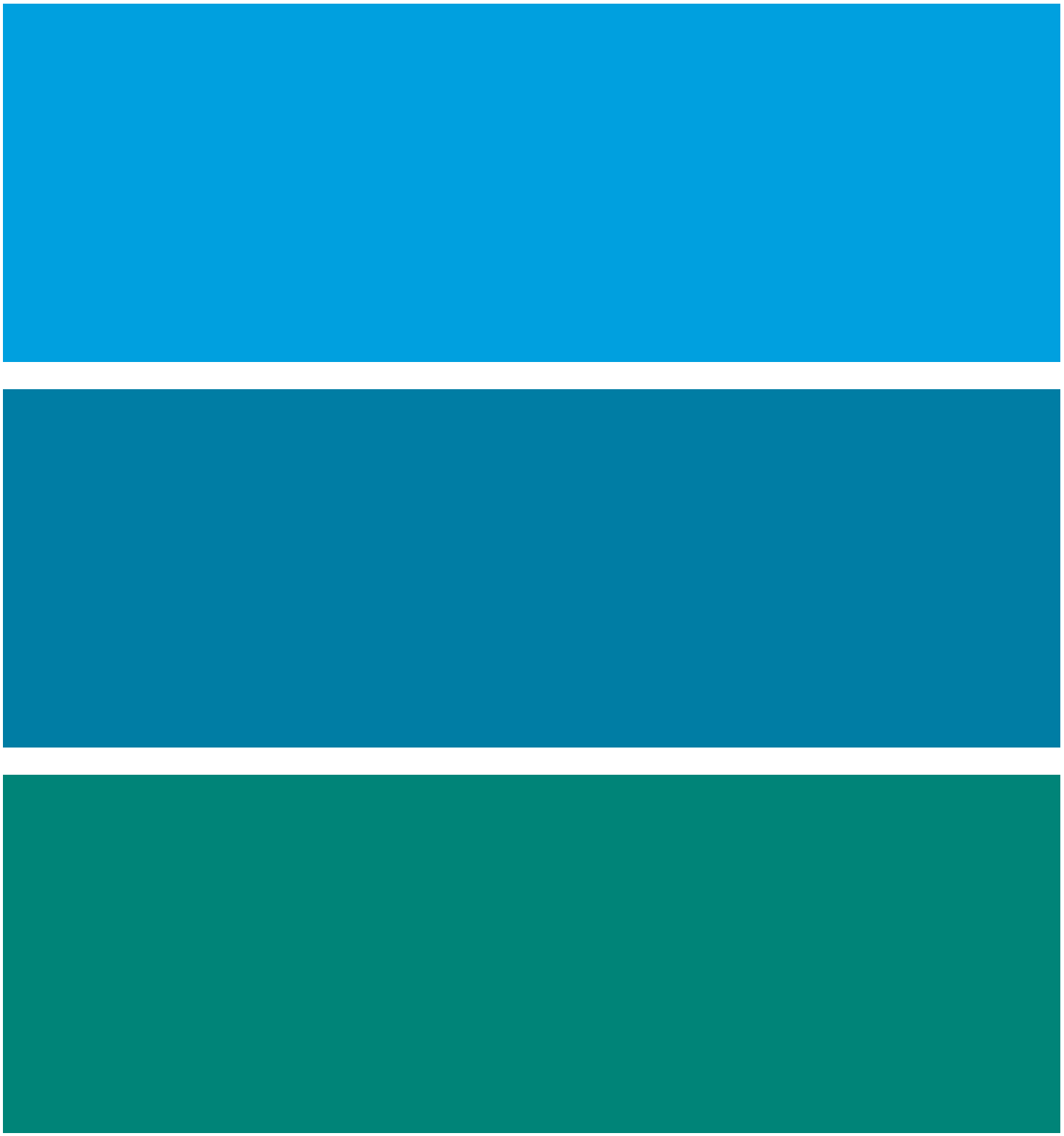
THE CROWN ESTATE

Requirements for providing survey data to The Crown Estate via the Marine Data Exchange



Requirements Note Version 1.4, January 2016

Printed versions of this document are not controlled. The latest electronic version can be provided upon request.

Contents

For more information please contact the data management team at The Crown Estate:
datasupport@thecrownestate.co.uk / Tel 020 7851 5349

Document Control

Version	Date	Author	Reviewed by	Description of change
1.0	July 2012	Pete Edmonds	Round 3 Data Workshop	Baseline version
1.1	September 2012	Pete Edmonds	None	Changed terminology from “Project” to “Series”
1.2	November 2012	Pete Edmonds	None	Minor edits and corrections
1.3	January 2016	Chelsea Bradbury	Pete Edmonds	Addition of data guidelines for meteorological masts and remote sensing devices; greater detail about the confidentiality process; MEDIN metadata standard elements and metadata comparison table updated; example data structures for specific surveys inserted
1.4	February 2016	Chelsea Bradbury	Pete Edmonds	Minor edits and corrections

datasupport@thecrownestate.co.uk
Tel. 0207 851 5000

1. Executive Summary

- The Marine Data Exchange (MDE) is the system used to store, manage and disseminate offshore survey data provided to The Crown Estate from the renewables industry. A set of consistent standards and requirements ensure that this system is maintained to a high level with enhanced discoverability of data.
- The data provided must be structured into a hierarchy of "Series", "Datasets" and "Reports" according to Figure 1. If providing data on physical media, this structure should also be used.
- A series can be defined as a group of datasets and reports collected in relation to a survey or data collection campaign, linked by a common research question. A series must be described by a metadata record and must contain all reports and datasets related to that survey or campaign, including report appendices and GIS shapefiles where applicable.
- A dataset is defined as a 'pillar' of data and information related to a *single* parameter, instrument or method. Datasets must be described by a metadata record. A series can contain one or more datasets.
- A report is defined as a set of documents relating to part, or the whole of a survey or data collection campaign. Final versions rather than drafts should be provided and reports must be described by a metadata record. A series can contain one or more reports.
- The Crown Estate's requirements for metadata are:
    - Series – MEDIN Discovery Standard
    - Dataset – MEDIN Discovery Standard
    - Report – The Crown Estate Report Metadata
- Metadata is used to return search results. Therefore, to improve discoverability, The Crown Estate require all mandatory elements to be completed, with emphasis upon an accurate title, abstract, temporal and spatial extent, and a minimum of three keywords (one from each of the following controlled vocabularies: MEDIN Parameter Discovery Vocabulary, TCE Renewable Energy Development Areas, TCE Renewable Energy Development Phases).
- Where data are spatially referenced, their coordinate reference system should be fully defined, and where transformation parameters do not exist (e.g. a bespoke engineering grid), transformation parameters between the coordinate reference system and WGS 1984 (EPSG code: 4326) should be provided.
- Where appropriate, The Crown Estate will use the MEDIN data guidelines as a pro forma for quality control of data received. This means that, as a minimum, The Crown Estate expects that data provided contain all of the mandatory details identified by the appropriate MEDIN data guideline.
- For all data collected after 1st January 2014, The Crown Estate requires data to be provided in full compliance with the appropriate MEDIN data guideline (where one exists). For the avoidance of doubt, full compliance with the guidelines will not be expected to be applied to data collected before 1st January 2014.

datasupport@thecrownestate.co.uk
Tel. 0207 851 5000

2. Introduction

When a renewable energy developer enters into an Agreement[1] with The Crown Estate an obligation to provide data collected in respect of the development is included in the data clause. This document should be taken as a generic set of requirements, with the Agreements being referred to for specific definitions of the obligation.

This document focusses on survey data. All survey data are provided to The Crown Estate via the Marine Data Exchange (MDE). The more complex nature of survey data means that it is necessary for The Crown Estate to provide a specialised system which incorporates a comprehensive quality assurance mechanism, ensuring maximum value is derived from survey data for both The Crown Estate and developers alike.

To standardise data management and quality, The Crown Estate has adopted the MEDIN[2] discovery metadata standard, and also uses the MEDIN Data Guidelines to quality assure data. MEDIN is a partnership of over 30 UK organisations committed to improving access to marine data and the MEDIN data model is accepted as best practice across government Data Archive Centres.

All data is provided and handled according to The Crown Estate's Marine Data Policy. This Policy advises how we work with developers to help them discharge with regards to the data clause. It also outlines the principles by which we take account of commercial sensitivity and generally seek to make data available to the public.

3. Scope

The scope of this document is to provide requirements for providing survey data in accordance with Agreements including defining:

a. High level data structure
b. High level data definition: Series, Datasets and Reports
c. Metadata requirements
d. Minimum acceptable standards for survey data
e. The process for providing survey data

Outside of this scope are requirements relating to non-survey data e.g. specific GIS/CAD data requirements for turbine layout plans and cable route options, commercial contracts, health and safety information, specifications for survey data collection and the Round 3 non-spatial data room.

This document is provided for information purposes only and in no way replaces, varies or supplements the terms of any Agreement. It should be used to supplement the data clause.

[1] "Agreement" means *inter-alia* an Exclusivity Agreement, Zone Development Agreement, Agreement for Lease or Lease held with The Crown Estate, or a temporary consent to perform intrusive seabed works.

[2] Marine Environmental Data and Information Network

datasupport@thecrownestate.co.uk
Tel. 0207 851 5000

4. Supporting Appendices

This document is supplemented by a number of Guidance Notes – included here as Appendices:

Appendix No.	Subject	Scope
Appendix 1	The Marine Data Exchange	1. An introduction to the Marine Data Exchange, including why the system was created, and how the published data is discovered and used.
Appendix 2	Series anticipated from an offshore wind development	1. Provide a high level breakdown of anticipated Series to be provided to the Marine Data exchange for an offshore wind development.
Appendix 3	Detailed Data Structure Examples	1. This provides some guidance and a template to structuring a variety of survey types, including more complex surveys such as geophysical surveys that have numerous outputs and deliverables.
Appendix 4	QC Protocol	1. The Crown Estate's check list for performing quality control on received data. 2. The minimum acceptance criteria related to data
Appendix 5	Metadata	1. An introduction and background to metadata standards for marine data 2. The Crown Estate's preferred metadata standard for marine data.
Appendix 6	Survey and Information Management Standards	1. Provides links to information regarding survey and information management procedures 2. Recommend adoption of consistent approach throughout the Series life-cycle

5. Requirements

5.1 Themes

For the definition of "data" for a specific development, please refer to the relevant Agreement[3]. In generic terms, data themes that are required include:

- Environmental and ecological
- Physical environment
- Human environment
- Engineering & technical

[3] It is important to note that under some Agreements "data" can mean all data required by the developer for the purposes of developing the relevant project.

datasupport@thecrownestate.co.uk
Tel. 0207 851 5000

These themes refer to primary observations, investigations, modelling, monitoring, and associated reports. For a detailed breakdown of anticipated survey data to be provided to the MDE – please see Appendix 2.

5.2 Stages of development

Data are required during the entire life-cycle of a Project including: feasibility, consenting, pre / during and post-construction (monitoring), and during / post decommissioning[4].

5.3 Data Structure

The data shall be structured according to *Figure 1*. Although the majority of data shall be provided online, The Crown Estate may sometimes approve a data delivery offline on physical media. Any delivery to The Crown Estate, whether online or offline, must always follow the structure defined in Figure 1.

A workable example of this structure can be seen in Figure 2, and guidance for structuring more complex Series, such as geophysical, can be found in Appendix 3.

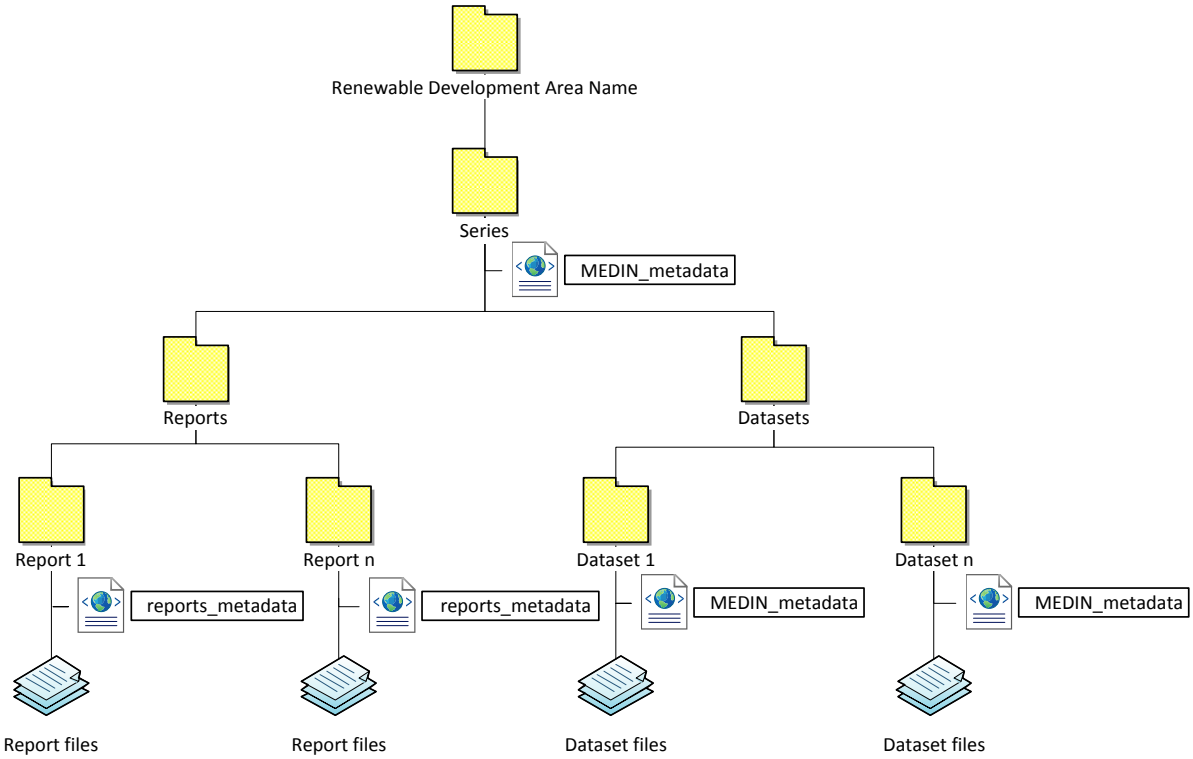


Figure 1 - Survey data structure implemented by The Crown Estate

[4] It is important to refer to the relevant Agreement, as Agreements may differ on this point.

datasupport@thecrownestate.co.uk
Tel. 0207 851 5000

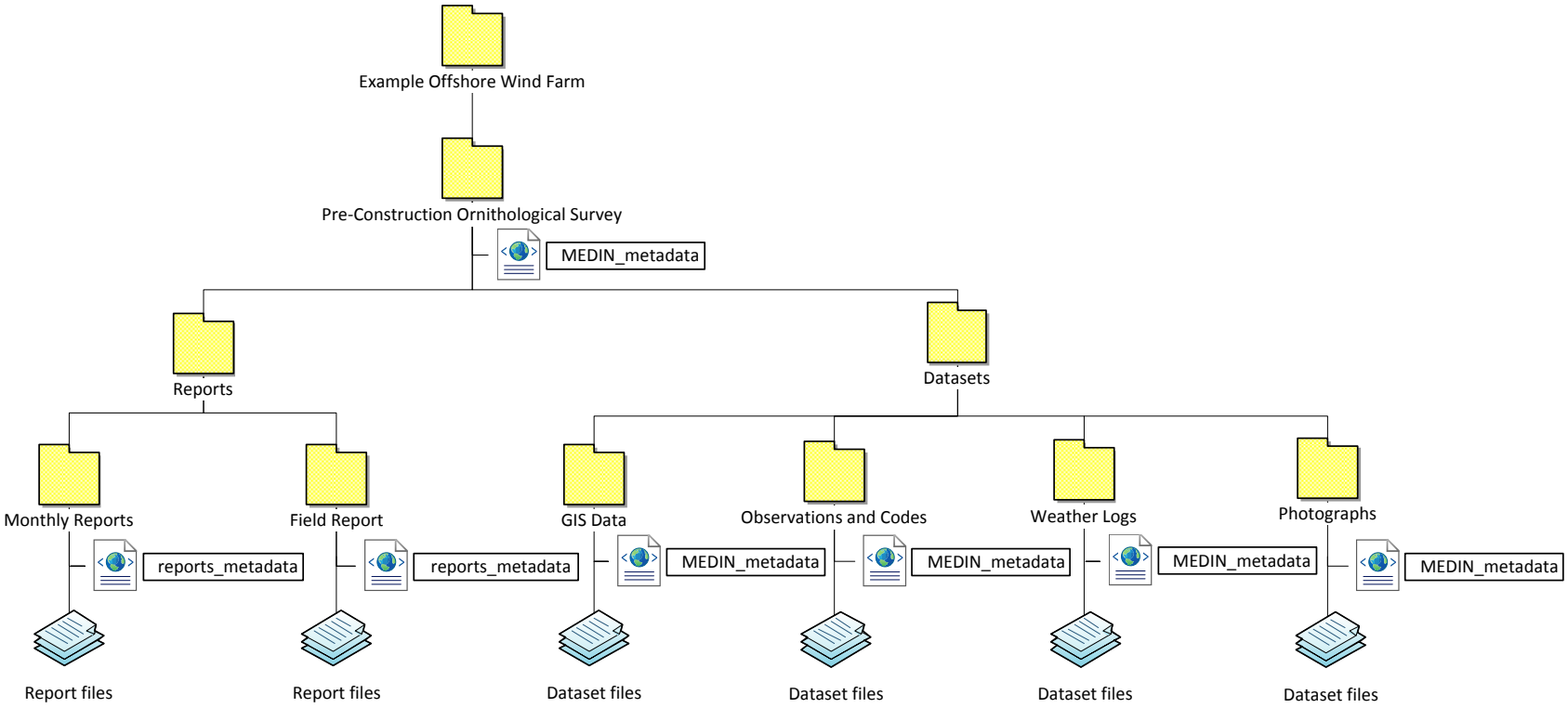


Figure 2 - Example structure for an ornithology survey

datasupport@thecrownestate.co.uk
Tel. 0207 851 5000

5.3.1 Series

A series can be defined as a collection of datasets and reports related to a survey or data collection campaign, linked by a common research question. A series can be viewed as a folder holding related datasets and reports together and must be described by a MEDIN discovery metadata record. It is important that the series contains all final reports, report appendices, datasets and other deliverables, such as GIS shapefiles, where relevant.

The data clause does not extend to requiring every version of a dataset and/or report and each dataset and/or report should be assessed to confirm the final versions of reports, before being provided. As a guide, a series provided to The Crown Estate will contain:

- Metadata for the series, dataset and report to an agreed standard;
- A base data product (cleaned and QA'd version of the raw data, pre-analysis);
- Raw data where appropriate;
- A final version of the analysed/modelled data (e.g. 10m resolution digital elevation model);
- QC and calibration reports;
- Supporting data (e.g. survey equipment, logs, GIS shapefiles); and
- Final report(s).

Example Series: 2020, Originator, Example Offshore Wind Farm, Pre-construction Ornithological Survey

5.3.2 Datasets

Datasets are only required within a series if data has been collected or created. A dataset is defined as a 'pillar' of data and information related to a single parameter, instrument or method. Therefore, individual dataset packages should be provided for each parameter, instrument or method, along with a MEDIN Discovery Standard metadata file. Within a series, there can be one or more datasets and special attention should be given to ensure that *all* datasets are provided, especially for surveys where numerous equipment types are used and data outputs generated, such as geophysical surveys (please see Appendix 3 for an example).

Example Dataset 1: 2020, Originator, Example Offshore Wind Farm, Pre-construction Ornithological Survey – Observation Data

Example Dataset 2: 2020, Originator, Example Offshore Wind Farm, Pre-construction Ornithological Survey – Weather Logs

Example Dataset 3: 2020, Originator, Example Offshore Wind Farm, Pre-construction Ornithological Survey – GIS Data

Example Dataset 4: 2020, Originator, Example Offshore Wind Farm, Pre-construction Ornithological Survey – Photographs

5.3.3 Reports

Also within the series there can be one or more reports. A report is defined as a document relating to part, or the whole of a survey or data collection campaign. Reports must be accompanied by a TCE Reports Metadata.

Example Report 1: 2020, Originator, Example Offshore Wind Farm, Pre-construction Ornithological Survey – Monthly Reports

Example Report 2: 2020, Originator, Example Offshore Wind Farm, Pre-construction Ornithological Survey – Field Reports

5.4 Metadata

The Crown Estate's requirements for metadata are:

Level	Standard	Where can this be created?
Series	MEDIN Discovery Standard	Using the MEDIN Online Tool or the Metadata Maestro desktop tool. See here.
Dataset	MEDIN Discovery Standard	Using the MEDIN Online Tool or the Metadata Maestro desktop tool. See here.
Report	The Crown Estate Report Metadata	This can be created using the link on the Marine data Exchange. See here.

Metadata standards provide a definition of a set of minimum elements that must be populated to create a conformant metadata record. Not only do standards improve the ability to discover, understand and evaluate datasets, they also ensure that the originator understands the data that has been created by ensuring that certain elements are considered (e.g. spatial reference system of the dataset).

Metadata is critical in the storage and discoverability of data on the MDE. In addition to storing and managing survey data, the MDE acts as a platform for disseminating and publishing data into the public domain (see Section 6.1. for information about the publishing and confidentiality process), where it can be freely downloaded. Here the metadata acts as a catalogue and is the first item returned in a search prior to the dataset or report being downloaded. Therefore, it is crucial that the information is accurate and facilitates discoverability.

The Crown Estate requires all mandatory fields within the metadata to be complete accurately, ensuring that the title follows the MEDIN recommended format (date, originator, location, survey type), the spatial and temporal information is correct, abstracts are informative and a minimum of three keywords are provided: MEDIN Parameter Discovery Vocabulary, TCE Renewable Energy Development Areas, TCE Renewable Energy Development Phases.

All mandatory fields are clearly marked in each of the metadata creation tools. For more information see Appendix 5. For help with MEDIN metadata, please contact the MEDIN Metadata Helpdesk at medin.metadata@mba.ac.uk or 01752 633291. An example MEDIN Discovery Standard metadata record is provided below:

Example MEDIN metadata

5.5 Standards for Data

5.5.1 Coordinate Reference Systems

Where data are spatially referenced, their coordinate reference system should be fully defined, and where transformation parameters do not exist (e.g. a bespoke engineering grid), transformation parameters between the coordinate reference system and WGS 1984 (EPSG code: 4326) should be provided.

5.5.2 MEDIN Data Guidelines

Where appropriate, The Crown Estate will use the MEDIN data guidelines (http://www.oceannet.org/marine_data_standards/medin_data_guide.html) as a pro forma for quality control of data received. The data guidelines set out the requirements for information that must be recorded when a certain type of data is being collected (e.g. cetacean sighting and identification data) to ensure that the data can be understood, interpreted and used with confidence now and in the future. These guidelines instil good practice, aid organisations in specifying formats that data should be returned in and ensure that all relevant attributes are included to ensure that the data collected is of maximum value.

For all data collected after 1st January 2014, The Crown Estate will require data to be provided in full compliance with the appropriate MEDIN data guideline (where one exists). For the avoidance of doubt, full compliance with the guidelines will not be expected to be applied to data collected before 1st January 2014. Although The Crown Estate does not require data to be delivered in the MEDIN data Guideline Excel template, all mandatory elements (as defined within the Guidelines) must be available within the datasets and reports that make up a complete series.

For help and advice in using these guidelines, please contact MEDIN.

5.5.3 Data Guidelines for Meteorological Masts and Remote Sensing Devices

When data is submitted that refers to wind measurements recorded from Meteorological Masts and Remote Sensing Devices like LIDAR, The Crown Estate will use the following guidelines as a pro forma for quality control of data received.

The following wind data is expected:

Data	Description
Raw Data	Raw data collected in 10-minute intervals from each sensor. This should provide information about: • Mean wind speed • Maximum wind speed • Minimum wind speed • Standard deviation of wind speed • Mean wind direction • Maximum wind direction • Minimum wind direction • Standard deviation of wind direction • Air temperature • Air pressure
Cleaned Data *The complete meteorological dataset with all calibrations applied and instances of erroneous data, sensor degradation, and sensor icing removed from the dataset. The dataset should be comparable to the data that would be used for a pre-construction energy assessment.*	Cleaned data collected in 10-minute intervals from each sensor. This should provide information about: • Mean wind speed • Maximum wind speed • Minimum wind speed • Standard deviation of wind speed • Mean wind direction • Maximum wind direction • Minimum wind direction • Standard deviation of wind direction • Air temperature • Air pressure

To accompany the data and to provide a site description, the following information is expected:

Site Information	Description
Device coordinates and datum	The coordinate reference system should be fully defined, and transformation parameters to WGS 1984 (EPSG code: 4326) should be provided.
Description of mast or LIDAR	This should provide details of the manufacturer, type of instrument and height from mean sea level.
Description of mast equipment	This should describe the equipment used to record data, such as information about the sensors including device calibration certificates and logger programs.
Installation record	This should provide details of the installation date and process
Maintenance records	All maintenance records should be provided to inform any inconsistencies in the data, including details of replacement sensors.
Sensor and mast relationships	For meteorological masts, the type, location, height and orientation of all sensors relative to the mast should be provided.

For further guidance, please see the MEDIN data guidelines for Fixed Position Meteorological Measuring Instruments.

6. Process of providing data

The process for providing data to The Crown Estate is outlined in *Figure 3*.

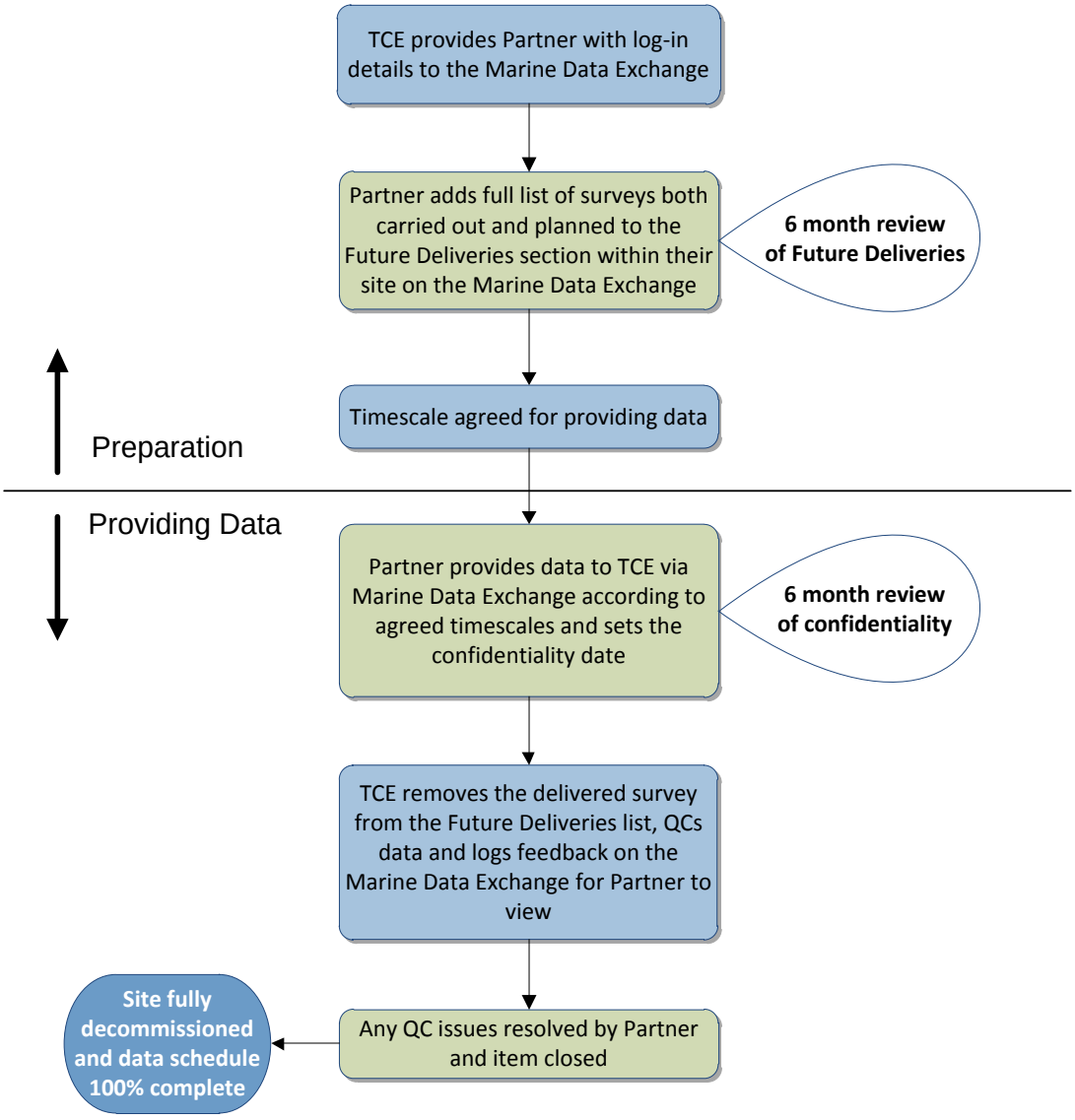


Figure 3 - Process for providing data to The Crown Estate

6.1 Publishing Data and Confidentiality Review Dates

In order to promote sustainability and our core values of Commercialism, Integrity and Stewardship, The Crown Estate's aim is to make as much data publically available as possible. This will facilitate collaboration and research for the benefit of industry and others, whilst also improving the understanding of the marine environment and the impacts of different developments on it. Alongside this aim is The Crown Estate's need to preserve the integrity of its business relationships with tenants and others.

A secure and traceable process for managing confidentiality of data has been implemented for data held on the MDE. Control over the confidentiality of data rests primarily with the providers of data. Periodic reviews of the status of data are undertaken in consultation with Developers and details of the decision to make data public or not are recorded.

The process is as follows:

1. Developers are required to set a sensible 'confidentiality review date' for each dataset or report that is uploaded to the MDE.
2. Once this review date arrives, a notification is sent to the MDE administrators
3. A discussion between the Developer and The Crown Estate is triggered surrounding the status of the dataset or report.
4. The decision of whether to publish the data or not is logged in the MDE and if necessary, a new date for review is set.

Within a series, different confidentiality statuses can be applied to each individual dataset or report. This ensures that a sensitive dataset or report can be held in confidence whilst the rest of the series is made publically available.

The confidentiality status of each Series, Dataset and Report can be seen quickly and easily in the MDE, along with the date that any data were made publically available[5].

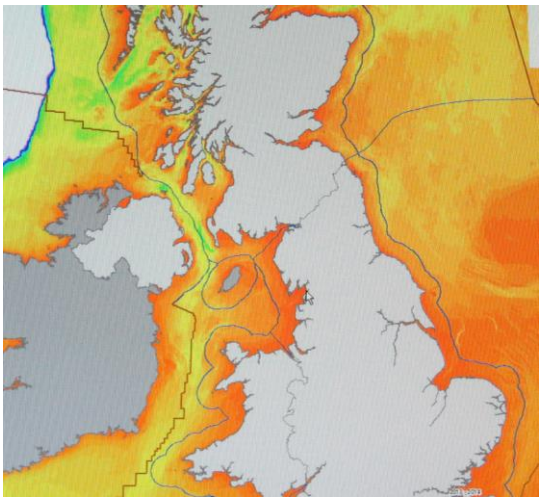
For more information surrounding the data clause, please refer to the Marine Data Policy.

[5] Please note that data which has been migrated from the previous COWRIE Data Management System retains the same status as was recorded in that system. I.e. data that was publically available in the COWRIE Data Management System will be publically available in the MDE.

Appendices

Appendix 1 - The Marine Data Exchange

Protecting and sharing a pool of valuable knowledge



A young industry that requires extensive research and environmental study

Offshore renewables is a 'young' industry that requires extensive research and environmental study before starting new projects. These studies can take a long time and be expensive to carry out, creating barriers to entry and an additional financial burden for companies entering the sector. If a developer exits the industry or cancels a project, the data that they have built up can be lost, meaning other organisations cannot benefit from it and may even have to re-collect it at further cost.

The Marine Data Exchange ensures that information from offshore renewables can be captured and shared for the benefit of the industry

Recognising the potential benefits of more systematically capturing and sharing data, in 2002 The Crown Estate decided to insert a 'data clause' into all new offshore leases. The aim was to ensure that survey data, information or knowledge coming out of industries such as wind, wave and tidal energy, was captured for the long term, and could also be shared with the wider industry and made publicly available.

The Crown Estate quickly recognised that the volume and variety of data, and need for its active management, would pose a challenge. It was clear that a new, robust, secure and flexible IT system was needed to deal with the storage and discovery of the data. The Crown Estate decided to invest the time and resources to build an integrated, multi-disciplinary data management and dissemination system, which has now become the Marine Data Exchange (MDE).

The MDE plays a vital role in de-risking the offshore renewables industry. It provides a resource, a structure and a control to boost future development potential, increase current value and to save money. By actively advocating transparency and sharing of data, knowledge and information The Crown Estate wish to ensure that learning is shared quickly and effectively to avoid repeating mistakes and overcome uncertainties or issues quickly. All of which could enable offshore renewables to become more efficient and effective.

Discoverability and accuracy of the information made available through the Marine Data Exchange is critical in enabling others to benefit from it.

The data published through the MDE enables students, universities, innovators and researchers to learn and understand more about offshore renewables and the marine environment. The MDE and its content are open to a variety of individuals and organisations, spanning from relative beginners to experts in their field. Therefore, it is critical that all data publically available on the MDE is coherent and complete, so that it can be re-used with confidence by others. Furthermore, the ability to discover data relies upon the content of the metadata which is uploaded with the data by the offshore tenant. As with any cataloguing system it is paramount that the content describing the resource is accurate, so that when a search criterion is defined, the search results are relevant and all of the appropriate information is returned. For these reasons, The Crown Estate processes all datasets, reports and their metadata through a quality checking process to ensure that the information can be discovered and is useable both now, and in the future.

For more information please contact: mde@thecrownestate.co.uk

To access data that has been made publically available please visit

www.marinedataexchange.co.uk

Appendix 2 – Series anticipated from an offshore wind development

This guide has been compiled of two documents:

BVG Associates. 2010. A Guide to an Offshore Wind Farm. Swindon and CEFAS. 2004. Offshore Wind Farms: Guidance Note for Environmental Impact Assessment in Respect of FEPA and CPA Requirements, Version 2 - June 2004.

The below table demonstrates the survey themes anticipated to be carried out during the lifetime of an offshore wind farm development (this will also largely apply for wave and tidal energy developments). Many surveys will be repeated as the project progresses from the feasibility/consenting phase through to decommissioning/re-power.

Survey Name	Content
PHYSICAL ENVIRONMENT: GEOPHYSICAL, PHYSICAL PROCESSES, GEOTECHNICAL, METOCEAN, SEDIMENT SAMPLING	
Geophysical surveys of the site and cable routes	• (Single) multi-beam echo soundings or swathe bathymetry, side scan sonar, magnetometer, sub bottom profiler (sparker, boomer, etc.). • Please see the geophysical survey example in Appendix 3
Meteorological data	• Meteorological and oceanographic conditions at the site; sensors and auxiliary systems measuring wind speed, wind direction, temperature, pressure, humidity, solar radiation and visibility. For The Crown Estates wind data requirements please see section 5.5.3 of the Requirements Document and for further details please refer to the MEDIN data guidelines for Fixed position Meteorological Measuring Instruments.
Geotechnical survey	Test procedures, Qualification testing, Equipment verification, Daily operation reports and field logs, Borehole records, Core photographs, Vibra-core, Cone, DGPS and gyrocompass calibrations, Lab report, Cross section diagrams, Raw and derived PCPT test results, standard , Penetration tests, PSA. The Crown Estate strongly recommends contacting the British Geological Survey about possibly depositing borehole samples. This would be discussed on a case by case basis based upon location, terminal depth and proximity to other samples.
Metocean surveys	Data on wave, tide, current and suspended sediment concentrations
Coastal processes	Interpretation of hydrodynamic, sediment and morphodynamic regime
ENVIRONMENTAL & ECOLOGICAL: BENTHIC, BIRDS & MAMMALS, SEDIMENT SAMPLING	
Intertidal habitat appraisal	Wildlife surveys to consider potential ecological impact
Sediment and benthic sampling surveys	• Baseline control sites (then on-going monitoring) • Beam trawl sampling for epifaunal species • Video observations of colonisation of the monopoles and scour protection • Further e.g. grab sampling, retention of benthic fauna, sampling of the sediment particle size analysis, sampling of the sediment for contaminants, drop-down camera sampling • Ongoing monitoring – a minimum of three replicates at each station • Survey of the intertidal habitats where trenching is used to install the export cables in the intertidal zone
Seabed morphology and scour survey	• Two (one winter and one summer) high resolution swath-bathymetric surveys per annum of the wind farm intra-array and export cable route – pre-construction baseline, then twice a year • Immediately after construction is complete, 100% coverage swath bathymetric survey around a sample of adjacent turbines. Repeated at six monthly intervals for a period of three years (revision report every 1.5 years).

Marine mammals survey	•Aerial, boat-based visual surveys, mammal tagging and PAM assessment of the potential impacts of the wind farm to mammal populations •Mitigation of marine mammals during installation of foundations • Additional monitoring throughout the life cycle of the project or as deemed necessary
Ornithological surveys	• Aerial, boat-based visual surveys, radio tagging surveys • Comparison between development phases i.e. pre-construction to post-construction • Please see Section 5.3 Figure 2 in the main document for guidance on structuring ornithological series.
Seabird survey for locally identified species	Terrestrial, inshore boat, and offshore boat surveys; reanalysis of bird population data to apply in modelling to calculate mortality rates in relation to wind farms; feeding ecology, desk reviews
Marine ecology	• Investigation of benthic grab samples, epifaunal beam trawling and drop-down video sampling surveys to provide data and analysis informing on the marine ecology • Pelagic survey • Information on the macro benthic properties
Electromagnetic field	Desk-based study
Environmental Statement	Full Environmental Statement including appendices
HUMAN ENVIRONMENT: VISUALS LANDSCAPE & SEASCAPE, COMMERCIAL FISHERIES, SOCIO-ECONMIC, NOISE, UXO, ARCHAEOLOLOGICAL, AVIATION RADAR	
Archaeological assessment	•Investigation of geophysical, geotechnical and diver/ROV data to identify archaeological sites and exclusion construction zones. • Subsequent monitoring: results of watching briefs, a 3 year reporting cycle to monitor effectiveness of seabed construction exclusion zones
Underwater noise study	Modelling assessment of underwater noise during piling operations
Seascape and visual technical study	Magnitude and significance of change to the character of the regional seascape, landscape and general views
Shipping and navigation risk assessment	An assessment of the potential impact of the proposed development on the routeing and safe navigation of vessels. Baseline information relating to ports, navigational aids, sailing directions, wrecks, oil and gas infrastructure, exercise areas and met ocean data.
Maritime traffic survey	Radar, Automatic Identification System (AIS) and visual observations.
Magnetometer UXO survey	Geophysical interpretation of sonar and magnetometer charts together with detailed interrogation of the digital raw data
Marine fish survey	A survey determining the general marine fish status (numbers, distribution) of the local fish populations which will be used for comparison with the post construction surveys (adult and juvenile fish surveys); fish and shellfish resources Raw, processed and interpreted data as well as report.
Human activities	Any surveys should be related to benthic surveys: • The distribution, intensity of recreational activities • The distribution, intensity and type of commercial and recreational fishing activity
Socio-economic study	Accesses the impacts of a wind farm or coastal infrastructure, e.g. a port, such as changes in employment, transportation or recreation, or changes in the aesthetic value of a landscape.
Noise and vibration from the turbines	• Measurements at various frequencies at locations immediately adjacent to and between turbines
Radiological assessment of disturbed sediment	A desk-based assessment to determine the risk to human health from the sediment that could be disturbed during the wind farm construction
Cables (if jetting of the export cable in the intertidal zone is agreed)	Monitoring of suspended sediment concentrations within the area of jetting and at a control point outside the area.
Decommissioning programme	Proposed decommissioning measures

Appendix 3 – Detailed Data Structure Examples

This section is intended to provide guidance when structuring and uploading some of the more complex surveys where multiple equipment types are used and many different datasets and reports are generated, such as geophysical surveys.

Geophysical Survey Example

The following example for geophysical survey data has been derived from the BGS Marine Survey Folder Structure for geological and geophysical survey data, and adapted to suit the requirements of the MDE.

Please refer to Figure 1 when structuring and uploading geophysical survey data.

All supporting data should be found in the Ancillary Data package. This includes Sound Velocity Profiles (SVP) and Conductivity, temperature and Depth (CTD) data files.

A Geophysical Equipment Type dataset package should be provided for each of the equipment used, for example, side scan sonar, boomer, pinger and single beam echo sounder. Ideally processed SEGY/XTF files should be produced for each equipment type, specific to the start and end of lines, but where this is done with no post-processing they should be stored as *Raw Data. Processed Images* refers to the Tif images.

Where a CODA system is used to record multiple equipment types, all CODA files should be stored in one dataset package. It is important to keep all of the original CODA files together so that they can be readable for processing and interpretation. The CODA dataset package should contain the files produced by the software package during acquisition and processing phases of the data e.g. CODA project file (.prj) – referred to as *Software Files* in the following diagram.

The Multibeam dataset package should be used to store multibeam echo-sounder data (raw, processed, products and images). A new dataset package should be created for each multibeam equipment type.

Within each multibeam dataset package the *Bathymetry Products* refers to ESRI ArcGrid, bathymetry images, Fledermaus SD files and XYZ. Although these files should be uploaded into the same dataset package, a coherent structure should be maintained by zipping these products into individual folders.

Backscatter Products/Images within the Multibeam package refers to the GeoTIFFs, which should be provided here. The *Software Files* folder name should include the name of the software used (e.g SoftwareFiles_CARIS). If raw data files are produced by the software package during the acquisition and processing phases, they should be included here rather than as *Raw Data*.

All data relevant to the course of the ship and equipment position during data acquisition should be provided under the Navigation dataset package.

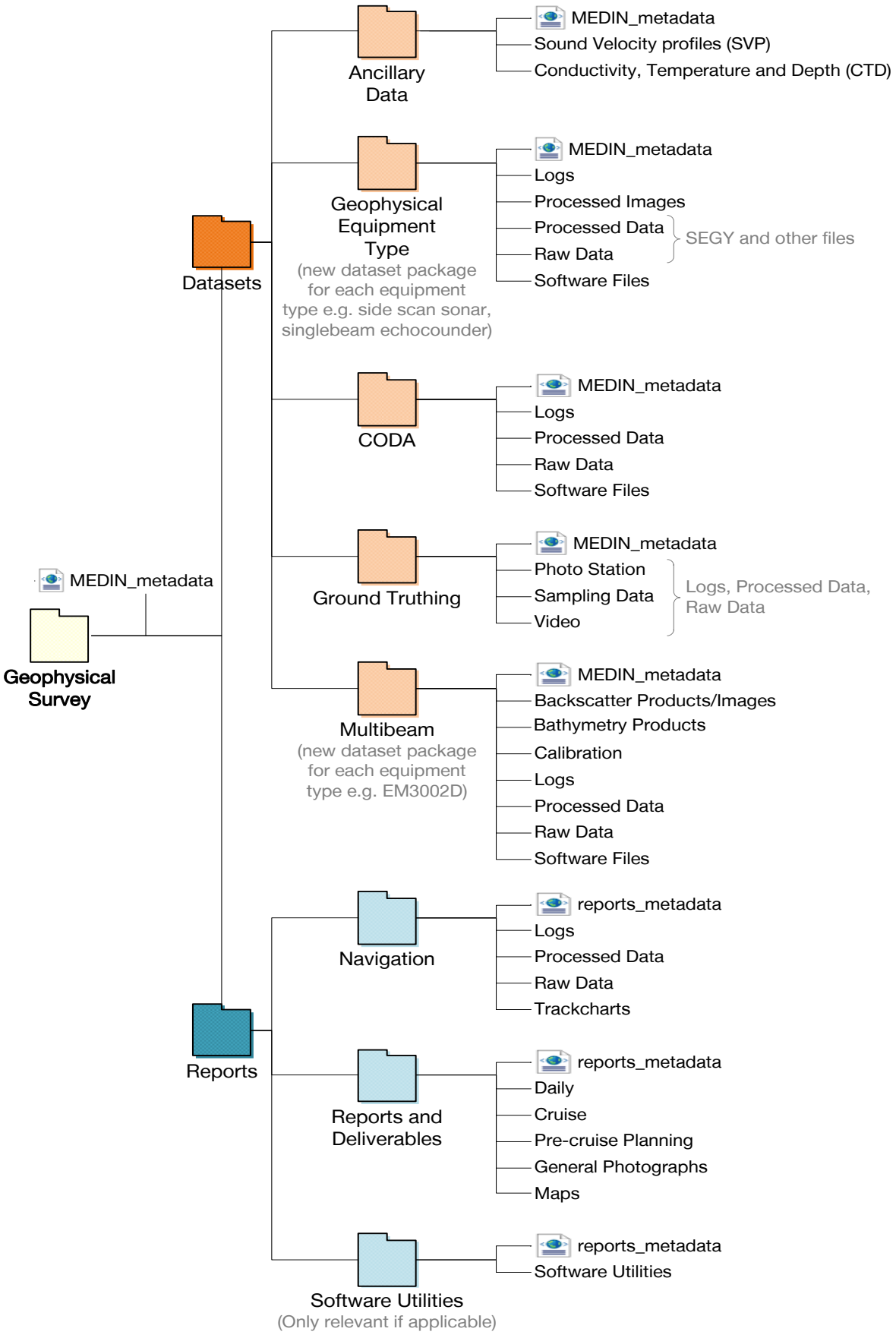


Figure 1 - Example geophysical survey data structure

Wind Data Survey Example

Wind data is treated slightly differently from other survey data due to its continuous nature. A Series on the MDE would not be considered complete until all datasets and reports have been provided and passed QA – this process will be on-going until the met mast is decommissioned or the lidar is recovered and the data collection campaign complete.

As per the Agreement, The Crown Estate intends to publish wind data after a two year confidentiality period. For this reason, it is recommended that datasets are structured into discrete annual packages containing all raw and cleaned data.

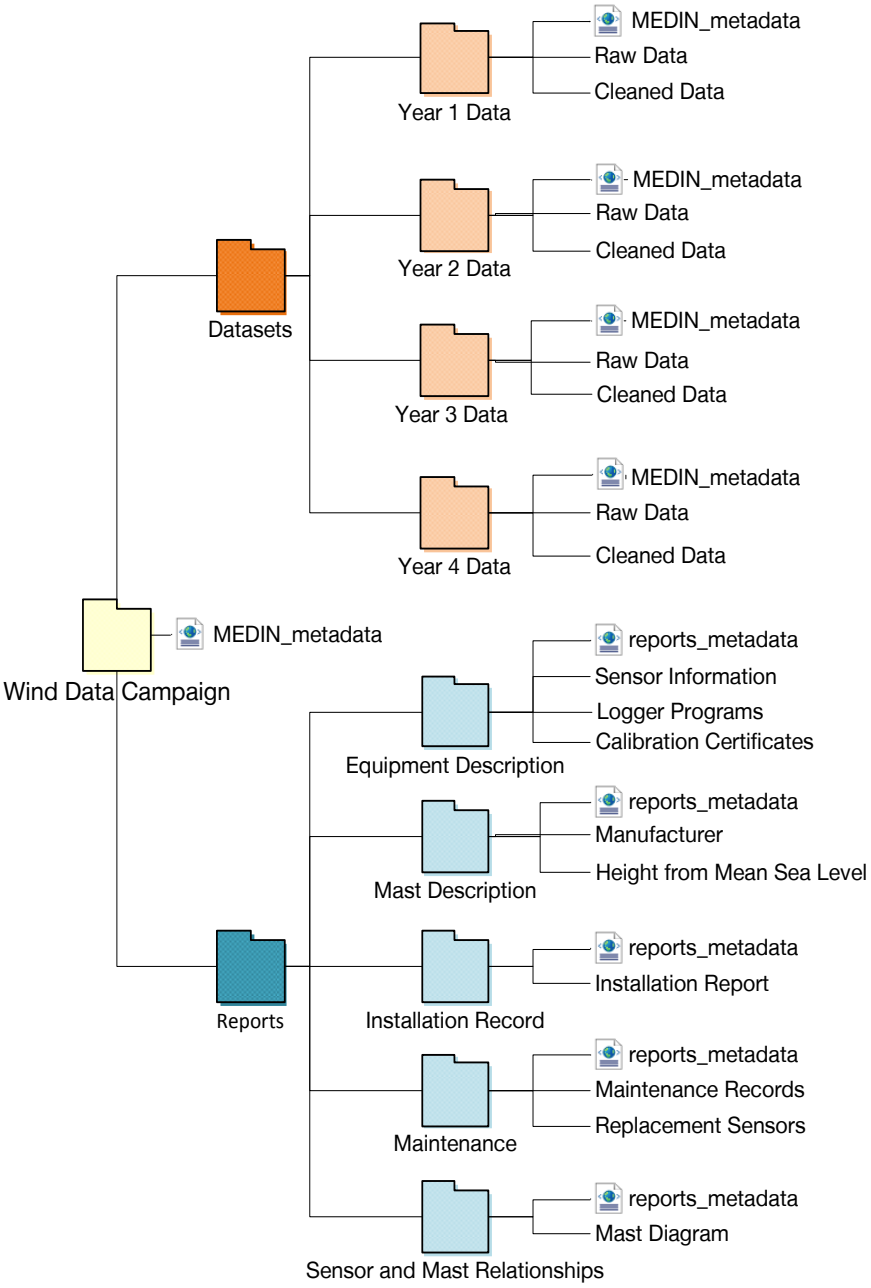


Figure 2 - Example wind data survey structure

Appendix 4 – Quality Assurance Protocol

All data uploaded to the Marine Data Exchange (MDE) will be assessed through a high level Quality Assurance (QA) process. This is documented directly onto the MDE and any outstanding QA issues are discuss during regular meetings with the data provider.

Where applicable, The Crown Estate will carry out the following checks on each Series, Dataset and Report:

1. Check the metadata file
   - ✓ Correct metadata file present
   - ✓ Metadata content appropriate i.e. title, abstract, keywords, contact details
2. Check the coordinate system
   - ✓ Defined and transformable to WGS 84
3. Check the content of the Datasets/Reports
   - ✓ No human or gross errors
   - ✓ No file naming issues
   - ✓ Resolution and coverage of the data is as expected
   - ✓ QC and calibration reports are available where there is raw data
   - ✓ All data collected is MEDIN compliant, as checked against the relevant MEDIN Data Guidelines where available
4. Check the confidentiality date
   - ✓ It has been set for each Dataset and/or Report

Appendix 5 - Metadata standards: guidance for third parties

Contents

For more information please contact Peter Edmonds, Spatial Data Manager at The Crown Estate:

peter.edmonds@thecrownestate.co.uk / *Tel 020 7851 5349*

1 Introduction

1.1 Scope

This document provides high-level guidance on metadata for Series, Datasets and Reports with technical guidance for implementation in the Annexes. Metadata standards for describing services (e.g. a web mapping service, see ISO 19119) are not included.

The guidance note is intended to describe the provenance of metadata standards, the links between them and their application within best practice survey data management, and to recommend adoption of the guidance within developer organisations.

It is acknowledged that there are more generic metadata standards designed for resources other than survey data (e.g. websites or books), for example Dublin Core (ISO 15836) and profiles thereof, such as the e-Government Metadata Standard (e-GMS). However, the standards described in this document, and adopted by The Crown Estate, are designed to be applicable to a variety of resources.

1.2 Background

Metadata is often defined as 'data about data' or 'information about data'.

Metadata is a critical resource that allows a user to discover, understand the provenance of and evaluate how appropriate data are for a particular use case. It can also be used to track the currency of data, the steps that have been taken in processing the data, and contact details for the metadata and /or data. These kinds of information allow a user to correctly understand the data origins and content, and to re-use the information where necessary for a wide variety of tasks, such as stepping back through the creation process, fixing problems or using parameters.

1.3 Glossary and abbreviations

Discovery metadata	Defines the minimum set of elements required to enable a dataset to be effectively searched for.
e-GMS	e-Government Metadata Standard.
Element	A discrete item of information.
INSPIRE	The Infrastructure for Spatial Information in Europe (INSPIRE) Directive aims to create a European Union (EU) spatial data infrastructure and came into force on 15 May 2007. It will be implemented in stages.
ISO	International Organisation for Standardisation.
MaRS	Marine Resource System.
MEDIN	Marine Environmental Data and Information Network.
Metadata profile	Specific definition of a metadata standard designed for a particular use.

Metadata record	A collection of metadata elements relating to a resource.
Schema definition	Describes, defines and establishes the relationships between different elements of metadata e.g. a UML model.
Schematron	Rule based validation language, based on XSLT that validates XML documents against assertions that detect the presence or absence of patterns in an XML tree.
Standard	Defines the metadata elements and relationships between them, the core metadata elements that must be populated and may consider the technical encoding to in which the metadata should be transferred.
Technical encoding	Defines the technical encoding used to store and transfer metadata records e.g. a set of XSD files.
XML	Extensible Mark-up Language.
XSD	XML Schema Definition/Document.
UML	Unified Modelling Language.

2 Metadata standards

2.1 The business case for metadata

Metadata standards are used because they allow for interoperability and encourage sharing of metadata between different data suppliers, consumers and warehouses. They provide standard definitions of elements, define relationships between elements, define a set of minimum elements that must be populated to create a conformant metadata record, and potentially provide a technical encoding by which to store and share the metadata record. Not only do standards improve the ability to discover, understand and evaluate datasets, they also ensure that the originator understands the data that has been created by ensuring certain elements are considered (e.g. spatial reference system of the dataset).

2.2 Metadata basics

A minimum set of metadata elements, such as those considered to be mandatory in the MEDIN Discovery metadata (see section 2.4 of this appendix) provide a simple and consistent means by which to audit, understand, correct, or alter the dataset to ensure that it is in an appropriate format for a specific use case.

Metadata should form an integral part of any dataset. The definition of a geographic dataset should be considered to include: data, metadata and a defined coordinate reference system. Without one of these items, a geographic dataset should not be considered complete.

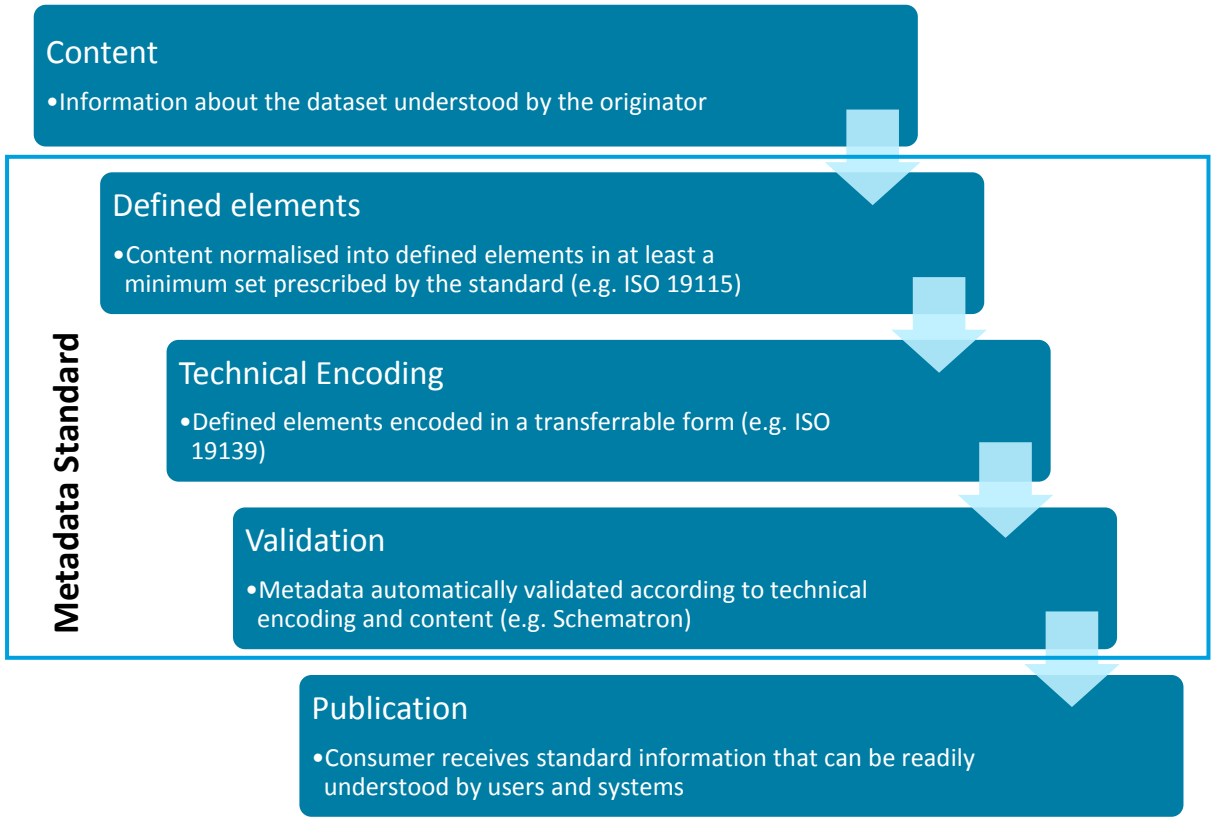


Figure 1 - Component parts of a metadata standard

2.3 The Crown Estate and metadata

The Crown Estate manages substantial amounts of marine data through the Marine Data Exchange (MDE) and the Marine Resource System (MaRS). Having assessed the available standards that are widely used and relevant to the UK marine sector, the Marine Environmental Data Information Network (MEDIN) discovery standard for metadata v2.3.7 (see section 2.4) was found to be the most suitable.

Suppliers of data to The Crown Estate are required to generate metadata which conforms to this MEDIN discovery metadata standard, (see Annex C for technical guidance) which will help to ensure the maximum interoperability between data held and produced, and that all parties gain maximum benefit more generally from improved data management.

2.4 MEDIN discovery metadata standard v2.3.7

MEDIN is a partnership of over 30 UK organisations committed to improving access to marine data. The MEDIN standard conforms to the most up-to-date and relevant UK, EU and international standards and provides additional benefits such as marine thesaurus services to improve an organisation's ability to standardise keyword terms. A more detailed description of different metadata standards can be found in Annex A.

Guidelines for the creation of metadata to this standard can be found in Annex C.

MEDIN Discovery Metadata requires the following set of minimum elements (M indicates a mandatory element, C indicates a conditional element) encoded in ISO 19139 (see appendices):

Element	Description
Resource title (M)	The title is used to provide a brief and precise description of the resource. The following format should always be used: Date, Originating Organisation/programme, Location, Type of Survey.
Resource abstract (M)	The abstract should provide a clear and brief statement of the content of the resource.
Resource type (M)	Identify the type of resource e.g. a dataset or series
Resource locator (C)	Formerly named online resource. CONDITION: If the resource is available online you must provide a web address (URL) that links to the resource.
Unique resource identifier (M)	MEDIN recommends the use of a code + a codespace e.g. Code: 5639287,Code space: www.bodc.ac.uk
Coupled Resource (C)	An INSPIRE element referring to data services such as a data download or mapping web services. You should supply the Unique resource identifiers of the relevant datasets. CONDITION: mandatory if linkages to the datasets on which the service operates on are available.
Resource language (C)	Describes the language(s) of any textual information contained within the resource. CONDITION: If the resource contains textual information.
Topic category (C)	This indicates the main theme(s) of the data resource using the ISO controlled vocabulary e.g. biota, environment, structure. CONDITION: mandatory for datasets and series.
Keywords (M)	Should consist of two sub-elements: the keywords and reference to the controlled vocabulary the keywords are taken from. The reference to the controlled vocabulary is automatically added when using any of the online or desktop metadata tools. The Crown Estate requires a minimum of three keywords from each of the following controlled vocabularies: MEDIN Parameter Discovery Vocabulary, TCE Renewable Energy Development Areas, TCE Renewable Energy Development Phases. For example: Fish behaviour, Offshore Area of Search, Pre-construction Monitoring.
Geographic bounding box (C)	Four sub-elements represent the geographical bounding box of the resource's extent. The co-ordinates of this bounding box should be expressed as decimal degrees longitude and latitude. CONDITION: mandatory element for datasets.

Element	Description
Spatial reference system (M)	Describes the system of spatial referencing (typically a coordinate reference system) used in the resource. This should be derived from the EPSG register of geodetic parameters (www.epsg.org/Geodetic.html).
Temporal reference (M)	The temporal extent of the resource (e.g. the time period over which data were collected) and the date of publication (i.e. the date at which it was made publicly available) are mandatory, dates of creation and revision are optional.
Lineage (C)	Lineage includes the background information, history of the sources of data used and can include data quality statements. CONDITION: mandatory element for datasets and series.
Spatial resolution (C)	Provides an indication of the spatial resolution of the data. MEDIN recommends providing the average distance (i.e. resolution) between sampling locations in metres. CONDITION: should be provided for datasets and series where a resolution distance can be specified.
Limitations on public access (M)	This element describes any restrictions imposed on the resource for security and other reasons using the controlled ISO vocabulary.
Conditions applying for access and use (M)	This element describes any restrictions and legal restraints on using the data. Any known constraints such as fees should be identified. If no conditions apply, then “no conditions apply” should be recorded.
Responsible party (M)	MEDIN mandates that the roles of ‘Originator’ and ‘Custodian’ (data holder) and ‘Metadata point of contact’ are provided. Other types of responsible party may be specified from the controlled vocabulary if desired.
Frequency of update (C)	This describes the frequency that the resource (dataset) is modified or updated and should be included if known. CONDITION: mandatory element for datasets and series
Conformity (C)	Please note that this element is under discussion with the authors of GEMINI2 and within the Location Programme metadata subgroup. It is currently felt that the MEDIN interpretation is INSPIRE compliant. Indicates whether a resource conforms to a product specification or other INSPIRE thematic specification. CONDITION: if wishing to create INSPIRE conformant metadata. The Crown Estate recommends this is completed.

Element	Description
Metadata Date (M)	This describes the last date the metadata was updated on. If the metadata has not been updated it should give the date on which it was created.
Metadata standard name (M)	Identify the metadata standard used to create the metadata. It is recommended that the term “MEDIN Discovery Metadata Standard” is used to comply with this MEDIN standard.
Metadata standard version (M)	Identify the version of the metadata standard used to create the metadata. It is recommended that the term “2.3.3” is used to comply with this MEDIN standard.
Metadata language (M)	Describes the language(s) elements of the metadata according to the ISO code list of languages e.g. “eng” for English.

2.5 The Crown Estate Report Metadata

To gain the greatest possible benefits in searching that are provided by metadata, The Crown Estate has also chosen to implement a metadata standard for reports. This standard is based on the MEDIN standard and is comprised of the following elements, all of which are mandatory:

Attribute name	Description	Type
Title	Name given to the report set.	Free text
Abstract	Brief narrative summary of the report set.	Free text
Reference Date	Reference date for the data resource: Creation, Publication or Revision.	Date
Topic Category	Main theme(s) of the data resource. Topic Categories are specified on a list maintained by ISO.	Controlled list
Keyword	The Crown Estate require a minimum of three words; one from each of the controlled vocabularies available using the online tool. MEDIN Parameter Discovery Vocabulary, TCE Renewable Energy Development Areas, TCE Renewable Energy Development Phases.	Controlled list

Report metadata can be created via the Marine Data Exchange.

3 Introduction to the Annexes

The appendices are intended to provide a more detailed description of the different available metadata standards, guidance to technical staff on how a consistent metadata standard might be implemented, and an overview of the benefits of adopting an internationally standardised technical encoding (ISO 19139).

- Annex A describes the development of metadata standards and justifies The Crown Estate's decision to adopt the MEDIN discovery metadata standard v2.3.7;
- Annex B provides a comparison between the different sets of minimum defining elements prescribed by the various standards discussed;
- Annex C describes options for authoring and transforming between different metadata standards;
- Annex D covers standard technical encodings of metadata and validation options.

Annex A The development of metadata standards

A.1.1 Glossary and abbreviations

AGI	Association of Geographic Information
ANSI	American National Standards Institute
ANZLIC	Australia and New Zealand Location Information Council
CSDGM	Content Standard for Digital Geospatial Metadata
DTD	Document Type Definition
GEMET	GEneral Multilingual Environmental Thesaurus
NAP	North American Profile
NGDF	National Geospatial Digital Framework
UK	United Kingdom
UML	Unified Modelling Language
XSD	XML Schema Definition/Document
XSL	Extensible Stylesheet Language
XSLT	Extensible Stylesheet Language Transformations

A.1.2 Metadata standards structure

Metadata standards tend to be hierarchically layered in three component parts: The element definitions (schema definition), the technical encoding, and the minimum set of elements (profile).

The table below gives a brief description and examples of the different component parts of metadata standards and how they fit together.

Component	Description	Example
Schema definition	Describes, defines and establishes the relationships between different elements of metadata.	A UML model schema that defines relationships between defined elements of metadata and provides a framework of how to populate metadata. Example: ISO 19115 UML model.

Technical encoding	Defines the technical encoding used to store and transfer metadata records.	A set of XML schema files (e.g. XSD's) defining how to populate, store and transfer metadata. Example: ISO 19139.
Profile	Prescribes the schema definition (and possibly the technical encoding) to be used to create metadata records along with specifying the obligation to populate particular schema definition elements.	A unit (e.g. country or industry) specific set of rules defining how and what to author in metadata relating to that unit. Example: UK GEMINI v2.

In addition to the technical structure of metadata, there is also a hierarchy for the use to which metadata is created for. In general terms there are 3 levels of resolution that metadata can be created for: discovery, evaluation and application/use.

- Discovery metadata – defines the minimum set of elements required to enable a dataset to be effectively searched for.
- Evaluation metadata – defines the same as discovery metadata plus additional information to enable the user to evaluate the dataset e.g. detailed information on the instrumentation and their calibration parameters used to collect wind data.
- Application/use metadata – defines evaluation metadata plus additional information to enable to user to apply the data to a specific use. This might include information such as a UML model to define the potential relationships to other datasets.

At the point of publication, there is no standard relevant to the UK that defines metadata beyond discovery.

A.1.3 Introducing the options

When deciding which geospatial metadata standard to adopt, The Crown Estate considered the different options that exist, their relative merits and constraints, the level of conformance with emerging data/metadata standards such as INSPIRE, the level of industry adoption and the ease of creation and validation.

There are a number of metadata standards available for businesses to use. The most common have been identified and compared. A more detailed comparison of the set of minimum elements for these standards can be found in Annex B. Below are the metadata standards and profiles considered most relevant to the UK marine sector, that are discussed in this section. Where one standard is born out of another (e.g. the MEDIN discovery metadata standard is a child profile of UK GEMINI), it will conform to its parent standard by default. This is why it is advantageous to have an understanding of the provenance of different metadata standards.

- FGDC (CSDGM)
- ISO 19115

- INSPIRE (metadata implementing rules)
- UK GEMINI – Version 2.1
- MEDIN discovery metadata standard – Version 2.3

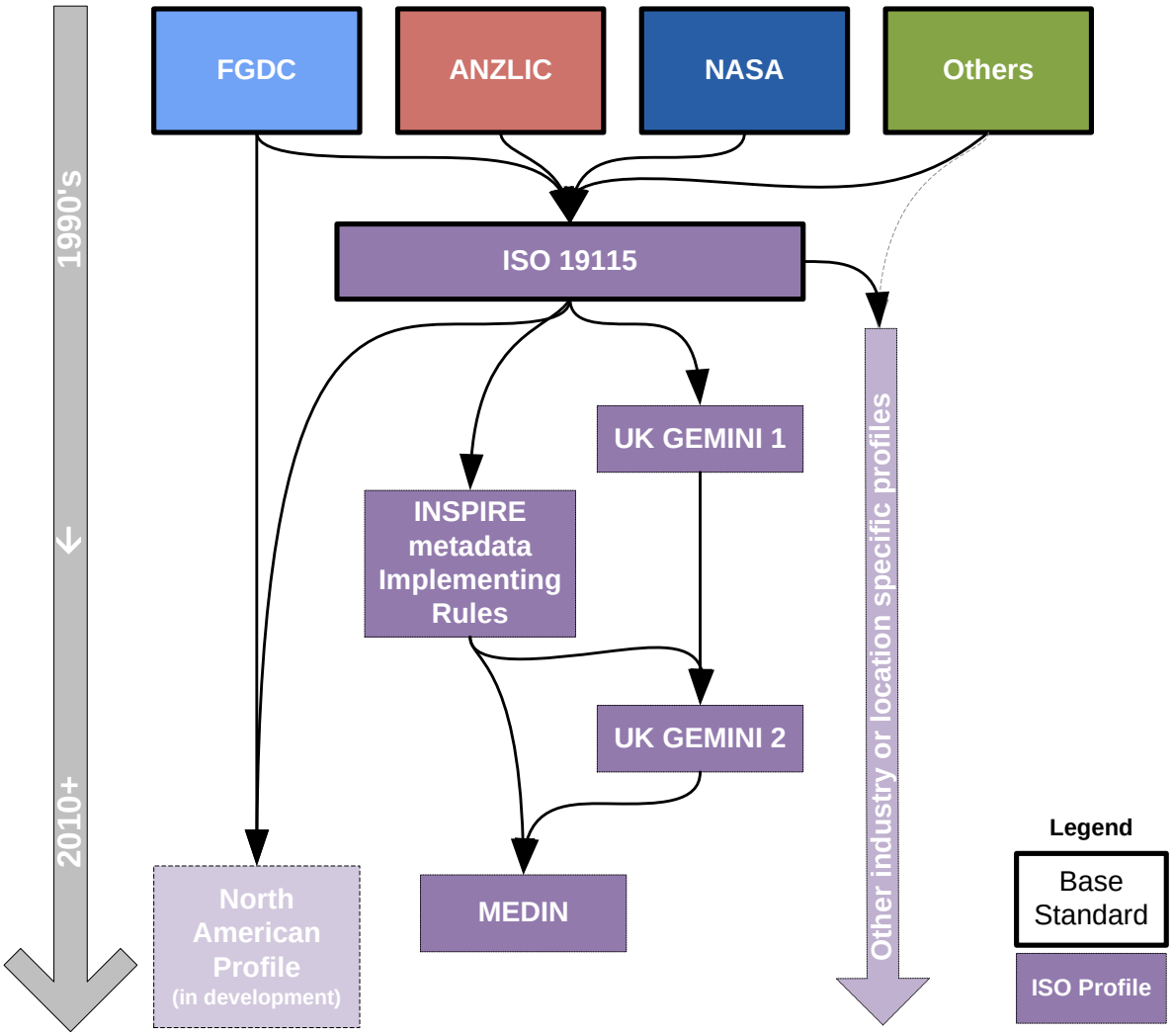


Figure 2 - High level overview of the provenance of metadata standards

A.1.3.1 FGDC

This standard was first adopted by FGDC in 1994, the latest version was published in June 1998.

The FGDC standard (officially called the Content Standard for Digital Geospatial Metadata (CSDGM), V2 (FGDC-STD-001-1998)) is the US Federal Metadata standard. FGDC were tasked to develop procedures and assist in the implementation of a distributed discovery mechanism for national digital geospatial data. All US Federal agencies are mandated to use this standard to document digital geospatial data. Many US States and local governments have also adopted this standard.

The FGDC standard defines an extensive set of metadata elements, a large number of which are included in the set of minimum elements of the standard. The benefits of FGDC over other standards such as ISO 19115 (see A.1.3.2) are in the provision for the description of the geographic database (e.g. an attribute table in a shape file). This facility is not present in ISO 19115 but is the subject of a separate ISO standard: 19110 – the methodology for feature cataloguing.

Although this standard has been widely adopted in the USA by both public and private bodies, the extensive nature of the set of minimum elements and their bias to specific federal applications means that it can be considered unsuitable for use in the UK and wider international geospatial community. In addition, whilst there is some standardisation in the technical encoding of this metadata, the encoding is far less well defined than other standards meaning that it is more difficult to provide a constrained and automated validation processes. This standard generally does not conform to more recent standards. Creation of metadata to this standard can be facilitated through a number of tools. More information can be found at www.fgdc.gov/metadata/geospatial-metadata-tools.The current FGDC standard provided one of the starting points for the development of ISO 19115 (A.1.3.2) and following the adoption by ANSI of ISO 19115 FGDC were tasked with developing a new standard: The North American Profile [of ISO 19115] (NAP). NAP is currently in development. Whilst the NAP is in development and aims to address the problems with the current FGDC standard, aiming to come into line with the wider international geospatial community, it is unlikely to conform to emerging European specific data standards such as INSPIRE. Further information can be found at www.fgdc.gov/metadata/geospatial-metadata-standards.

A.1.3.2 ISO 19115

The latest version of ISO 19115 was published 19 March 2014. It is an international standard to provide information about the identification, the extent, the quality, the spatial and temporal extent, the content, the spatial reference, the portrayal, distribution, and other properties of digital geographic data. It was developed by ISO in response to emerging, disparate metadata standards such as those produced by FGDC and ANZLIC. It provides the base schema definition and set of minimum elements and has been widely adopted as industry best practice and the base standard on which many more specific profiles are built. Further information can be found at http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=53798.

An extensive set of metadata elements are defined and several benefits are provided over FGDC including numerically coded pick lists, topic categories for high level metadata classification and prescribing a unique identifier for the metadata record. The set of minimum elements are concise, well defined, generalised and easy to populate. This standard does not identify a technical encoding in isolation, but ISO 19139 (see Annex D) defines the technical implementation of ISO 19115. In conjunction, the two ISO standards provide an excellent metadata standard, upon which many profiles have been built. As a basis for many other standards, ISO 19115 is not necessarily conformant with the minimum set of minimum elements of other standards. Creation of metadata to this standard can be facilitated through a number of tools. More information can be found at www.fgdc.gov/metadata/iso-metadata-editor-review.

A.1.3.3 INSPIRE

The latest version of the INSPIRE metadata implementing rules were published 18 February 2009. The INSPIRE directive aims to create an EU spatial data infrastructure. This will enable the sharing of environmental spatial information among public sector organisations and better facilitate public access to spatial information across Europe. The INSPIRE directive came into force on 15 May 2007 and will be implemented in various stages, with full implementation required by 2019. All public geospatial data must conform to the INSPIRE implementing rules under EU law at various different stages of the implementation. Further information can be found at http://inspire.jrc.ec.europa.eu/index.cfm.

The INSPIRE implementing rules define the use of both ISO 19115 and ISO 19139 as the basis for creating metadata for datasets (for services ISO 19119 is defined in place of ISO 19115) and it extends the set of minimum elements, adding (but not limited to) the inclusion of the INSPIRE themes from the GEneral Multilingual Environmental Thesaurus (GEMET) thesaurus as mandatory along with a conformance statement indicating whether and how the dataset conforms to the INSPIRE directive. The INSPIRE metadata implementing rules conform to ISO 19115 and ISO 19139.

Creation of metadata to this standard can be facilitated through the INSPIRE Geoportal: http://inspire.ec.europa.eu/index.cfm/pageid/381.

A.1.3.4 UK GEMINI standard – version 2.2

UK GEMINI was first published in 2004 by collaboration between the then e-Government Unit (e-GU), the Association for Geographic Information (AGI) and the UK Data Archive. The latest version was published December 2012.

UK GEMINI was designed to replace the National Geospatial Data Framework (NGDF) metadata standard with metadata that was compatible with ISO 19115. In 2006 a working group was set up to revise the standards. The new criteria included:

- Meet requirements for metadata of the EU INSPIRE Directive;
- Be conformant with the International Standard ISO 19115 (within the limits of the requirements of INSPIRE)
- Be consistent with the GEMINI 1.0 standards where possible

- Be compatible with the e-Government Metadata Standard where possible

Version 2.2 of UK GEMINI 2 includes changes requested by the above Working Group, including the extension to cover geospatial services, the removal of unnecessary elements and the addition of elements to conform to ISO 19115 mandatory core set and the INSPIRE Metadata Implementing Rules.

Further information can be found at http://www.agi.org.uk/uk-gemini/ .

UK GEMINI is widely conformant to existing standards and has a good set of minimum defining elements. It does not specify a technical encoding that should be followed when creating metadata, so in this way, UK GEMINI 2 is not fully conformant to the INSPIRE metadata implementing rules unless it is encoded in ISO 19139. Creation of UK GEMINI 2 metadata can be facilitated through means already cited in section A.1.3.2.

A.1.3.5 MEDIN discovery metadata standard – version 2.3.8

The latest version was published 14 March 2012. The MEDIN standard is a profile of ISO 19115 and has been born out of UK GEMINI. All data released by the MEDIN portal must comply with MEDIN standards. Further information can be found at http://www.oceannet.org/marine_data_standards/medin_disc_stnd.html.

The MEDIN standard has been designed to comply with UK GEMINI and provides several benefits over UK GEMINI specific to the marine data community such as marine thesaurus services to improve the ability to standardise keyword terms. The major different between MEDIN and UK GEMINI is that it prescribes the use of ISO 19139 for encoding metadata, making the standard fully compliant with INSPIRE, UK GEMINI 2, ISO 19115 and ISO 19139.

MEDIN offers the most extensive, up-to-date and relevant metadata standard for the UK marine sector and provides tools and advice on the creation of MEDIN discovery standard metadata. Metadata can be authored using the online tool: http://www.dassh.ac.uk/medin_metadata/login or by following the guidelines and using the conversion tool in conjunction with ArcCatalog, available from http://www.oceannet.org/submit_metadata/creating.html. MEDIN also provide a desktop tool for the creation of metadata http://www.oceannet.org/marine_data_standards/metadata_maestro.html. In addition to creation tools, MEDIN also provides a schematron file (see Annex D) to intelligently test metadata for validity against the MEDIN standard. This could be implemented in an automated work stream and increase the efficiency and validity of metadata production.

Annex B Metadata comparison

The metadata comparison below presents a comparison of the sets of minimum metadata elements of the different standards referenced in this document.

The scope of the comparison is restricted to metadata for datasets and where an element is conditional, the condition being that the metadata is being created for a dataset, the obligation will be represented as mandatory.

Broad descriptions are given, but the definition of elements varies slightly between standards. Please refer to the standard documentation for a full and proper definition as per the standard.

Metadata elements	Description	TCE notes	GEMINI V2.2	INSPIRE	ISO 19115	MEDIN	FGDC
Title	Name given to the data resource.	The title is used to provide a brief and precise description of the resource which in most cases will be a dataset. 'Date' 'Originating organisation/programme' 'Location' 'Type of survey'.	M	M	M	M	M
Abstract	Brief narrative summary of the data resource.	The abstract should provide a clear and brief statement of the content of the resource	M	M	M	M	M
Dataset Reference Date	Reference date for the data resource: Creation, Publication or Revision.	The date of publication is mandatory. Date of last revision or date of creation may also be provided.	M	M	M	M	M
Dataset Language	Language used in the data resource.	Recommend populating to conform with INSPIRE.	C (when contains textual information – multiple values allowed)	M	C (when contains textual information – single value allowed)	C (when contains textual information – multiple values allowed)	X
Resource Type	Scope to which metadata applies e.g. dataset, series, service.	If the resource type is anything other than dataset or series, it is out of scope.	M	M	M	M	X
Topic Category	Main theme(s) of the data resource.	This element is mandatory for INSPIRE and must be included for INSPIRE compliance. This indicates the main theme(s) of the data resource.	M	M	M	M	M
Keyword	Specific terms relating to the resource.	Recommend using a published thesaurus (e.g. GEMET).	M	M (must include at least 1 INSPIRE theme from GEMET thesaurus)	M	M	M
Originating Controlled Vocabulary (includes title, creation, publication, update date)	Name of the formally registered thesaurus or a similar authoritative source of keywords or extent identifier.	MaRS has its own taxonomy. Guidance can be provided.	C – required if keywords originate from a controlled vocabulary	C – required if keywords originate from a controlled vocabulary	C – required if keywords originate from a controlled vocabulary	M – Bespoke thesauri can be created	X
Temporal Extent	Date range for the content of the data resource.	It is recommended that all known temporal references of the resource are included.	M	M	O	M	M
Lineage	Information about the events or source data used in the construction of the data resource.	Lineage includes the background information, history of the sources of data used and can include data quality statements.	M	M	O	M	M
Geographic Bounding Box	North, east, west and south limit of the data resource extent, expressed in decimal degrees in any geographic coordinate reference system with a Greenwich Prime Meridian.	Can be populated automatically with ArcCatalog.	M	M	C – At least one of bounding box or extent by identifier must be populated	M	M

Metadata elements	Description	TCE notes	GEMINI V2.2	INSPIRE	ISO 19115	MEDIN	FGDC
Geographic Extent by Identifier	Extent of data resource by country or subdivision of country.		O	O	C – At least one of bounding box or extent by identifier must be populated	O	O
Spatial Reference System	Name or description of the system of spatial referencing used in the data resource.	Describes the system of spatial referencing (typically a coordinate reference system) used in the resource.	O	O	O	M	M
Spatial Resolution	Measure of the granularity of the data by distance or equivalent scale.	ISO 19139 allows for a “nil reference” which states the information is not available.	C	C – If available	O	M	M
Resource Locator	Location (address) for on-line access using a Uniform Resource Locator (URL) address or similar addressing scheme.	Formerly named online resource. If the resource is available online you must provide a web address (URL) that links to the resource.	C – when on-line access is available	C – when on-line access is available	C – when on-line access is available	C – when on-line access is available	C - when on-line access is available
Data Format	Format in which the digital data can be provided.	Recommended best practice is to select a value from a controlled vocabulary, for example PRONOM.	O	O	O	O	O
Data Point of Contact	Name of Organisation, email address, role of responsible organisation.	At least roles of Originator, Point of contact, and Distributor must be provided. Recommend using generic email address.	M – at least distributor role	M	M	M – at least originator, custodian and distributor roles if available	O
Frequency of Update	Frequency with which modifications and deletions are made to the data resource after it is first produced.	This describes the frequency that the resource (dataset) is modified or updated and should be included if known.	O	M	M	M	M
Use limitations	Restrictions imposed on the data resource for security and other reasons.	This element describes any restrictions imposed on the resource for security and other reasons i.e. not for navigation.	M	M	M	M	M
Use/Access/Other Constraints	Restrictions and legal restraints on using the data resource.	This element describes any restrictions and legal restraints on using the data i.e. copyright.	M	M	M	M	M
Unique Resource Identifier	Value uniquely identifying the data resource.	To conform to INSPIRE, must include a code and codes pace (website). See documentation for example.	M	M	M	M	M
Resource Type	Scope to which metadata applies – i.e. dataset/service.	The resource type must be a dataset or a series for any data supplied to The Crown Estate. Other resource types include Service.	M	M	M	M	M

Metadata elements	Description	TCE notes	GEMINI V2.2	INSPIRE	ISO 19115	MEDIN	FGDC
Conformity Statement/Specification	Degree to which the data conforms with a data standard (e.g. INSPIRE – data implementing rules not released).	MaRS recommends INSPIRE conformance where possible.	C – Required if conforming to INSPIRE	M	C – Required if conforming to INSPIRE	C – Required if conforming to INSPIRE	N/A
Metadata Identifier	Unique identifier for the metadata record.		M	O	O	M	X
Metadata Standard Name	Name of the Metadata standard.		M	O	O	M	M
Metadata Language	Language used for documenting the metadata.		C – Required if conforming to INSPIRE	M	C	M	X
Metadata Date	Date the metadata was: Created, Published, Revised.		M	M	M	M	M
Metadata Point of Contact	Party responsible for the creation and maintenance of the metadata. Must include organisation name, role (point of contact), email address.	Recommend using generic email address.	M	M	M	M	M
Metadata Security Information	Restrictions on metadata because of national security.	Specific to US federal data.	X	X	X	X	M

Annex C Metadata creation

C.1.1 Introduction

As indicated earlier in this guidance note, there are various methods by which to author metadata. This section is focussed on methods of creating MEDIN discovery standard metadata v2.3.8. The various options will be discussed in terms of their ease of use, the opportunities for development, automation, validation and transformation of and between metadata standards.

Essentially, creation of metadata to the MEDIN standard can be facilitated by any ISO 19115 metadata creator/editor. The pitfalls of using any option are that there is unlikely to be the facility to indicate where metadata is mandatory according to MEDIN, and that the output may not conform to the ISO 19139 XML encoding and therefore require further validation. This reduces the usability of the tools and increases the level of knowledge required to create metadata. In essence, the easier a tool makes it to create conformant metadata, the quicker the task can be completed.

C.1.2 Creation

C.1.2.1 The options

The options that will be discussed in this section are:

- The INSPIRE Geoportal metadata creator
- The MEDIN online metadata creator
- The MEDIN desktop tool: Metadata Maestro
- ArcCatalog (version 9.3 & 10)

C.1.2.2 INSPIRE geoportal

http://inspire-geoportal.ec.europa.eu/

The INSPIRE Geoportal provides an online interface for creating metadata to the INSPIRE metadata implementing rules. This is largely similar to MEDIN, but caution must be taken to ensure that all of the minimum elements as prescribed by MEDIN are populated.

The tool is relatively easy to use and provides indication as to which fields are mandatory by selecting validation to be on. The tool also provides easy access to the GEMET thesaurus from which it is possible to select the INSPIRE data themes, however it is not possible to directly select the MEDIN keywords, so they have to be entered manually. There is no provision to enter a data format.

The output from this tool conforms with ISO 19139, but the metadata standard name is cited as ISO 19115, version 2003/Cor.1:2006. This will need to be altered either manually, or by writing a simple Extensible Stylesheet Language Transformation (XSLT) to transform this to cite MEDIN.

This tool does not provide any validation at a higher level than indicating the minimum set of required elements and does not provide the ability to import or export existing metadata

in different formats, or create templates to reduce the amount of repetition between the authoring of different metadata records. It is not possible for individual users to develop this tool.

C.1.2.3 MEDIN online metadata creator

http://www.dassh.ac.uk/medin_metadata/login

The MEDIN online metadata creator provides a web-based interface for creating MEDIN compliant metadata. The tool is relatively easy to use and provides indication as to the obligation of elements that are made available to populate. It provides many services and lists to help the user input keywords including a list of published INSPIRE implementing rules and the INSPIRE themes, along with a link to the EPSG registry for unambiguously selecting the coordinate reference system. All metadata created by a specific user are saved, so it is possible to edit this record and export it again as a new metadata record, and to save time where there is general overlap in the information input e.g. metadata contact, geographic bounding box etc.

For beginners, this is the recommended tool to use.

C.1.2.4 MEDIN desktop Metadata Maestro

http://www.oceannet.org/marine_data_standards/metadata_maestro.html

The MEDIN desktop Metadata Maestro provides an offline interface for creating MEDIN compliant metadata. The tool is relatively easy to use and provides indication as to the obligation of elements that are made available to populate. It provides many services and lists to help the user input keywords including a connection to the MEDIN vocabulary services and the INSPIRE themes, along with a link to the EPSG registry service for unambiguously selecting the coordinate reference system. It is possible to create drafts and templates to save time where there is general overlap in the information input, to edit these and export as a new metadata record.

Metadata Maestro also provides the ability to perform a powerful batch validation and transformation using XSLT and an XML text editor to directly edit the XML code (though the latter two options are only recommended for advanced users who have experience of XML).

For intermediate users or those that require an offline method, this is the recommended tool.

C.1.2.5 ArcCatalog

ArcCatalog, the data management component of the ArcGIS desktop suite, provides several tools to aid the creation of metadata. Extensive documentation can be found at:

V9.3.x http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=About_metadata

V10
http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//00qp0000001p000000.htm

V9.3.x

ArcCatalog supports the creation and maintenance of FGDC and ISO 19115 standard metadata by default. It populates elements from both or either of these standards into the same XML document, encoded in the proprietary ESRI-ISO schema. There are several options associated with metadata creation in ArcCatalog: the default editor; when metadata should be created; and which elements are populated and maintained automatically via synchronisers.

The metadata tab of the options dialogue box provides the ability to control the first two of these options. It is recommended that the ISO wizard is selected as the default metadata editor.

The synchroniser options are more hidden. The user has to select tools>customise> search for “metadata” in the categories tab> drag the set working synchronisers icon to the toolbar to expose the options. There are three default synchronisers (4 if you have productivity suite): FGDC, ISO and Geography Network (plus a Gemini synchroniser if you have productivity suite). All are turned on by default. It is recommended that the ISO synchroniser is left active, the others turned off as they duplicate information and confuse the metadata record.

Finally, there are different stylesheets that can be selected to view the metadata in more human-readable form. These simply convert the XML “on-the-fly” into html and display in a web-style browser. For ERSI-ISO encoded metadata, created with the ISO wizard, select the ISO stylesheet from the dropdown list on the metadata toolbar, for ISO 19139 metadata select the ISO 19139 stylesheet from the drop down. A MEDIN stylesheet has also been produced and can be easily added to this list following ESRI guidelines.

ESRI provide guidance and the option of using ArcObjects to create custom synchronisers and metadata editors. These could be utilised to create a MEDIN specific metadata creation facility, fully integrated within an ESRI workflow.

V10, 10.1 and 10.2

Version 10, version 10.1 and version 10.2 implements metadata subtly differently from 9.3.x. Metadata here is based on “styles” which essentially expose certain elements from a core pool depending on the standard selected. Therefore transformation into MEDIN metadata can be undertaken using a stylesheet provided by MEDIN, which transforms the ESRI ArcGIS metadata XML format to the ISO 19139 XML format that is used to encode MEDIN metadata.

Use service pack 3 and above to set up ArCatalaog v10, 10.1 or 10.2 to create MEDIN metadata.

For the tools required to create MEDIN discovery metadata see
http://www.oceannet.org/submit_metadata/creating.html

Annex D Technical encoding and metadata validation

Having a well-structured, well defined and internationally adoptable technical encoding for metadata enables the automation of validation and interoperability between systems that ingest, create, and manage metadata records. For this reason, XML has been widely adopted as the de-facto format in which metadata are stored and transferred. There are two main encodings considered here: ESRI-ISO and ISO 19139.

Although Annex C touched on validation, none of the tools mentioned thus far provide the ability to validate a metadata record beyond indicating which elements are mandatory, conditional, or optional. XML does however provide ample opportunity for validation to be carried out. It is possible to validate an XML document against a Document Type Definition (DTD) which allows for validation against a defined grammar and structure. This is the only level of validation document defined for the ESRI-ISO metadata encoding which is one of the major limitations of this encoding as it does not allow code lists to be validated against, and the encoding itself is lacking in data-type constraint (ie all elements are encoded as free-text, so, for example, a date cannot easily be checked for format (e.g. 2010-06-03 vs. 03/06/2010)).

ISO 19139 defines a set of XML Schema Definition (XSD) documents that can be used in place of a DTD and are far more powerful in that they define namespaces, use rich data-typing so that a more complex interrogation can be made to check structure, content (e.g. date format), and code lists can be explicitly defined within an XSD to ensure that some content can be validated.

Lastly, a schematron can be used which looks for XML tree patterns and element content to validate against and provides a much more 'content rich' validation, providing more information to the author as to where content has been found that differs from that expected (e.g. a start date is before an end date in the temporal extent).

Whilst these steps are relatively straightforward to implement, they rely on the document to validate against being available, and having the tool available to process the validation.

ISO 19139 provides a set of XSD documents which can be validated against. MEDIN provide a schematron file that can be validated against (available at http://www.oceannet.org/submit_metadata/validating.html). These options can be incorporated and automated in an XML pipeline or incorporated during the metadata creation process by a relatively experienced XML developer with experience of metadata standards. There are however some tools available that allow for this validation such as XMLspy and Oxygen XML.

For examples of well written metadata records, go to http://www.oceannet.org/submit_metadata/creating.html

For more information please contact Peter Edmonds, Spatial Data Manager at The Crown Estate:

peter.edmonds@thecrownestate.co.uk / Tel 020 7851 5349

Survey and Information Management Procedures, Guidelines, Standards and Best Practice

1 Introduction

This guidance note provides renewable energy developers with links to information regarding survey and information management procedures, guidelines, standards and best practice. The references are not exhaustive but present key documents and information on sources of information on survey and information management. Discussion of the content and suggestions of alternative sources are welcomed.

By presenting the procedures etc. that could be adopted, it is intended that a consistent approach to this matter can be applied throughout the project life-cycle and across the renewable energy industries.

Consideration of the procedures etc. when drafting technical specifications for surveys will ensure that data are collected and deliverables are presented in appropriate formats. This will maintain the quality of the data sets and promote data interoperability. In addition it ensures the maximum longevity for the data and the potential for re-use and re-analysis will be enhanced, reducing risks and costs in the future. They should also be referred to when designing data and information management systems.

2 Management Procedures, Guidelines, Standards and Best Practice

2.1 MEDIN

Marine Environmental Data and Information Network (MEDIN) is a partnership of UK organisations committed to improving access to marine data (http://www.oceannet.org/). Their partners are both public and private sector and include The Crown Estate. MEDIN reports directly to the Marine Science Coordination Committee (MSCC).

MEDIN provides a variety of data guidelines and standards for metadata generation. The MEDIN data model is accepted best practice across government Data Archive Centres and will be implemented by The Crown Estate in the archiving of observed survey data provided by renewable energy developers and by those granted a small works consent.

The data guidelines set out the requirements for information that must be recorded when a certain type of data is being collected (e.g. cetacean sighting and identification data) to ensure that the data can be understood, interpreted and used with confidence now and in

the future. These guidelines instil good practice, aid organisations in specifying formats that data should be returned in that can be readily used and include all relevant attributes and ensure that the data collected are of maximum value. More information and the data guidelines are available from http://www.oceannet.org/marine_data_standards/medin_data_guidelines.html. The Crown Estate strongly recommends the adoption of these standards and will be using these when performing quality checks on data provided by renewable energy developers.

The MEDIN discovery metadata standard (http://www.oceannet.org/marine_data_standards/medin_disc_stnd.html) sets out the information that should be recorded to enable efficient discovery and cataloguing of data. The standard complies with internationally accepted standards for metadata and forms the minimum requirement for provided data to The Crown Estate. The benefit of using this standard is an easily transferrable, easily validated catalogue of information to help the user understand a data set and data packages on a consistent basis using controlled vocabularies and accepted best practice.

The Crown Estate strongly recommends contacting MEDIN for help and advice with marine survey data management.

2.2 COWRIE

Prior to the Marine Data Exchange (MDE), survey data held by The Crown Estate relating to marine data for offshore renewable developments was managed through The Crown Estate's Collaborative Offshore Wind Research Into the Environment (COWRIE) website. COWRIE was set up by The Crown Estate as an independent body to carry out research into the impact of offshore wind farm development on the environment and wildlife. The list of COWRIE reports can be viewed here: http://www.thecrownestate.co.uk/energy-and-infrastructure/downloads/cowrie/.

All of this data is now managed entirely through the MDE and any information that was previously stored on the COWRIE system has been migrated across.

2.3 Other Key Guidance

The below table presents other key documents to support survey and information management.

Standard / best practice / guidance type	Title	Author	Date	Notes / Web links
Biological	Establishing best practice for the documentation and dissemination of marine biological data	MBA	2008	http://www.thecrownestate.co.uk/media/5902/2008-09%20Establishing%20best%20practice%20for%20the%20documentation%20and%20dissemination%20of%20marine%20biological%20data.pdf

Coastal Processes Modelling	COWRIE COAST -07-08, "Coastal Processes Modelling for Offshore Wind Farm Environmental Impact Assessment: Best Practice Guide"	COWRIE	2008	http://www.thecrownestate.co.uk/media/5903/2009-09%20Coastal%20process%20modelling%20for%20offshore%20wind%20farm%20EIA%20-%20best%20practice%20guide.pdf
Geophysical	Marine Survey Data Management Handbook, Internal Report IR/08/024 BGS Specifications	BGS	2008	http://www.bgs.ac.uk/downloads/browse.cfm?sec=1&cat=112
Hydrographic	Marine Guidance Note 371. Offshore Renewable Energy Installations (OREIs) - Guidance on UK Navigational Practice, Safety and Emergency Response Issues.	MCA	2008	The Crown Estate strongly recommends referring to Annex 2, Section 6 of this note. http://www.dft.gov.uk/mca/mgn371.pdf
Hydrographic	Manual On Hydrography. Publication M-13. 1st Edition,	IHO	2005	http://download.eiva.dk/online-training/Manual%20on%20Hydrography%20-%20IHO.pdf
Hydrographic	IHO Standard S-44 STANDARDS FOR HYDROGRAPHIC SURVEYS	IHO	2008	http://www.iho.int/iho_pubs/standard/S-44_5E.pdf
Hydrographic	IHO Standard S-57 Transfer Standard for Digital Hydrographic Data	IHO	2000	http://www.iho.int/iho_pubs/standard/S-57Ed3.1/31Main.pdf
Hydrographic	IHO Standard S-100 Universal Hydrographic Data Model	IHO	2010	http://www.iho.int/mtg_docs/com_wg/TSMAD/TSMAD_Misc/S-100InfoPaper_FinalJan2011.pdf
Geodetic	Guidance notes on the use of co-ordinate systems in data management on the UKCS	UKOOA	1999	https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/15094/2888-guidance-notes-use-of-coordinate-systems.pdf
Data Management	Marine Data Guidelines	MEDIN	2011	http://www.oceannet.org/marine_data_standards/medin_data_guidelines.html
Data Management	Metadata standards	MEDIN	2011	http://www.oceannet.org/marine_data_standards/medin_disc_stnd.html

London
The Crown Estate
16 New Burlington Place
London
W1S 2HX
T 020 7851 5000

Edinburgh
The Crown Estate
6 Bell's Brae
Edinburgh
EH4 3BJ
T 0131 260 6070

Windsor
The Crown Estate
The Great Park
Windsor, Berkshire
SL4 2HT
T 01753 860 222

www.thecrownestate.co.uk
@TheCrownEstate