



Management of OSPAR data

OSPAR Agreement 2024-03¹

Data and information management at OSPAR is becoming increasingly complex. This document outlines the overarching structures and processes of OSPAR data management activities to increase transparency and effectiveness of the process and help structure future work.

The integrated nature of the framework and documentation that underpin OSPAR's data related activities is complex. This document looks to outline all parts of the process so that, irrespective of the point of access, users can more easily access and understand data management at OSPAR.

This document also reduces the chances of work being repeated by facilitating the understanding of existing, standardised content and templates, by making them clearly available and accessible. This document and its associated sub-documents are part of an ongoing process to refine the available information of each part of the data management process to minimise gaps and promote standardisation by ensuring all areas are clearly documented to meet the standards expected by OSPAR for data activities.

The document serves as guidance for those involved in any aspect of OSPAR data and information handling, from development of new datastreams and the drafting of data calls, through to data processing and publication of assessments with associated data products, to end users looking for clarity on the OSPAR process.

¹ English only

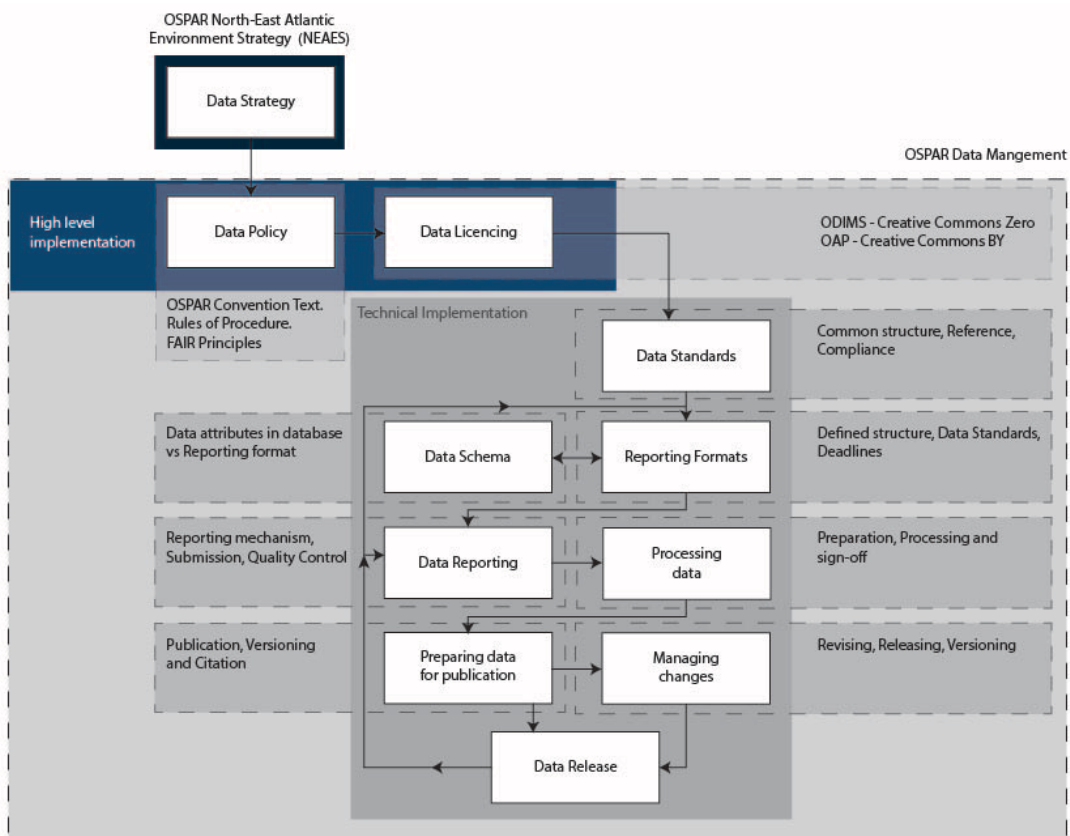
Contents

1. Background	3
2. Data Policy.....	3
OSPAR Convention text.....	4
OSPAR Rules of Procedure	4
3. OSPAR Data and Information Licencing.....	6
4. How OSPAR data are structured	6
5. CEMP Guidelines	6
6. Reporting Formats.....	7
7. Defining reporting formats	7
8. Data schema.....	7
9. Data standards	7
10. Deadlines.....	8
11. How data are reported	8
12. Reporting procedure for data	8
13. Re-submission deadline	9
14. Processing of data	9
15. Data approval.....	9
16. Preparing Data for Publication.....	9
17. Versioning.....	10
18. Improving OSPAR data.....	10
19. FAIR Principles.....	11
20. Implementation of the FAIR Principles	12

1. Background

a. The figure below gives context on the various elements covered in this document. The high-level data strategy provides the steer for all underlying activities and is derived from the North-East Atlantic Environment Strategy 2030 (NEAES 2030):

“We will ensure that data collection and assessment programmes are kept under continuous review, so they are up to date and fit for purpose as both technology and our understanding of the marine environment develop. Monitoring and assessment of the marine environment require the effective use and management of data and information to support the production of robust assessments. This will be achieved through the OSPAR Data and Information Management System (ODIMS) and the OSPAR Assessment Portal (OAP), allowing links to be made with other providers and consumers of OSPAR data and information. We are committed to ensuring that the data we use are findable, accessible, interoperable, reusable and reproducible.”



2. Data Policy

a. The OSPAR Data Policy is taken from the text of the OSPAR Convention and the Rules of Procedure; exact texts are extracted below and referenced. The OSPAR Convention² provides detail on access to information (Article 9) and data submission (Annex IV, Article 2):

² <https://www.ospar.org/convention/text>

*OSPAR Convention text**Article 9 - Access to information:*

1. The Contracting Parties shall ensure that their competent authorities are required to make available the information described in paragraph 2 of this Article to any natural or legal person, in response to any reasonable request, without that person's having to prove an interest, without unreasonable charges, as soon as possible and at the latest within two months.
2. The information referred to in paragraph 1 of this Article is any available information in written, visual, aural or data-base form on the state of the maritime area, on activities or measures adversely affecting or likely to affect it and on activities or measures introduced in accordance with the Convention.
3. The provisions of this Article shall not affect the right of Contracting Parties, in accordance with their national legal systems and applicable international regulations, to provide for a request for such information to be refused where it affects: (a) the confidentiality of the proceedings of public authorities, international relations and national defence; (b) public security; (c) matters which are, or have been, sub judice, or under enquiry (including disciplinary enquiries), or which are the subject of preliminary investigation proceedings; (d) commercial and industrial confidentiality, including intellectual property; (e) the confidentiality of personal data and/or files; (f) material supplied by a third party without that party being under a legal obligation to do so; (g) material, the disclosure of which would make it more likely that the environment to which such material related would be damaged.
4. The reasons for a refusal to provide the information requested must be given.

Annex IV - On the Assessment of the Quality of the Marine Environment, Article 2

- a) For the purposes of this Annex, the Contracting Parties shall: (a) cooperate in carrying out monitoring programmes and submit the resulting data to the Commission.

OSPAR Rules of Procedure

- b. The OSPAR Rules of Procedure³ formally govern the work of the OSPAR Commission. The relevant detail for data and information is in Annex 3 of the Rules of Procedure:

OSPAR Data Release Arrangements

- a. OSPAR is committed to making as much information as possible publicly available, consistent with achieving other similarly important goals of public policy;
- b. OSPAR and its Contracting Parties wish to collaborate to the greatest possible extent with other agencies working in the field of monitoring and observing the marine environment.;
- c. Data-handling arrangements should be in place to ensure properly documented/quality controlled and comparable datasets and made available for those needing them; *“assist the maintenance of sound, comprehensive, high-quality, accessible data banks, which can be relied on for their accuracy and integrity”*;
- d. Data-handling arrangements should also make efficient use of resources and be clear and transparent, while protecting the privacy and confidentiality of individuals and commercial interests.

³ <https://www.ospar.org/documents?d=34012>

Collection and handling of data

- a. The OSPAR monitoring programmes rely upon data derived from publicly funded monitoring by OSPAR Contracting Parties;
- b. OSPAR will ensure that its specifications of programmes for collecting and evaluating data on the state of the marine environment, on the activities and measures which can affect it and on the activities and measures adopted under the Convention make proper provision for: a. the design of monitoring activities; b. reporting on the collection of data, including the identification of a thematic data centre to hold and manage the data; c. the documentation, quality control and building of comprehensive data sets.
- c. For the following programmes, these elements are set out in the following OSPAR Agreements (as they may be amended from time to time by agreements of OSPAR), and the following bodies (subject to any changes that may be made by ordinary agreements of OSPAR) fulfil the role of thematic data centre:
 - i. Comprehensive Atmospheric Monitoring Programme (CAMP) the Principles for the Comprehensive Atmospheric Monitoring Programme (Agreement 2001-7); thematic data centre: The Norwegian Institute of Air Research (NILU);
 - ii. Coordinated Environmental Monitoring Programme (CEMP) the Coordinated Environmental Monitoring Programme (Agreement 2004-16) and the Requirements for the Submission of National Comments to ICES when Submitting Monitoring Data (Agreement 2003-11); thematic data centre: International Council for the Exploration of the Sea (ICES);
 - iii. Comprehensive Study of Riverine Inputs and Direct Discharges (RID) the Principles for the Comprehensive Study of Riverine Inputs and Direct Discharges (Agreement 1998-5); thematic data centre: the OSPAR Secretariat.
- d. For future OSPAR monitoring products specified in the Joint Assessment and Monitoring Programme, the OSPAR agreements on the form, timing and development of these products will ensure that these elements are adequately specified, together with the identification of a thematic data centre.

Release of data

- a. There are different data release arrangements depending on the datastream. It was understood that the chosen data release arrangement would be chosen upon definition of the datastream, see §e below;
- b. CAMP and RID data, freely available following publication of the corresponding annual reports;
- c. CEMP data available following quality control;
- d. Meteorological/Hydrographic data are not directly available where restricted by national legislation or international agreement;
- e. When establishing future datastreams, the corresponding OSPAR Agreement shall specify the agreement for data release in line with the CEMP, CAMP or RID options;
- f. When using data from OSPAR thematic data centres, Contracting Parties will ensure that data are treated consistently, irrespective of source. This also applies to quality assurance procedures;
- g. Where an OSPAR product is based on data with multiple levels of restriction the most stringent level of restriction shall be applied to all basic data on which the product is based.

OSPAR documents & published material

The arrangements do not apply to:

- a. data contained in documents to be presented to OSPAR or subordinate bodies, (governed by rules 56-60), or;
- b. data and information obtained from scientific literature or published reports that are already in the public domain (in line with copyright rules).

3. OSPAR Data and Information Licencing

a. OSPAR 2019 noted the outcome of JL2019 (OSPAR SR §9.1) which included the advice from JL on the revision of the OSPAR Conditions of Use (OSPAR 9/1 §12): *“JL advised that the OSPAR Data and Information Management System (ODIMS) should make use of the Creative Commons Zero (CC-0) licence and the OSPAR Assessment Portal should make use of the Creative Commons BY (CC-BY) licence.”*

- i. OSPAR data are made available with the least amount of restriction possible. Data made available via the OSPAR Data and Information System (ODIMS – odims.ospar.org) are bound by Creative Commons 0 1.0 Universal Public Domain Dedication (CC0 - <https://creativecommons.org/publicdomain/zero/1.0/>). This means the data can be copied, modified, distributed and used for commercial purposes without asking permission or requiring citation.
- ii. Information made available via the OSPAR Assessment Portal (OAP – oap.ospar.org) are bound by Creative Commons Attribution 4.0 International (CC-BY – <https://creativecommons.org/licenses/by/4.0/>). This means the information presented can be copied and redistributed in any medium or format, but must be given appropriate credit and any changes made must be indicated. There is no inferred endorsement of resultant products.

4. How OSPAR data are structured

a. OSPAR data requirements are developed following a documented structure. More information, as well as published documents on the Coordinated Environmental Monitoring Programme (CEMP) and the associated CEMP Guidelines can be found here: <https://www.ospar.org/work-areas/cross-cutting-issues/cemp>

5. CEMP Guidelines

a. Each datastream has a set of defined requirements, covering the purpose, objectives, and assessment process. These requirements should be captured in the appropriate CEMP Guideline, following the template.

b. The CEMP Guideline has a section dedicated to data related requirements, “Data reporting, handling and management”. These are defined by the experts and include reporting format and guidance, metadata schema, data confidence and data flows.

6. Reporting Formats

a. The defined reporting format from the CEMP Guideline requires an element of looking forward, considering the data flows and should consider any further onward obligations the reported data could respond to. This ensures the relevant functionality exists at the start point, to achieve all of the objectives for each datastream. The following are considerations that need to be made through the process:

7. *Defining reporting formats*

a. Reporting formats for datastreams will be drafted by expert groups and will be approved by the Committee or other subsidiary body that is responsible for issuing the data call. Consideration is to be given to the datastream's position within the OSPAR data ecosystem as a whole to enable it to be correctly managed to achieve OSPARs data goals.

b. The reporting format is an explicit data structure to be used when reporting data. Associated guidance for completing the structure must also be made available when issuing a data call to increase clarity and maximise the chances of success for data reporters.

c. Experts drafting a reporting format are responsible for checking compliance with existing data standards and aligning with OSPAR licencing expectations.

d. If the data gathered under a data call will not be made openly available (ref. OSPAR Convention, Article 9, §3), the implications of this must be considered, as sharing without limitation is the default position. The consideration of an anonymised data product or a timed delay in publication could be suitable responses to ensure data continue to be made available. Having data exposed via ODIMS increases the visibility and interoperability of datasets. If data are not to be made openly available, they will be handled through a different process to ensure that data are kept private. It might be that updates to content are required further down the chain of data processing and presentation, so a thorough consideration of the data flow is important.

8. *Data schema*

a. This is a clear definition of how data are to be structured and the associated field types/lengths. There can be differences between the way data are collected in a reporting format and how they are then stored within a database or compiled data environment; this is usually reflected in field naming due to best practices for database management. The individual elements of the data schema should be available for parameters that are covered by existing parameters outlined in the data standards.

9. *Data standards*

a. The common set of OSPAR data standards are being developed based on changing data requirements and refinement against existing standards. The following presents elements to be considered when developing data requirements and reporting formats:

- i. data collection parameters to facilitate presentation and reuse of data, e.g., latitude and longitude should be collected in Decimal Degrees, using CRS WGS84.
- ii. application of existing data standards applicable to the topic.
- iii. adaptations outside the remit of existing data standards should be avoided insofar as possible as the standards form the basis for data storage and dissemination and making changes could have a wide-reaching impact. Missing elements need to be carefully documented and reported.

- iv. additions to data standards can be considered. If a new element is proposed, then retrospective amendments and/or other impacts should be outlined when the proposal is made. This also applies when updating reporting formats as it can make historic data redundant and limit trend assessments over time.
- v. reporting formats should include notes for handling 'missing information, e.g., NI, ND, N/A, and zero values. A definition of each field with example data and a glossary of terms can facilitate this.

10. Deadlines

- a. Deadlines for reporting are agreed by each Committee or responsible subsidiary body, typically at the meeting preceding the year of data delivery. For regular, annual data calls the deadline date is often kept the same year after year to support the national planning actions needed by Contracting Parties to respond to a call.
- b. If a deadline is being frequently missed, consideration should be made to changing the deadline to facilitate ongoing work that may be impacted by late delivery. It is the responsibility of the Committees to review data reporting using the data reporting updates presented by the Secretariat in the annual paper on the status of data reporting to Heads of Delegation, which is also available in real-time via the OSPAR Data Submission site.
- c. Consideration of data sharing requirements and how deadlines could impact other data requirements must be considered, along with coordination with the overall OSPAR timeline. Examples include *"Practical Arrangements between the International Atomic Energy Agency and the OSPAR Commission on sharing data on concentrations of radioactive substances in the OSPAR Maritime Area"* (Agreement 2009-06⁴) and "Dumping and placement of waste or other matter at sea" data and their onward sharing with the London Convention.

11. How data are reported

- a. Consideration of how data are reported and handled are important to limit potential introduction of errors. A step forward was made in this area with the introduction of the OSPAR Data Submission SharePoint site⁵. This allows Contracting Parties to have a single-point of access for reporting in response to OSPAR related data calls. The site continues to be refined and developed based on user feedback.

12. Reporting procedure for data

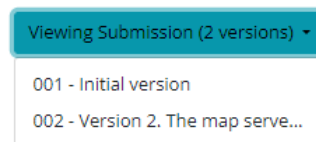
- a. All data should be reported in accordance with the corresponding data call and associated reporting format.
- b. All data from OSPAR-managed datastreams will be reported using the OSPAR Data Submission site.
- c. Each time data are submitted a corresponding reporting record is completed to log when each reporting was received and other pertinent information. This information is then used as the basis of the 'Status of data reporting' document.

⁴ <https://www.ospar.org/documents?d=32814>

⁵ <https://osparcsp.sharepoint.com/sites/datasubmission>

13. Re-submission deadline

- a. In order for the whole process to flow, reporting deadlines must be adhered to. The data team aim to begin work on data processing or distribute data to Expert Assessment Panels (EAPs) within the month following the agreed deadline. If timings allow, efforts will be made to include late data, but it is not always possible and is a discretionary decision by those processing the data.
- b. Reporting timelines will be developed for each Committee to provide clarity on the process each datastream goes through. This should include the request date when the data calls are issued to Contracting Parties; the Contracting Party reporting deadline; the window for the submission of revisions; date of upload (depending on approval processes for each datastream, see [§15. Data approval](#)); and publication date.
- c. There is a mechanism, utilising the versioning control function on ODIMS, to update previously published data and keep all versions. This also allows historic datasets to be supplemented with previously missing information in a traceable fashion:



Screen shot from <https://odims.ospar.org/>

14. Processing of data

- a. Each datastream managed by the Secretariat can have a Data Processing Routine document associated with it. There is a data processing routine template presented in sub-document *OSPAR_Data_Processing_Routine_Template-vX.docx*⁶. The document captures information on the following:
 - i. Reporting deadline/Processing deadline;
 - ii. Links to reporting format and guidance for reference;
 - iii. Folder structure, for local data management; and
 - iv. Step-by-step process to check and ensure the reported data adhere to the reporting format.

15. Data approval

- a. Each datastream has an agreed approval procedure which ties back to the Rules of Procedure. *OSPAR_datastream_breakdown_vX.docx*⁶ has a field entitled "Data approval procedure" outlines when data can be made available. This sub-document responds to the task of ensuring that there are no gaps in information in each datastream process.

16. Preparing Data for Publication

- a. These are the steps that data must undergo following delivery to the data managers. This considers the processing, approval and publication steps. Data are uploaded to ODIMS, as appropriate, following approval for publication, in accordance with the data release arrangements. There is a degree of preparation involved in making the data available, following approval for it to be published,

⁶ These are internal/working documents that may be periodically published via: <https://www.ospar.org/work-areas/cross-cutting-issues/data-and-information>

this is outlined in sub-document [OSPAR_preparing-data-for-publication_vX.docx](#)⁶. The document captures the technical steps required to prepare the data as a submission package for ODIMS with associated styling and metadata files.

17. Versioning

a. The role of the versioning functions of ODIMS and OAP are to allow for a running catalogue of content to be maintained with changes and updates being recorded. It allows persistent references to specific OSPAR published content to remain so those making a reference can do so in the knowledge that the link will remain indefinitely. Submissions to ODIMS should not be deleted once they have been approved; in cases where data are ‘wrong’ and this could be to the detriment of those using it, it is possible to ‘reject’ the submission, this ‘hides’ the content. The number of the rejected version will persist but will appear as a gap in the timeline, e.g., 1, 2, 4, 5. This will allow users to request the ‘missing’ data if required and an explanation of the known issues given, but they will not otherwise be readily available.

b. Additional care will be required when forming the submissions for upload to avoid multiple versions based on input error from administrators.

c. Note that the versioning function allows two different types of links to a dataset/assessment; a link that will always reference the latest dataset (a general link to the content), for example:

i. ODIMS: https://odims.ospar.org/en/submissions/ospar_encounters_2019_01/

ii. OAP: <https://oap.ospar.org/en/ospar-assessments/committee-assessments/hazardous-substances-and-eutrophication/mercury-losses-chlor-alkali-industry/mercury-losses-chlor-alkali-industry-2017/>

and a link to a specific version of a dataset/assessment:

iii. ODIMS: https://odims.ospar.org/en/submissions/ospar_encounters_2019_01_002/

iv. OAP: <https://oap.ospar.org/en/versions/mercury-losses-chlor-alkali-industry-2017-en-1-0-0/>

d. Even though data are completely open to reuse without citation, often a citation is requested. A template data citation is offered via ODIMS https://odims.ospar.org/en/data_policy/:

Data Source: OSPAR Data and Information Management System,
<https://odims.ospar.org/>(add direct link to dataset or metadata)

e. OAP licencing requires that content is cited, and an example citation is provided via <https://oap.ospar.org/en/data-policy/> following the Title, Author, Source, License (TASL⁷) method:

Example attribution: [Page Title], OSPAR Commission, OSPAR Assessment Portal
[https://oap.ospar.org/en/\[add direct link to content page\]](https://oap.ospar.org/en/[add direct link to content page])), licenced under
<https://creativecommons.org/licenses/by/4.0/>

18. Improving OSPAR data

a. Responding to the North-East Atlantic Environment Strategy, and particularly the commitment “to ensuring that the data we use are findable, accessible, interoperable, reusable and reproduceable”

⁷ https://wiki.creativecommons.org/wiki/best_practices_for_attribution

or FAIR⁸. The Secretariat has begun the process of periodically reviewing OSPAR data against the FAIR principles⁹ to measure how well exposed and accessible OSPAR published content is.

19. FAIR Principles

a. The FAIR Data Principles are intended to improve “**F**indability, **A**ccessibility, **I**nteroperability, and **R**euse of digital assets” and “*apply to metadata, data, and supporting infrastructure... Most of the requirements for findability and accessibility can be achieved at the metadata level. Interoperability and reuse require more efforts at the data level.*”¹⁰.

The high-level principles:

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

b. It should be noted that under these high-level principles there are detailed qualifiers and that these can be delivered through other OSPAR work, such as the provision of INSPIRE compliant metadata:

Findable:

- F1. (Meta)data are assigned a globally unique and persistent identifier
- F2. Data are described with rich metadata (defined by R1 below)
- F3. Metadata clearly and explicitly include the identifier of the data they describe
- F4. (Meta)data are registered or indexed in a searchable resource

Accessible:

- A1. (Meta)data are retrievable by their identifier using a standardised communications protocol
 - A1.1 The protocol is open, free, and universally implementable
 - A1.2 The protocol allows for an authentication and authorisation procedure, where necessary
- A2. Metadata are accessible, even when the data are no longer available

Interoperable:

- I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (Meta)data use vocabularies that follow FAIR principles
- I3. (Meta)data include qualified references to other (meta)data

⁸ Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>

⁹ <https://www.go-fair.org/fair-principles/>

¹⁰ <https://www.go-fair.org/fair-principles/fairification-process/>

Reuseable:

- R1. (Meta)data are richly described with a plurality of accurate and relevant attributes
 - R1.1. (Meta)data are released with a clear and accessible data usage license
 - R1.2. (Meta)data are associated with detailed provenance
 - R1.3. (Meta)data meet domain-relevant community standards

20. *Implementation of the FAIR Principles*

- a. By implementing the FAIR guiding principles, the goal is the exposure of content in a machine-to-machine structure or format requiring minimal human intervention, allowing OSPAR digital assess to be reused to their full extent.
- b. Measuring the “FAIRification” process is an area that continues to be refined and developed. Recent publications, such as “FAIR in action – a flexible framework to guide FAIRification”¹¹ can help by giving further structure and highlighting additional resources in the process of implementing and measuring progress toward achieving and maintaining the goals of the NEAES Strategy to 2030.
- c. Measurement against the categories is a useful activity to take stock of how certain actions feedback into the overall process and ensure a positive overall outcome for OSPAR data and information. The Secretariat will therefore periodically assess progress in responding to each of the principles through the process of achieving FAIRification and look to provide a qualitative measure of compliance. Details in sub-document *OSPAR_FAIR_principles_compliance_vX.docx*⁶

¹¹ Welter, D., Juty, N., Rocca-Serra, P. *et al.* FAIR in action - a flexible framework to guide FAIRification. *Sci Data* **10**, 291 (2023). <https://doi.org/10.1038/s41597-023-02167-2>