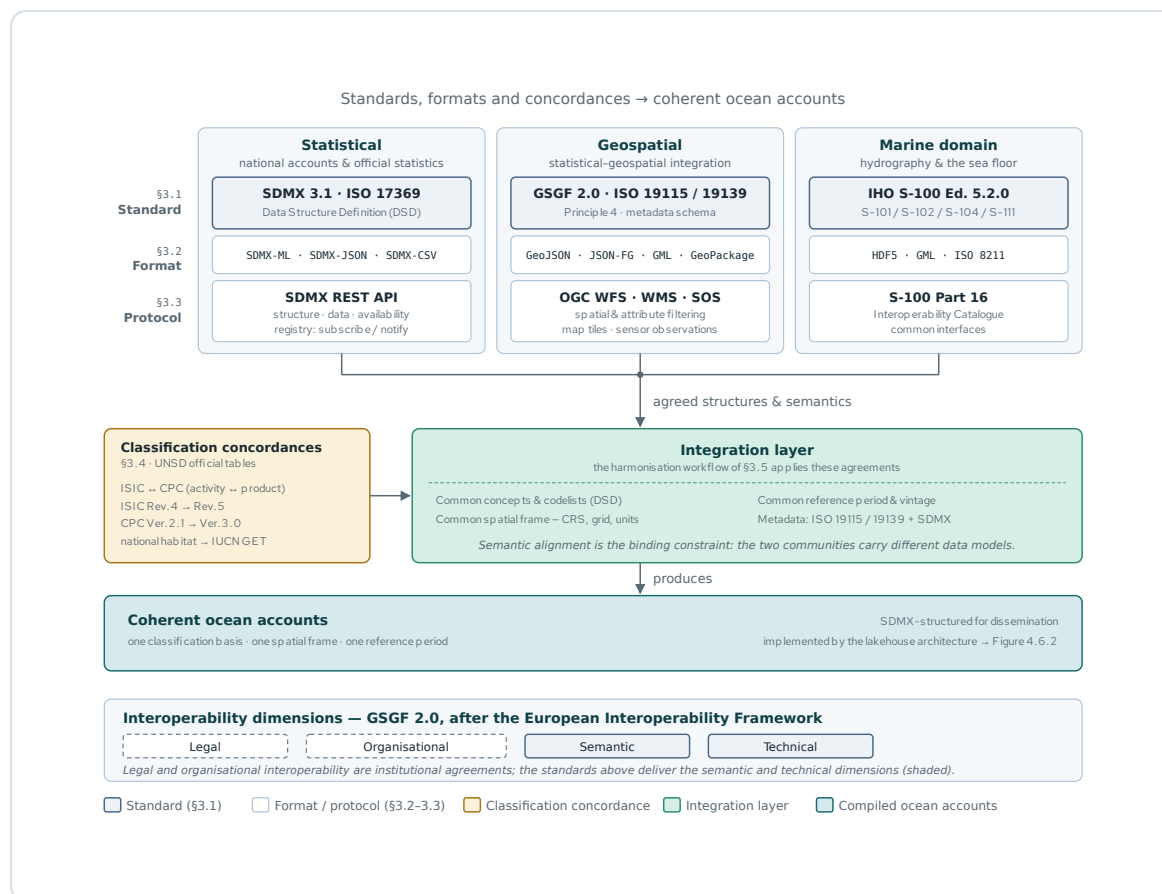


Figure 4.6.1 Statistical, geospatial and marine standards, with classification concordances, converge on an integration layer that yields ocean accounts on one classification, spatial and temporal basis. Colour marks role, not family. The system that runs this integration is Figure 4.6.2.

Source: TG-4.6 §3.1 (data standards, Tables 3.1.1–3.1.2), §3.2 (exchange formats), §3.3 (interoperability protocols) and §3.4 (classification concordances); SDMX 3.1 information model (May 2025), ISO 17369; GSGF v2.0 (UN-GGIM / UNSC 2025); ISO 19115 / ISO/TS 19139; IHO S-100 Edition 5.2.0 (June 2024) and the S-101 / S-102 / S-104 / S-111 product specifications (December 2024); UNSD ISIC Rev.4–Rev.5 and CPC Ver.2.1–Ver.3.0 correspondence tables; IUCN Global Ecosystem Typology.

3.1 Data Standards

¶1 Data standards provide the foundational agreements on how data should be structured, described, and interpreted. For ocean accounting, three families of standards are particularly important: statistical data standards, geospatial data standards, and marine domain-specific standards. The application of these standards supports the quality dimensions described in **TG-0.7 Quality Assurance Principles**, particularly coherence and comparability.

Statistical Data and Metadata Exchange (SDMX)

¶1 The **Statistical Data and Metadata Exchange (SDMX)** initiative provides the international standard for exchanging statistical data and metadata². Developed by the seven sponsor organisations (BIS, ECB, Eurostat, IMF, OECD, United Nations, and World Bank), SDMX has been designated as an ISO International Standard (ISO 17369)³. The SDMX 3.1 version, released in May 2025, introduces enhanced support for geospatial data, microdata, and improved data structure definitions⁴.

¶2 SDMX is built around an **information model** that defines the structural metadata needed to describe statistical data⁵. At its core is the **Data Structure Definition (DSD)**, which specifies the dimensions, attributes, and measures that characterise a particular data collection. For ocean accounting, DSDs would define the structure of fisheries statistics, marine economic activity data, and

environmental flow accounts. The structure of physical flow accounts is described in **TG-2.6 Ocean Economy Investment**.

¶3 The SDMX framework supports three process patterns for data exchange⁶:

1. ¶4 **Bilateral exchange** – counterparties agree on specific formats and processes
2. ¶5 **Gateway exchange** – multiple organisations agree on a common exchange format
3. ¶6 **Data-sharing exchange** – open standards enable any organisation to access and use data

¶7 For ocean accounting, the data-sharing model is particularly relevant as it enables national statistical offices, environmental agencies, and research institutions to publish data in standard formats that can be automatically discovered and integrated.

¶8 SDMX provides three transmission formats (SDMX-ML, SDMX-JSON, SDMX-CSV)⁷, detailed in Section 3.2 below. The CSV format is particularly accessible to non-specialist users whilst remaining losslessly convertible to the other formats⁸. SDMX registry services provide “visibility into the data and metadata existing within the community, and support the access and use of this data and metadata by providing a set of triggers for automated processing”⁹. A registry-based architecture enables ocean accounting compilers to discover available data sources, retrieve structural metadata, and automate data collection workflows.

Geospatial Standards

¶1 The **Global Statistical Geospatial Framework (GSGF)**, version 2.0 released by UN-GGIM and the Statistical Commission in 2025, provides the overarching framework for integrating statistical and geospatial information¹⁰. The GSGF defines five principles, of which Principle 4 on “Statistical and geospatial interoperability” is directly relevant to ocean accounting data integration¹¹.

¶2 Data interoperability, as defined in the GSGF, refers to “the ability of different systems, organizations, and applications to exchange, interpret, and use data in a coordinated manner”¹². The framework identifies four interoperability dimensions derived from the European Interoperability Framework (EIF)¹³. Table 3.1.1 below summarises these dimensions.

Dimension	Description
Legal interoperability	Enabling organisations under different legal frameworks to share data.
Organisational interoperability	Aligning business processes and responsibilities.
Semantic interoperability	Exchanging data through common terminology and formats.
Technical interoperability	Linking systems through standard interfaces and formats.

¶3 For ocean accounting, semantic interoperability is particularly challenging as statistical and geospatial communities have “evolved their own general data models, metadata capabilities, architectures, and data infrastructures, creating differences in fundamental terminology over time”¹⁴.

¶4 The **Open Geospatial Consortium (OGC)** provides the principal geospatial interoperability standards relevant to ocean accounting¹⁵. Table 3.1.2 below summarises those standards.

Standard	Description
Web Feature Service (WFS)	For vector data exchange, applicable to administrative boundaries, marine protected area polygons, and ecosystem extent units.
Web Map Service (WMS)	For raster map images and tiles, useful for bathymetric visualisation and land cover products.
Geography Markup Language (GML)	The XML encoding standard for geospatial features.

¶5 Two additional ISO standards are essential for geospatial metadata management. **ISO 19115** (*Geographic information – Metadata*) provides the standard schema for describing geographic datasets and services, defining the metadata elements needed for data discovery, evaluation, and

use¹⁶. Its XML implementation, **ISO/TS 19139**, specifies the encoding format in which ISO 19115 metadata should be stored and exchanged, enabling machine-readable metadata interchange between systems¹⁷. Together, these standards ensure that geospatial data used in ocean accounts carry sufficient provenance and quality information for informed integration.

¶6 The GSGF recommends that organisations “host appropriate technical infrastructures which support the use of the relevant standards where systems and services are linked through standard interfaces, services, and data formats”¹⁸.

Marine Domain Standards: IHO S-100

¶1 The **International Hydrographic Organization (IHO) S-100** standard provides the Universal Hydrographic Data Model for marine geospatial data¹⁹. S-100 Edition 5.2.0, adopted in June 2024, builds upon the earlier Edition 5.1.0 (October 2023) and establishes “a contemporary hydrographic geospatial data standard that can support a wide variety of hydrographic-related digital data sources, and is fully aligned with mainstream international geospatial standards, in particular the ISO 19100 series”²⁰. In December 2024, IHO Member States approved operational editions of key S-100-based Product Specifications, including S-101 (Electronic Navigational Charts), S-102 (Bathymetric Surface), S-104 (Water Level Information), and S-111 (Surface Currents)²¹.

¶2 S-100 is designed to support applications “that go beyond the scope of traditional hydrography – for example, high-density bathymetry, seafloor classification, marine GIS”²². The standard comprises multiple parts covering conceptual schema, feature catalogues, spatial schemas, metadata, portrayal, and encoding formats including ISO 8211, GML, and HDF5²³.

¶3 Part 16 of S-100 specifically addresses the **Interoperability Catalogue Model**, which “enables interoperability between disparate technologies through the use of common interfaces”²⁴. For ocean accounts that integrate hydrographic data (bathymetry, tides, currents) with ecosystem and economic data, S-100 provides the interoperability framework for the marine domain component.

3.2 Exchange Formats

¶1 Standardised exchange formats enable data to be transmitted between systems in machine-readable form. The choice of format affects both the technical ease of data integration and the preservation of semantic content.

SDMX Transmission Formats

¶1 The three SDMX transmission formats introduced in Section 3.1 serve distinct use cases²⁵:

- ¶2 **SDMX-ML** (XML-based) provides the most complete expression of the SDMX information model. The Structure Specific Data format is “the sole XML format for data exchange” following the deprecation of legacy formats in SDMX 3.0²⁶. Appropriate for formal data exchanges where complete metadata preservation is required.
- ¶3 **SDMX-JSON** is “well-suited to web-based applications and modern programming environments”²⁷ and is increasingly preferred for API-based data access.
- ¶4 **SDMX-CSV** provides a “simple columnar format” that offers accessibility for practitioners without specialised tools whilst still enabling conversion to other SDMX formats “without loss”²⁸. CSV may serve as a practical intermediate format for data validation and review.

Geospatial Formats

¶1 For geospatial ocean accounting data, several format standards are relevant:

¶2 **GeoJSON** and **JSON-FG** (OGC Features and Geometries JSON) provide JSON-based formats for geographic features that are “well-suited to web mapping applications”²⁹. These formats are appropriate for ecosystem extent boundaries, marine protected area polygons, and other vector features.

¶3 **GML** (Geography Markup Language) provides the full XML encoding for geographic information conforming to ISO 19136³⁰. GML is required for formal data exchange in contexts where complete geometry representation and coordinate reference system information must be preserved.

¶4 **HDF5** (Hierarchical Data Format version 5) is adopted by IHO S-100 for gridded and imagery data³¹. HDF5 is appropriate for bathymetric grids, ocean temperature fields, and other multi-dimensional environmental data.

¶5 **GeoPackage** provides an SQLite-based container format that can store both vector and raster data. GeoPackage is increasingly used for distributing compiled geospatial datasets.

3.3 Interoperability Protocols

¶1 Interoperability protocols define how systems communicate to discover, query, and retrieve data. Effective protocols enable automation of data collection workflows. Automation reduces manual processing and improves timeliness.

SDMX RESTful Web Services

¶1 SDMX defines a **RESTful web services API** for querying data and structural metadata³². The API enables:

- ¶2 Structure queries – retrieving Data Structure Definitions, codelists, concept schemes
- ¶3 Data queries – retrieving data observations filtered by dimension values
- ¶4 Availability queries – discovering what data exists for given parameters

¶5 The RESTful API design means queries are expressed through URL patterns, enabling integration with standard web tools. For example, a query for fisheries catch data might specify the dataflow, reference area, time period, and species codes through URL parameters.

¶6 The SDMX registry architecture supports “automated processing” through subscription and notification services³³. Compilers can subscribe to data updates and receive notifications when new data becomes available, enabling near-real-time data integration workflows.

OGC Web Services

¶1 The OGC web services introduced in Section 3.1 (Table 3.1.2) function as interoperability protocols: WFS supports spatial and attribute filtering for retrieval of ecosystem assets within specified bounds, whilst WMS supports rendered map visualisation of spatial context. The **Sensor Observation Service (SOS)** complements these by providing access to sensor data and observations³⁴, relevant for integrating ocean monitoring data (water quality, temperature, wave height) into ocean accounts.

¶2 The GSGF emphasises that technical infrastructure should support “standard interfaces, services, and data formats, including metadata standards”¹⁸. Implementation of OGC services enables ocean accounting systems to integrate with national spatial data infrastructures.

FAIR Data Principles

¶1 The **FAIR Principles** (Findability, Accessibility, Interoperability, Reusability) provide guiding principles for scientific data management that are increasingly applied to statistical data³⁵. The GSGF explicitly incorporates FAIR principles, noting that the Interoperability principle “emphasises meta-data” and that “(meta)data should be ready to be exchanged, interpreted and combined in a (semi)automated way”³⁶.

¶2 For ocean accounting, FAIR principles translate to practical requirements summarised in Table 3.3.1 below.

Principle	Practical requirement
Findability	Data should be registered in searchable catalogues with persistent identifiers.

Principle	Practical requirement
Accessibility	Data should be retrievable through standard protocols with clear access conditions.
Interoperability	Data should use standard formats, vocabularies, and include qualified references to other data.
Reusability	Data should have clear provenance and appropriate licenses.

¶3 The FAIR principles align with quality dimensions in **TG-0.7 Quality Assurance Principles** and should guide the design of ocean accounting data management systems.

3.4 Classification Concordances

¶1 Classification concordances (also called crosswalks or mapping tables) enable translation between different classification systems. For ocean accounting, concordances are essential for:

- ¶2 Linking economic activity data (ISIC) with product data (CPC)
- ¶3 Mapping national ecosystem classifications to international reference classifications
- ¶4 Converting between historical and current classification versions
- ¶5 Integrating data collected under different classification frameworks

¶6 The classification systems relevant to ocean accounting are described in **TG-0.2 Overview of Relevant Statistical Standards**.

Activity and Product Classifications

¶1 The relationship between **ISIC** (activity classification) and **CPC** (product classification) is fundamental to economic statistics. As the CPC documentation explains, “each subclass of the CPC consists of goods or services that are generally produced in a specific class or classes of the ISIC”³⁷. The CPC documentation includes explicit correspondences to ISIC, enabling ocean accounting compilers to link product flows with producing industries.

¶2 The CPC serves as a “central” classification providing “a framework for international comparison and promotes harmonization of various types of statistics related to goods and services”³⁸. The classification was “developed primarily to enhance harmonization among various fields of economic and related statistics and to strengthen the role of national accounts as an instrument for the coordination of economic statistics”³⁹.

¶3 With the endorsement of **ISIC Rev.5** and **CPC Ver.3.0** by the UN Statistical Commission in March 2024, compilers should note that the division-level structure for ocean-relevant activities has been maintained, though with increased detail at lower classification levels⁴⁰. For ocean accounting, key concordance relationships include:

- ¶4 **ISIC Division 03 (Fishing and Aquaculture)** corresponds to **CPC Division 04 (Fish and Other Fishing Products)**
- ¶5 **ISIC Division 50 (Water Transport)** corresponds to relevant groups within **CPC Divisions 64–67 (Transport and Supporting Services)**, with specific correspondence at the group and class level documented in official UNSD concordance tables

¶6 The UNSD has developed official correspondence tables for the ISIC Rev.4-to-Rev.5 and CPC Ver.2.1-to-Ver.3.0 transitions⁴¹. These tables should be used when bridging data compiled under different classification versions. Compilers working with historical data collected under ISIC Rev.4 and CPC Ver.2.1 should apply these official correspondence tables rather than constructing ad hoc mappings, to ensure consistency and comparability across time.

¶7 Detailed guidance on ocean-relevant economic classifications is provided in **TG-0.2 Overview of Relevant Statistical Standards**.

Ecosystem Type Concordances

¶1 The **IUCN Global Ecosystem Typology (GET)** provides the reference classification for ecosystem types in SEEA Ecosystem Accounting⁴². However, “countries may have their own national classification system of ecosystems (or ecological areas) that could be used for the extent accounts. In such cases, developing a bridge or concordance (often called a schema crosswalk in GIS) of this national classification system with the GET reference classification may facilitate comparability across countries”⁴³.

¶2 For marine ecosystems specifically, concordances may be required between:

- ¶3 National marine habitat classifications and GET Marine Realm types
- ¶4 EUNIS habitat classification and GET Ecosystem Functional Groups
- ¶5 National land cover classifications and GET transitional realm types (for coastal areas)

¶6 The SEEA Biophysical Modelling Guidelines note that the GET “represents a global typological framework that applies a process-based approach to ecosystem classification across the whole planet”⁴⁴. The hierarchical structure (Realms, Biomes, Ecosystem Functional Groups) provides natural aggregation levels for concordance development.

¶7 Guidance on ecosystem classification for ocean accounts is provided in **TG-4.1 Remote Sensing and Geospatial Data** Section 3.2.4 and **TG-3.1 Asset Accounts**.

Concordance Development Process

¶1 Developing classification concordances involves:

1. ¶2 **Structural comparison** – examining the hierarchical structure and level of detail in each classification
2. ¶3 **Conceptual alignment** – identifying where definitions align or diverge
3. ¶4 **Mapping relationships** – defining one-to-one, one-to-many, or many-to-many relationships
4. ¶5 **Documentation** – recording mapping rationale and any approximations
5. ¶6 **Validation** – testing concordances with actual data

¶7 The UNFC (United Nations Framework Classification for Resources) provides an example of formalised concordance development: “mapping schemes have been developed showing the link between the UNFC2009 and the SPE and CRIRSCO classifications” for mineral and petroleum resources⁴⁵.

Temporal Concordances

¶1 Classification systems evolve over time through revision processes. ISIC has progressed through Revisions 3, 3.1, 4, and now 5. The CPC has progressed through Versions 1.0, 1.1, 2.0, 2.1, and now 3.0. The UN Statistics Division maintains official correspondence tables between versions to enable time series compilation across classification changes.

¶2 For ocean accounting, temporal concordances are required when:

- ¶3 Compiling time series that span classification revisions
- ¶4 Integrating historical data collected under previous classifications
- ¶5 Comparing international data collected under different version implementations

¶6 The ISIC documentation emphasises that “continuity, i.e., comparability between the revised and preceding versions of ISIC, has always been a major concern expressed by the Commission”⁴⁶. The official ISIC Rev.4-to-Rev.5 correspondence table, developed between March and October 2024 following the endorsement of ISIC Rev.5 explanatory notes, should be consulted when compiling time series that bridge these classification versions⁴⁷. Similarly, the CPC Ver.2.1-to-Ver.3.0 correspondence table is available through the UNSD classifications portal⁴⁸.

3.5 Data Harmonisation Workflow

¶1 Implementing data harmonisation for ocean accounts requires a systematic workflow that transforms heterogeneous source data into consistent, integrated accounting tables. This section describes the compilation procedure and provides a worked example.

Compilation Procedure

¶1 The data harmonisation workflow consists of four sequential phases:

¶2 **Phase 1: Classify** – Assign source data to standardised classifications. For ocean accounting, this involves:

- ¶3 Mapping economic activity data to ISIC Rev.5 divisions and groups
- ¶4 Mapping product flows to CPC Ver.3.0 categories
- ¶5 Mapping ecosystem types to IUCN GET Ecosystem Functional Groups
- ¶6 Mapping spatial units to the national spatial framework and coordinate reference system

¶7 Classification mappings should be documented in concordance tables that record the source classification, target classification, mapping type (1:1, 1:many, many:1, many:many), and any assumptions or approximations.

¶8 **Phase 2: Map** – Transform source data values to common units and temporal reference periods. This includes:

- ¶9 Converting spatial coordinates to the accounting area's coordinate reference system (typically WGS84 for horizontal coordinates)
- ¶10 Aligning temporal references to accounting period start/end dates (calendar year, fiscal year)
- ¶11 Converting physical units to standard measurement units (e.g., tonnes for biomass, hectares for extent, cubic metres for water volumes)
- ¶12 Converting monetary values to constant prices using deflators where time series require comparability

¶13 Mapping transformations should preserve source values in metadata to enable quality checks and re-transformation if specifications change.

¶14 **Phase 3: Validate** – Check integrated data for consistency, completeness, and plausibility. Validation includes:

- ¶15 Range checks: values fall within plausible bounds (e.g., catch cannot exceed vessel capacity)
- ¶16 Balance checks: supply equals use in flow accounts
- ¶17 Spatial checks: geometries are valid, no gaps or overlaps in exhaustive spatial frameworks
- ¶18 Temporal checks: time series are continuous, no unexplained breaks
- ¶19 Cross-source confrontation: data from multiple sources describing the same phenomena are reconciled

¶20 Validation should generate exception reports that flag issues for resolution. The quality assurance framework in **TG-0.7 Quality Assurance Principles** provides detailed validation procedures.

¶21 **Phase 4: Integrate** – Compile validated data into accounting tables. Integration produces:

- ¶22 Asset accounts combining physical extent, condition indicators, and monetary valuations
- ¶23 Flow accounts recording extractions, emissions, and ecosystem service flows
- ¶24 Supply-use tables linking ocean economy production and consumption
- ¶25 Combined presentations integrating environmental, economic, and social dimensions (see **TG-3.8 Combined Presentations**)

¶26 *Table 3.5.1: Data harmonisation workflow phases*

Phase	Inputs	Transformation	Outputs	Quality Checks
1. Classify	Source data with native classifications	Apply concordance tables	Data assigned to standard classifications	Classification coverage: all source items mapped
2. Map	Classified data in source units/CRS	Unit conversions, coordinate transformations, temporal alignment	Data in common units, CRS, accounting periods	Transformation validity: reversible where possible
3. Validate	Mapped data from multiple sources	Range, balance, spatial, temporal checks	Validated data with exception reports	Consistency: cross-source confrontation resolved
4. Integrate	Validated data	Compile into accounting tables	Asset, flow, activity accounts; combined presentations	Completeness: all required cells populated or flagged

Worked Example: Harmonising Fisheries Data Using SDMX

¶1 This synthetic example demonstrates the data harmonisation workflow for integrating fisheries catch data from multiple agencies into ocean accounts. The scenario: a national statistical office is compiling fisheries accounts and must integrate data from three sources with different reporting frameworks:

- ¶2 **Source A:** National fisheries agency – species-level catch data by fishing zone, reported quarterly in tonnes live weight
- ¶3 **Source B:** Port authority – landing declarations by vessel and port, reported monthly in various product forms (whole fish, fillets, processed)
- ¶4 **Source C:** Customs administration – seafood export data using national tariff codes, reported monthly in kilograms and monetary value (FOB)

¶5 Phase 1: Classify

¶6 The compiler establishes an SDMX Data Structure Definition (DSD) for fisheries flow accounts with the following dimensions:

- ¶7 **REF_AREA:** Fishing zone (territorial waters, EEZ zones by bathymetry, high seas)
- ¶8 **TIME_PERIOD:** Accounting period (annual, with quarterly detail available)
- ¶9 **SPECIES:** Fish species using FAO ASFIS codes
- ¶10 **ACTIVITY:** Producing industry using ISIC Rev.5 codes (Group 031: Marine fishing)
- ¶11 **PRODUCT:** Output product using CPC Ver.3.0 codes (Division 04: Fish and other fishing products)

¶12 For each source, the compiler develops classification mappings:

- ¶13 Source A reports species using national codes, and a concordance table maps these to FAO ASFIS codes (mostly 1:1)
- ¶14 Source B records product forms, and a concordance table maps these to CPC Ver.3.0 with conversion factors to derive live weight equivalent (1:many mapping: one product code may correspond to multiple CPC classes depending on processing level)
- ¶15 Source C uses HS 2022 tariff codes, and the official UNSD concordance table maps these to CPC Ver.3.0 (many:many mapping requiring allocation based on supplementary information)

¶16 Phase 2: Map

¶17 Unit conversions and temporal alignment:

- ¶18 Source A data are already in tonnes live weight, and temporal alignment aggregates quarterly data to annual totals for stock accounts whilst preserving quarterly detail for

seasonal analysis

- ¶19 Source B landing weights require conversion to live weight equivalent using species-specific factors (e.g., fillets to whole fish factor 2.5 for snapper, 2.8 for tuna). Conversion factors documented in metadata
- ¶20 Source C export data in kilograms converted to tonnes product weight, and to live weight equivalent using the same species- and product-form factors as Source B, with monetary values (FOB) converted to national currency constant prices using fish price deflator

¶21 Spatial mapping:

- ¶22 Source A fishing zones map to ocean accounting spatial framework using GIS overlay of zone boundaries
- ¶23 Source B port locations geocoded using national address gazetteer, with landings attributed to fishing zones using vessel logbook data on fishing grounds (where available) or pro-rata allocation based on Source A zone proportions for species
- ¶24 Source C export port locations geocoded, with spatial attribution to production zones inferred from Source A and B integration

¶25 **Phase 3: Validate**

¶26 Validation checks applied:

- ¶27 **Range check:** No individual vessel catch exceeds vessel hold capacity (cross-reference with vessel register)
- ¶28 **Balance check:** Total catch (Source A) equals total landings (Source B) plus estimated at-sea discards, within tolerance threshold of 5%
- ¶29 **Cross-source confrontation:** Export volumes (Source C) reconciled with landing volumes (Source B) accounting for domestic consumption, processing losses, and stock changes

¶30 Discrepancies identified:

- ¶31 Source A reports 15,200 tonnes tuna catch in EEZ, whilst Source B records 14,100 tonnes landings. The difference is attributed to unreported landings (estimated 600 tonnes based on observer coverage) and at-sea discards (estimated 500 tonnes, 3.3% discard rate consistent with observer data)
- ¶32 Source C reports 8,200 tonnes tuna exports in product weight, or 9,100 tonnes live weight equivalent (the 900-tonne difference is weight removed in processing), which reconcile with Source B after accounting for 4,800 tonnes domestic consumption and 200 tonnes cold storage stock increase

¶33 Validation produces exception report documenting assumptions and highlighting areas requiring further investigation (unreported catch estimation method).

¶34 **Phase 4: Integrate**

¶35 Validated data compiled into SDMX-compliant fisheries flow account table with dimensions:

REF_AREA	TIME_PERIOD	SPECIES	ACTIVITY	PRODUCT	OBS_VALUE	UNIT
EEZ_ZONE_1	2024	YFT (Yellowfin tuna)	ISIC 031	CPC 04111	8,500	Tonnes
EEZ_ZONE_2	2024	YFT	ISIC 031	CPC 04111	6,700	Tonnes
TERRITORIAL	2024	SKJ (Skipjack)	ISIC 031	CPC 04112	3,200	Tonnes

¶36 The DSD includes attributes documenting data sources, estimation methods, and quality flags. The compiled data structure enables:

- ¶37 Dissemination through SDMX registry for automated retrieval by analysts
- ¶38 Integration with economic accounts using ISIC and CPC concordances
- ¶39 Comparison with international fisheries statistics using FAO ASFIS species codes
- ¶40 Spatial analysis of catches by zone for marine spatial planning

¶41 The worked example demonstrates how SDMX Data Structure Definitions provide a standardised framework for harmonising heterogeneous source data. That framework supports systematic quality checks and produces interoperable outputs for ocean accounting.

3.6 Modular Data Architecture Patterns

¶1 The technical infrastructure supporting ocean account data harmonisation must accommodate diverse data sources, multiple processing workflows, and varied dissemination requirements. This section describes modular architecture patterns that enable national statistical offices and ocean accounting programmes to build scalable, maintainable data systems. The guidance draws on established practices in statistical data management and modern data engineering, adapted to the specific requirements of environmental-economic accounting.

Data lake versus data warehouse

¶1 Two architectural paradigms are relevant for ocean accounting data infrastructure: data lakes and data warehouses. Each serves a distinct purpose, and many implementations will employ both in combination.

¶2 A **data lake** stores data in its original format (raw satellite imagery, CSV files from monitoring stations, spreadsheets from fisheries agencies, GIS shapefiles) without imposing a predefined schema. The data lake preserves full provenance and enables re-processing when methodologies change. For ocean accounting, a data lake is appropriate for ingesting heterogeneous source data from the diverse agencies described in Section 3.1, including remote sensing products (**TG-4.1 Remote Sensing and Geospatial Data**), survey microdata (**TG-4.2 Survey Methods for Ocean Economic Activity**), administrative records (**TG-4.3 Administrative Data Sources**), and citizen science observations (**TG-4.4 Citizen Science**).

¶3 A **data warehouse** stores data in a structured, schema-enforced format optimised for analytical queries and reporting. The warehouse contains harmonised data that has passed through the classify-map-validate-integrate workflow described in Section 3.5. For ocean accounting, the data warehouse holds the compiled account tables in SDMX-compliant structures, ready for dissemination and analysis. The data warehouse enforces the classification concordances (Section 3.4) and accounting identities that ensure internal consistency.

¶4 The recommended pattern for ocean accounting is a **lakehouse architecture** that combines both approaches: source data are ingested into a data lake, processed through harmonisation pipelines, and loaded into a structured warehouse layer for dissemination. This pattern provides the flexibility to accommodate new data sources without redesigning the warehouse schema, whilst maintaining the rigour required for official statistical products. Figure 4.6.2 sets out the end-to-end lakehouse data architecture, in which heterogeneous source systems feed a data lake, a classify-map-validate-integrate harmonisation pipeline processes the lake's data into a schema-enforced warehouse, and the warehouse publishes through a tiered dissemination layer, with international standards and the metadata, quality and governance framework as cross-cutting rails.

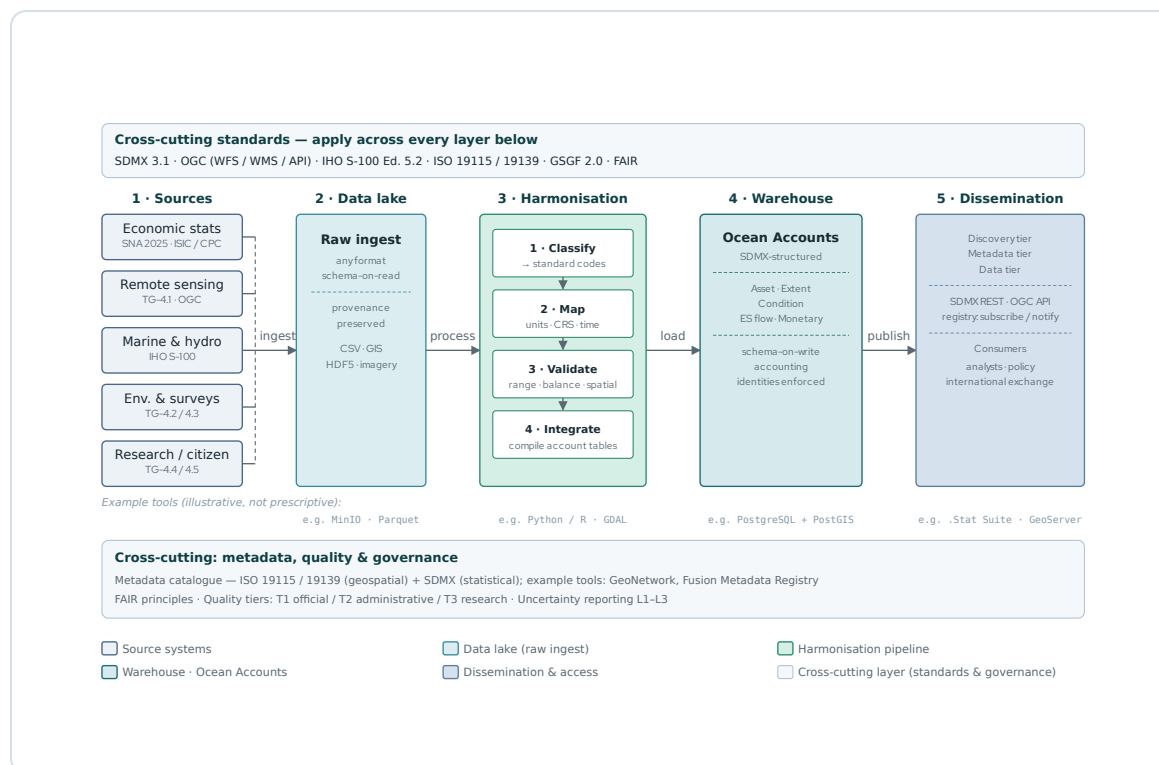


Figure 4.6.2 Lakehouse architecture moves ocean-account data from source systems through lake, warehouse, and dissemination layers under cross-cutting standards rails. Warehouse holds SDMX-structured Ocean Accounts. No adopted SEEA Ocean standard; GOAP TG is the interim framework.

Source: TG-4.6 draft §3.5 (data harmonisation workflow, Table 3.5.1) and §3.6 (modular data architecture patterns, Tables 3.6.1–3.6.2); SDMX 3.1 information model (May 2025); GSGF v2.0 (UN-GGIM / UNSC 2025); IHO S-100 Edition 5.2.0 (June 2024); ISO 19115 / 19139; SEEA CF (2012); SNA 2025.

¶5 **Table 3.6.1: Architecture pattern comparison for ocean accounts**

Characteristic	Data Lake	Data Warehouse	Lakehouse (Recommended)
Schema	Schema-on-read (flexible)	Schema-on-write (enforced)	Both: flexible ingest, enforced output
Data formats	Any (CSV, GIS, HDF5, imagery)	Structured (relational tables, SDMX)	Source formats preserved; output standardised
Processing	Batch and ad hoc	ETL pipelines	Harmonisation workflow (Section 3.5)
Users	Data engineers, researchers	Analysts, dissemination platforms	All stakeholders
Quality control	Minimal at ingest	Enforced at load	Validation at harmonisation stage
Suitable for	Raw data preservation, exploration	Official statistics, reporting	End-to-end ocean accounting

API design for data access

¶1 Application Programming Interfaces (APIs) enable automated data exchange between ocean accounting systems and external users. The API design should follow the SDMX RESTful web services specification (Section 3.3) for statistical data access, supplemented by OGC API standards for geospatial data. Key design principles include:

¶2 APIs should provide **three access tiers**: a discovery tier (what data are available, covering which areas and periods), a metadata tier (data structure definitions, classification codelists, quality reports), and a data tier (observation values with full dimensional context). This mirrors the SDMX

architecture of structure queries, availability queries, and data queries. Each tier should support standard HTTP methods and return responses in both SDMX-JSON and SDMX-CSV formats to accommodate different client capabilities.

¶3 **Versioning** is essential for APIs serving official statistics. Each API version should be maintained for a minimum of two years after a successor version is released, enabling downstream systems to migrate without disruption. Version identifiers should appear in the URL path (e.g., `/api/v2/data/ocean-accounts/...`) following REST conventions.

¶4 **Authentication and access control** should follow the data governance policies established by the national statistical office. Public-use indicators may be served without authentication, whilst microdata or geographically detailed data may require registration and licence acceptance, consistent with the FAIR principles described in Section 3.3.

Metadata standards

¶1 Complete metadata is essential for data discovery, evaluation, and reuse. Ocean accounting data should carry metadata conforming to two complementary standards:

¶2 **Geospatial metadata** follows ISO 19115 / ISO 19139 (introduced in Section 3.1). All geospatial data products used in ocean accounts (ecosystem extent maps, bathymetric grids, spatial unit boundaries) should carry ISO 19115-compliant metadata encoded in ISO 19139 XML, documenting the coordinate reference system, spatial resolution, temporal coverage, processing history, and quality assessment results.

¶3 **Statistical metadata** follows the SDMX information model (Section 3.1). All compiled ocean account tables should be accompanied by SDMX structural metadata defining the dimensions, attributes, and measures of each dataset. Reference metadata (quality reports, methodological descriptions) should follow the SDMX metadata structure definition format.

¶4 Where ocean accounting data combines geospatial and statistical components (as is typical), both metadata standards should be applied. The GSGF (Section 3.1) provides guidance on achieving interoperability between ISO 19115 and SDMX metadata, recommending that organisations maintain metadata catalogues that link geospatial and statistical metadata for the same underlying datasets.

Open-source stack recommendations

¶1 National statistical offices and ocean accounting programmes with constrained budgets can implement the lakehouse architecture using open-source software. The following stack has been demonstrated in environmental-economic accounting contexts and aligns with the standards described in this Circular:

¶2 *Table 3.6.2: Open-source technology stack for ocean accounting data infrastructure*

Layer	Function	Recommended Tools	Standards Supported
Data lake storage	Raw data preservation	MinIO (S3-compatible), Apache Parquet	Any format
Geospatial processing	Spatial data harmonisation	GDAL/OGR, PostGIS, QGIS	OGC WFS/WMS, ISO 19115
Statistical processing	Data transformation and validation	Python (pandas), R	SDMX (via sdmx1 library)
Data warehouse	Structured storage and queries	PostgreSQL + PostGIS	SQL, spatial SQL
API layer	Data dissemination	.Stat Suite (SDMX), GeoServer (OGC)	SDMX REST, OGC API

Layer	Function	Recommended Tools	Standards Supported
Metadata catalogue	Data discovery	GeoNetwork (ISO 19115), Fusion Metadata Registry (SDMX)	ISO 19139, SDMX
Visualisation	Dashboard and reporting	Apache Superset, Observable Framework	Web standards

¶3 The SDMX community maintains the **.Stat Suite**, an open-source platform for SDMX data dissemination that is already used by several national statistical offices and international organisations. For geospatial data, **GeoServer** provides OGC-compliant web services. These tools can be deployed on standard server infrastructure or cloud platforms, with costs scaling according to data volume and user load.

3.7 Science-to-Accounts Translation Protocols

¶1 Ocean accounts depend on data from scientific research programmes (oceanographic surveys, ecological monitoring, biodiversity assessments, and fisheries stock assessments) that are produced for scientific purposes and must be translated into the structured formats required by accounting frameworks. This section provides protocols for converting research datasets into account-ready inputs, managing uncertainty, and establishing quality tiers that reflect the maturity and reliability of different data sources.

Converting research datasets

¶1 Scientific datasets differ from administrative and statistical data in several respects that affect their suitability for ocean accounts. Research data may cover irregular spatial domains (transects, sampling stations) rather than exhaustive spatial frameworks, may use non-standard temporal reference periods (field seasons, project durations) rather than accounting years, and may employ bespoke classification schemes in place of international standard classifications. The conversion protocol addresses each of these differences systematically.

¶2 **Spatial harmonisation:** Research data collected at point locations (monitoring stations, sampling sites) must be interpolated or modelled to produce estimates for the exhaustive spatial units used in ocean accounts. The biophysical modelling guidance in **TG-2.1 Aggregate Biophysical Indicators of Environmental State** describes interpolation and modelling techniques. Compilers should document the interpolation method, the spatial support (area over which each estimate applies), and the estimation uncertainty. Remote sensing data (**TG-4.1 Remote Sensing and Geospatial Data**) can provide spatially complete coverage to supplement point-based research data.

¶3 **Temporal alignment:** Research datasets with non-standard temporal coverage must be adjusted to accounting periods. Where research data cover a different period (e.g., a biological survey conducted from March to November), compilers should assess whether temporal adjustment is needed (for stocks that change slowly, the survey period may be a reasonable proxy for the annual average) or whether correction factors should be applied (for seasonally variable quantities such as biomass or water quality). The temporal alignment should be documented in metadata, including any assumptions about seasonal patterns.

¶4 **Classification mapping:** Research datasets often use scientific taxonomic or habitat classification schemes that differ from the standard classifications used in ocean accounts. Compilers should develop and maintain concordance tables mapping research classifications to accounting classifications (IUCN GET for ecosystem types, FAO ASFIS for species, ISIC/CPC for economic activities). The concordance development process described in Section 3.4 applies. Where research classifications provide finer detail than accounting classifications, the mapping is straightforward (aggregation). Where research classifications are coarser or use different conceptual categories, expert judgement is required and should be documented.

¶5 Table 3.7.1: Science-to-accounts conversion steps

Step	Input	Transformation	Output	Documentation Required
1. Data assessment	Raw research dataset	Evaluate spatial, temporal, and thematic coverage	Gap analysis and fitness-for-purpose assessment	Assessment report with quality rating
2. Spatial harmonisation	Point/transect data	Interpolation, modelling, or aggregation to spatial units	Spatially complete estimates by accounting unit	Method, assumptions, uncertainty estimates
3. Temporal alignment	Survey-period data	Adjustment to accounting year	Annual estimates	Seasonal adjustment method, correction factors
4. Classification mapping	Scientific classifications	Apply concordance tables	Data classified to accounting standards	Concordance table with mapping rationale
5. Unit conversion	Research measurement units	Convert to accounting standard units	Data in standard units (tonnes, km ² , etc.)	Conversion factors and sources
6. Quality annotation	Converted dataset	Assign quality tier and metadata flags	Account-ready dataset with quality metadata	Quality tier justification

Uncertainty propagation

¶1 Scientific data carry measurement uncertainty that must be propagated through the accounting framework. Ocean accounts should report uncertainty for key aggregates, enabling users to assess the confidence they can place in derived indicators. Three levels of uncertainty reporting are recommended, in order of increasing sophistication:

¶2 **Level 1 – Qualitative assessment:** Each data input is assigned a qualitative uncertainty rating (low, medium, high) based on expert judgement of the data source, collection method, and processing steps. This minimum level is feasible for all ocean accounting programmes and should be documented in metadata for every compiled account.

¶3 **Level 2 – Confidence intervals:** Key aggregates (total ecosystem extent, total fish stock biomass, ocean economy GVA) are reported with confidence intervals derived from the uncertainty of input data. Where input uncertainties are characterised as standard errors or confidence intervals, propagation follows standard statistical methods (error propagation for sums, products, and ratios). This level requires quantitative uncertainty estimates for the major input datasets.

¶4 **Level 3 – Monte Carlo simulation:** For accounts involving complex non-linear transformations (such as ecosystem service valuation using benefit transfer, or carbon stock estimation using allometric equations), Monte Carlo simulation provides the most rigorous uncertainty propagation. Input distributions are sampled repeatedly, the accounting calculations are performed for each sample, and the resulting distribution of outputs characterises the aggregate uncertainty. This level requires specification of probability distributions for all major inputs and is appropriate for high-priority accounts where uncertainty is a policy concern.

¶5 Table 3.7.2: Uncertainty reporting levels

Level	Method	Input Requirements	Output	Recommended For
1	Qualitative assessment	Expert judgement	Low/Medium/High rating per indicator	All accounts (minimum requirement)
2	Confidence intervals	Standard errors for key inputs	95% confidence intervals for aggregates	Priority indicators, headline dashboard
3	Monte Carlo simulation	Probability distributions for all inputs	Full uncertainty distributions	High-stakes policy indicators, valuation accounts

Quality tiers

¶1 Not all data used in ocean accounts meet the standards of official statistics. To accommodate data of varying provenance and reliability whilst maintaining transparency, ocean accounts should classify data inputs into three quality tiers:

¶2 **Tier 1 – Official statistics:** Data produced by national statistical offices or designated official statistics producers, compiled according to the UN Fundamental Principles of Official Statistics. Tier 1 data have been through established quality assurance processes and carry the authority of the national statistical system. Examples include national accounts aggregates, population census data, and fisheries catch statistics compiled by the national statistical office. The quality assurance framework in **TG-0.7 Quality Assurance Principles** describes the quality dimensions applicable to official statistics.

¶3 **Tier 2 – Provisional and administrative data:** Data from government agencies (environmental ministries, fisheries authorities, port authorities) that follow documented collection and processing procedures but have not been formally designated as official statistics. Tier 2 data may be subject to revision and may not meet all quality dimensions of official statistics. Examples include environmental monitoring data, vessel tracking records, and protected area registries. Administrative data guidance is provided in **TG-4.3 Administrative Data Sources**.

¶4 **Tier 3 – Experimental and research data:** Data from research programmes, citizen science initiatives, remote sensing products, and modelled estimates that have been peer-reviewed or validated but are not part of the official statistical system. Tier 3 data are often the only source available for ecosystem condition indicators, ecosystem service estimates, and blue carbon stocks. Research data guidance is provided in **TG-4.5 Research Data**, and citizen science guidance in **TG-4.4 Citizen Science**.

¶5 Combined presentations and dashboards should display the quality tier for each indicator, enabling users to distinguish between well-established indicators based on official statistics and emerging indicators based on experimental estimates. Over time, investment in data systems should aim to progressively elevate key indicators from Tier 3 to Tier 2 and from Tier 2 to Tier 1.

Peer review workflows

¶1 Research data entering the ocean accounts should undergo peer review appropriate to the quality tier. For Tier 3 data, the minimum requirement is that the underlying research has been published in a peer-reviewed journal or technical report, or has been reviewed by a technical advisory committee with relevant domain expertise. For Tier 2 data, the data collection methodology should be documented and reviewed by the national statistical office or an equivalent quality assurance body. For Tier 1 data, the full quality assurance framework of the national statistical system applies.

¶2 Where ocean accounting programmes commission new research or monitoring to fill data gaps, the research design should be reviewed before data collection begins (to ensure the outputs will be compatible with accounting requirements) and the results should be reviewed before incorporation into accounts. The review should assess fitness for purpose (does the data measure what the

account requires?), methodological rigour (is the collection and processing method sound?), and reproducibility (could another team replicate the results?).

3.8 Time-Series Data Investment Guidance

¶1 Sustained monitoring and data collection are essential for compiling the time-series accounts that reveal trends in ocean health, economic activity, and social outcomes. This section provides guidance on identifying critical data gaps, evaluating the cost-benefit of sustained monitoring investments, and prioritising data collection to maximise the analytical value of ocean accounts.

Identifying critical gaps

¶1 A systematic gap analysis should be conducted as part of the initial compilation of ocean accounts, and updated periodically (at least every three years) as data systems evolve. The gap analysis should assess data availability against the full indicator set specified in the combined presentation dashboard (**TG-3.8 Combined Presentations**, Section 3.7) and the thematic account requirements of each relevant Circular.

¶2 For each indicator, the gap analysis should record: the current data source (if any), the quality tier (Section 3.7), the spatial and temporal coverage, the update frequency, the time lag between reference period and data availability, and the estimated cost of producing the indicator. Gaps should be classified into three categories:

¶3 **Complete gaps:** No data source exists for the indicator. Complete gaps are the most serious, as they prevent compilation of entire account components. Common complete gaps in ocean accounting include ecosystem condition indices for deep-sea ecosystems, monetary valuation of regulating services, and governance effectiveness indicators.

¶4 **Partial gaps:** Data exist but with insufficient spatial coverage (e.g., monitoring stations cover only a fraction of the coastline), temporal coverage (e.g., surveys conducted once rather than annually), or thematic detail (e.g., employment data available for the ocean economy as a whole but not disaggregated by industry). Partial gaps can often be addressed through statistical estimation or modelling, but sustained monitoring investment is needed for reliable time-series compilation.

¶5 **Quality gaps:** Data exist with adequate coverage but at an insufficient quality tier (e.g., Tier 3 experimental estimates where Tier 1 official statistics are needed for policy credibility). Quality gaps require investment in data collection methodology, quality assurance processes, or institutional capacity rather than new monitoring infrastructure.

¶6 *Table 3.8.1: Gap analysis template*

Indicator	Current Source	Quality Tier	Spatial Coverage	Temporal Coverage	Update Frequency	Gap Type	Prio
Ecosystem extent (coastal)	Satellite imagery	Tier 2	Full	2018-present	Annual	None	–
Ecosystem condition (coral)	Research surveys	Tier 3	30% of reefs	2020, 2023	Ad hoc	Partial (spatial + temporal)	High
Fish stock biomass	Stock assessment	Tier 2	Commercial species only	Annual	Annual	Partial (thematic)	Medi
Ocean economy GVA	National accounts	Tier 1	National	Annual	Annual	None	–
Ecosystem service valuation	None	–	–	–	–	Complete	High
Coastal community wellbeing	Census proxy	Tier 2	Full	Decennial	Decennial	Partial (temporal)	Medi

Cost-benefit of sustained monitoring

¶1 Investment in sustained ocean monitoring should be evaluated against the analytical value it generates for ocean accounts and the policy decisions it enables. The cost-benefit assessment should consider both the direct costs of data collection (equipment, personnel, survey operations, data processing) and the indirect benefits (improved policy decisions, avoided environmental damage, international reporting compliance).

¶2 Three principles should guide the cost-benefit assessment. The first is marginal value: the value of an additional year of data increases non-linearly with the length of the existing time series. The first five years of a monitoring programme establish baseline conditions, years five to ten reveal short-term trends, and only after ten or more years can long-term trends be distinguished from natural variability. Discontinuing a monitoring programme after a few years therefore sacrifices most of the accumulated investment value. The second is the integration multiplier: data that serve multiple accounts and indicators have higher value per unit cost than data serving a single purpose. Ecosystem extent data, for example, feeds into extent accounts, condition accounts (as a pressure indicator), ecosystem service accounts (as a scaling factor), and governance accounts (as a baseline for protected area assessment). The third is substitutability: some data can only be collected through direct observation (e.g., deep-sea biodiversity surveys), whilst others can be estimated from proxies or models (e.g., wave energy potential from reanalysis products). Investment priority should favour non-substitutable observations.

¶3 *Table 3.8.2: Monitoring investment priority matrix*

Priority Level	Criteria	Examples	Recommended Action
Critical	Non-substitutable; serves multiple accounts; policy-mandated	Fish stock assessments, water quality monitoring, ecosystem extent mapping	Sustain and strengthen; protect from budget cuts
High	Serves multiple accounts; partial substitutes available	Coral reef condition surveys, coastal erosion monitoring, marine debris surveys	Establish sustained programme; explore cost-sharing
Medium	Serves single account; substitutes partially available	Deep-sea biodiversity, offshore air quality, recreational fishing effort	Periodic surveys (3–5 year cycle); supplement with modelling
Low	Substitutable through modelling or proxies	Wave climate, sea surface temperature, chlorophyll-a concentration	Rely on global products (satellite, reanalysis); validate periodically

Priority investments

¶1 Based on the gap analysis and cost-benefit assessment, ocean accounting programmes should develop a prioritised data investment plan. The investment plan should be aligned with the national statistics development strategy and the broader ocean governance priorities. Priority investments for most countries implementing ocean accounts will typically include:

¶2 Establishing **annual ecosystem extent monitoring** using satellite remote sensing, which provides the foundational spatial data for extent accounts, condition assessments, and ecosystem service estimation. The relatively low marginal cost of processing freely available satellite imagery (Sentinel-2, Landsat) makes this a high-value investment. Guidance on remote sensing methods is provided in **TG-4.1 Remote Sensing and Geospatial Data**.

¶3 Strengthening **coastal water quality monitoring** networks to provide consistent, spatially representative measurements of nutrients, sediments, and pollutants. Water quality data serve the residual flow accounts (**TG-3.4**), ecosystem condition accounts, and the water-marine quality integration protocol described in **TG-3.8 Combined Presentations**, Section 3.8.

¶4 Developing **ocean economy satellite accounts** within the national accounts framework, enabling consistent measurement of ocean economy GVA, employment, and trade. This requires collaboration between the national statistical office and sectoral ministries to agree on the scope of ocean-related industries and compile thematic accounts following **TG-3.3 Economic Activity Relevant to the Ocean**.

¶5 Investing in **sub-national data disaggregation** to support provincial and municipal ocean accounting, as described in **TG-3.8 Combined Presentations**, Section 3.9. Many national datasets can be disaggregated to sub-national level at relatively low additional cost if the disaggregation requirement is built into data collection design from the outset.

Link to TG-4.7

¶1 The data investment priorities identified in this section should be integrated into the broader capacity development strategy described in **TG-4.7 National Data Coordination Architectures**. Sustained monitoring requires institutional capacity as well as financial investment: trained personnel, maintained equipment, quality assurance systems, and data management infrastructure. The capacity development Circular provides guidance on building and maintaining these institutional foundations, including workforce planning, training programmes, and partnerships with research institutions and international organisations. Data investment plans should be reviewed and updated alongside the capacity development strategy to ensure that monitoring ambitions are matched by implementation capacity.

4. Summary

¶1 Data harmonisation and interoperability underpin the integration of diverse data sources into coherent accounting frameworks. The main elements are:

- ¶2 **SDMX** provides the statistical data exchange standard, with DSDs defining data structures, transmission formats enabling machine-readable exchange, and registry services supporting data discovery
- ¶3 **Geospatial standards** from OGC and IHO S-100, supported by ISO 19115 metadata standards, enable integration of marine spatial data, with the GSGF providing the overarching framework for statistical-geospatial integration
- ¶4 **FAIR principles** guide data management to maximise findability, accessibility, interoperability, and reusability
- ¶5 **Classification concordances** enable translation between ISIC/CPC for economic data, between national and international ecosystem classifications, and across classification versions over time, with official UNSD correspondence tables available for the ISIC Rev.5 and CPC Ver.3.0 transitions
- ¶6 **Data harmonisation workflow** provides systematic procedures for classifying, mapping, validating, and integrating heterogeneous source data into ocean accounts

¶7 Implementation of these standards requires institutional investment in technical infrastructure, staff capacity building, and sustained collaboration across organisational boundaries. The GSGF notes that “both the statistical and geospatial communities operate their own general data models, metadata capabilities, architectures, and data infrastructures” and that bridging these requires “greater incorporation of geospatial processes, standards, and best practices in the statistical business processes and data management systems”⁴⁹.

¶8 For ocean accounting practitioners, a pragmatic approach involves:

1. ¶9 Adopting SDMX-CSV as an accessible intermediate format whilst building toward full SDMX-ML/JSON implementation

2. ¶10 Ensuring geospatial data conforms to OGC standards and carries ISO 19115-compliant metadata to enable integration with national spatial data infrastructures
 3. ¶11 Documenting all classification concordances used, including rationale and limitations, and using official UNSD correspondence tables for transitions between ISIC and CPC versions
 4. ¶12 Implementing FAIR principles progressively, prioritising persistent identifiers and standardised metadata
 5. ¶13 Establishing data harmonisation workflows with systematic quality checks at each phase
- ¶14 All ocean accounts benefit from harmonised data infrastructure. Asset accounts (**TG-3.1 Asset Accounts**) integrate physical and monetary data from diverse sources, requiring consistent classifications and units. Ocean economy investment accounts (**TG-2.6 Ocean Economy Investment**) depend on concordances between economic activity and product classifications. Combined presentations (**TG-3.8 Combined Presentations**) integrate environmental, economic, and social data, relying on interoperable formats and metadata standards. Data harmonisation and interoperability are therefore foundational to the ocean accounting framework.
- ¶15 Guidance on specific data integration challenges for individual account types is provided in the relevant thematic Circulars: **TG-4.1 Remote Sensing and Geospatial Data**, **TG-4.2 Survey Methods for Ocean Economic Activity**, **TG-4.3 Administrative Data Sources**, **TG-4.4 Citizen Science**, and **TG-4.5 Research Data**.

Implementation Considerations

¶1 For minimum institutional capacity, data infrastructure, and human skills requirements for implementing these data methods, see [TG-0.8 Implementation Readiness Assessment](#).

5. Acknowledgements

¶1 **Authors:** [To be confirmed]

¶2 **Reviewers:** [To be confirmed]

6. References

1. ¶1 Figure 4.6.1 illustrates the architecture by which data from statistical (SDMX), geospatial (GSGF/OGC), and marine (IHO S-100) standards, together with classification concordances, converge through an integration layer to produce ocean accounts.
2. ¶2 Statistical Data and Metadata Exchange (SDMX), "SDMX Standards: Section 1 - Framework for SDMX Technical Standards, Version 3.1" (May 2025), para 1.
3. ¶3 SDMX Framework Section 1, para 1. ISO 17369:2013 - Statistical data and metadata exchange (SDMX).
4. ¶4 SDMX Framework Section 1, Section 2.3, para 211-219. Major changes include support for geospatial data, micro-data, code list extension, and structure mapping improvements.
5. ¶5 SDMX, "SDMX Standards: Section 2 - Information Model: UML Conceptual Design, Version 3.1" (May 2025).
6. ¶6 SDMX Framework Section 1, Section 3.1, para 293-298.
7. ¶7 SDMX Framework Section 1, Section 5.
8. ¶8 SDMX Framework Section 1, Section 5.3, para 508.
9. ¶9 SDMX Framework Section 1, Section 3.1, para 300.
0. ¶10 United Nations Expert Group on the Integration of Statistical and Geospatial Information, "Global Statistical Geospatial Framework (GSGF), Version 2.0" (2025).
1. ¶11 GSGF v2, Principle 4: Statistical and geospatial interoperability in data standards, processes and organizations.
2. ¶12 GSGF v2, Principle 4, Definition section.
3. ¶13 European Commission, "European Interoperability Framework (EIF)".
4. ¶14 GSGF v2, Principle 4.
5. ¶15 Open Geospatial Consortium, Standards and Resources. See <https://www.ogc.org/standards/> (<https://www.ogc.org/standards/>)
6. ¶16 ISO 19115-1:2014 Geographic information – Metadata – Part 1: Fundamentals. See <https://www.iso.org/standard/53798.html> (<https://www.iso.org/standard/53798.html>)

7. ¶17 ISO/TS 19139:2007 Geographic information – Metadata – XML schema implementation.
8. ¶18 GSGF v2, Principle 4, Technology & Infrastructure section.
9. ¶19 International Hydrographic Organization, “S-100 – Universal Hydrographic Data Model, Edition 5.2.0” (June 2024).
0. ¶20 IHO S-100, Foreword.
1. ¶21 IHO, “Major Milestone Achieved in Transition to Smart Navigation with Operational Editions of S-100 Standards” (December 2024). Operational editions approved include S-101 (ENCs), S-102 (Bathymetric Surface), S-104 (Water Level Information), and S-111 (Surface Currents).
2. ¶22 IHO S-100, Foreword.
3. ¶23 IHO S-100, Part 0, Section 0-4.
4. ¶24 IHO S-100, Introduction.
5. ¶25 SDMX Framework Section 1, Section 5.
6. ¶26 SDMX Framework Section 1, Section 2.3 “Major Changes from 2.1 to 3.0”, under “XML, JSON, CSV and EDI Transmission formats”.
7. ¶27 SDMX Framework Section 1, Section 5.2.
8. ¶28 SDMX Framework Section 1, Section 5.3, para 508.
9. ¶29 OGC, “OGC Features and Geometries JSON (JSON-FG)”. See <https://github.com/opengeospatial/ogc-feat-geo-json> (<https://github.com/opengeospatial/ogc-feat-geo-json>)
0. ¶30 ISO 19136:2007 Geographic information – Geography Markup Language (GML).
1. ¶31 IHO S-100, Part 10c – HDF5 Data Model and File Format.
2. ¶32 SDMX Framework Section 1, Section 3.6; SDMX, “SDMX Standards: Section 6 – Technical Notes, Version 3.1”, Sections 9-11.
3. ¶33 SDMX Framework Section 1, Section 3.5.
4. ¶34 OGC Sensor Observation Service. See <http://www.opengeospatial.org/standards/sos> (<http://www.opengeospatial.org/standards/sos>)
5. ¶35 Wilkinson, M. D., et al. (2016). “The FAIR Guiding Principles for scientific data management and stewardship.” Scientific Data 3, 160018. <https://doi.org/10.1038/sdata.2016.18> (<https://doi.org/10.1038/sdata.2016.18>)
6. ¶36 GSGF v2, Principle 4. See also GO FAIR initiative, <https://www.go-fair.org/fair-principles/> (<https://www.go-fair.org/fair-principles/>)
7. ¶37 United Nations, “Central Product Classification (CPC) Version 2.1” (2015), Part One, Chapter IV.C, para 14.
8. ¶38 CPC Ver.2.1, Preface.
9. ¶39 CPC Ver.2.1, Part One, Chapter II.A, para 21.
0. ¶40 United Nations Statistical Commission, 55th Session (27 February – 1 March 2024). ISIC Rev.5 and CPC Ver.3.0 explanatory notes endorsed. ISIC Rev.5 maintains the division-level structure for fishing and aquaculture (Division 03) and water transport (Division 50) with increased detail at lower classification levels.
1. ¶41 United Nations Statistics Division, “Technical Note on the ISIC, Rev.4 – Rev.5 Correspondence Table” (2024). Available at <https://unstats.un.org/unsd/classifications/Econ> (<https://unstats.un.org/unsd/classifications/Econ>)
2. ¶42 United Nations, “System of Environmental-Economic Accounting – Ecosystem Accounting” (2021), Chapter 3.
3. ¶43 United Nations, “Guidelines on Biophysical Modelling for Ecosystem Accounting” (2022), para 93.
4. ¶44 Guidelines on Biophysical Modelling, para 92.
5. ¶45 United Nations, “System of Environmental-Economic Accounting 2012 – Central Framework” (2014), para 5.86 and footnote 56.
6. ¶46 United Nations, “International Standard Industrial Classification of All Economic Activities (ISIC), Revision 4” (2008), Historical Background.
7. ¶47 UNSD, “Technical Note on the ISIC, Rev.4 – Rev.5 Correspondence Table” (2024). The correspondence table was developed between March and October 2024 following the endorsement of ISIC Rev.5 at the 55th Session of the UN Statistical Commission.
8. ¶48 UNSD Classifications on Economic Statistics. Correspondence tables for CPC Ver.2.1 to Ver.3.0, ISIC Rev.5, and HS 2022 are available at <https://unstats.un.org/unsd/classifications/Econ> (<https://unstats.un.org/unsd/classifications/Econ>)
9. ¶49 GSGF v2, Principle 4.