Title	**Brief guidance notes for the production of discovery metadata for the Marine Environmental Data and Information Network (MEDIN)**
MEDIN Discipline	Discovery Metadata
Author(s)	Becky Seeley (DASSH), James Rapaport, (SeaZone), Olivia Merritt (SeaZone), Mark Charlesworth (BODC), Sean Gaffney (BODC).
Document Owner	Marine Environmental Data Information Network (MEDIN) Standards Working Group
Reviewed by	MEDIN Standards Working Group
Date reviewed	05 March 2019
Version	3.0
Date approved and published on MEDIN website	08 March 2019
Date last checked for accuracy	04 March 2019
Summary	The discovery metadata standard for resources submitted to the Marine Environmental Data Information Network.
Keywords	Discovery Metadata

Metadata standards are evolving at an international level and these guidelines are therefore subject to change.

It is recommended that you use a download of this document from the Marine Environmental Data and Information Network (MEDIN) website (www.medin.org.uk) rather than storing a local copy. A log of changes will be available on the website.

MEDIN Elements

1. Introduction

There has been a request by the wider community to produce a simpler guide to the MEDIN Discovery Metadata Standard. This document fulfils that need and is largely a thinned down version of the full guidance document, without any xml examples or annexes. Metadata standards are essential to enable easy discovery, evaluation and use of resources. In most cases within MEDIN the resource will be a dataset, however model outputs and services such as web mapping services and data download services are also included. Different sorts of standards are applied for discovering a dataset, service[1] or series (collectively known as resources), evaluating its fitness for purpose and in providing the information required to use it. This standard is one that sets out a specific format to record details of a resource so that in the future other people can easily discover resources that may be of use to them. It is therefore termed a 'metadata discovery standard' and this document sets out the format used by the Marine Environmental Data Information Network (MEDIN). All metadata released via the MEDIN portal must comply with a number of international and national metadata standards. The MEDIN metadata schema is based on the ISO 19115:2003 standard, and includes all core INSPIRE metadata elements. It also complies with the UK GEMINI 2.3 metadata standard. The xml produced conforms to the ISO 19139 standard for xml implementation.

This document is designed to assist those creating metadata for MEDIN and provides guidance on how to complete each element. Please refer to the INSPIRE metadata implementing rules, http://inspire.jrc.ec.europa.eu/ rules and UK GEMINI 2.3 specification https://www.agi.org.uk/gemini/40-gemini/1037-uk-gemini-standard-and-inspire-implementing-rules for additional information. In writing this document reference has been made to the Technical Guidance for the implementation of INSPIRE dataset and service metadata based on ISO/TS 19139:2007 (see guidelines at http://inspire.ec.europa.eu/document-tags/metadata)[2].

Metadata standards may change over time. It is recommended that this document is downloaded regularly to ensure the most current version is in use.

2. Data Discoverability

It is important that other users of MEDIN can find out how to access the raw data or products by using the information held in this standard. Therefore, where available, links should be provided to web pages and/or contact details of the person who holds the dataset. If there is a direct web link to the dataset or service then it should be stated in Element 5 ‘Resource Locator’. Further information such as related documents and links to other portals that may also hold information on the dataset, should be given in Element 19 ‘Additional Information’ and the contact details of the person who holds the dataset should be given in Element 22 ‘Responsible Party’.

Often it is difficult to decide if the data that has been collected constitutes one dataset or many - this is called ‘granularity’. It is important to get the level or ‘granularity’ correct otherwise it is possible to end up with either too many or too few records which makes it difficult for a user to find what they want via a portal. MEDIN has some practical guidance

---

[1] See Element 4 for definition of a service.

[2] INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119, 2009-02-18, Version 1.1, MD_IR_and_ISO_20090218.
Technical Guidance for the implementation of INSPIRE dataset and service metadata based on ISO/TS 19139:2007

to help you decide:

- The correct level for a dataset is a cruise, survey or a set of repeat observations with a common purpose.
- A dataset usually constitutes a specifically-funded piece of work.
- The dataset should be easily extractable from a database for a 3rd party.
- If you are searching for a dataset using a portal and get the result every time you search by different combinations of time, location and parameter then it is probably too coarse.

3. Elements for identifying a resource

Element 1 - Resource title (M)

Mandatory element. Only one occurrence allowed. Free text.

The title is used to provide a brief and precise description of the resource, which in most cases will be a dataset. MEDIN recommend the following format:

'Date' 'Originating organisation/programme' 'Location' 'Type of survey'. It is advised that acronyms and abbreviations are reproduced in full. Example: Centre for Environment, Fisheries and Aquaculture Science (Cefas).

If acronyms cannot be reproduced in full in the Title element, they must be fully expanded in one of the Resource Abstract or Alternative Resource Title elements.

Example 1

1992 Centre for Environment, Fisheries and Aquaculture Science (Cefas) North Sea 2m beam trawl survey.

Example 2

1980-2000 Marine Life Information Network UK (MarLIN) Sealife Survey records.

Element 2 - Alternative resource title (O)

Optional element. Multiple occurrences allowed. Free text.

The purpose of alternative title is to record any additional names by which the resource (e.g. dataset) may be known and may include short name, other name, acronyms or alternative language title e.g. Welsh language title of the same resource. If including acronyms in the text, they should be expanded in full if the full term has not been stated in the Resource title element.

Example

1980-2000 MarLIN Volunteer Sighting records.

Element 3 - Resource abstract (M)

Mandatory element. Only one resource abstract allowed. Free text.

The abstract should provide a clear and brief statement of the content of the resource (e.g. dataset). It shall be a minimum of 100 characters in length and shall not be a duplicate of

the title. Metadata creators should include what has been recorded, what form the data takes, what purpose it was collected for, and any limiting information, i.e. limits or caveats on the use and interpretation of the data. Background methodology and quality information should be entered into the Lineage element (Element 17). It is recommended that acronyms and abbreviations are reproduced in full e.g. Centre for Environment, Fisheries and Aquaculture Science (Cefas).

Restrictions relating to spatial resolution for metadata for services shall be expressed in Resource abstract if they exist, and not in Element 18 Spatial Resolution.

Example 1

Benthic marine species abundance data from an assessment of the cumulative impacts of aggregate extraction on seabed macro-invertebrate communities. The purpose of this study was to determine whether there was any evidence of a large-scale cumulative impact on benthic macro-invertebrate communities as a result of the multiple sites of aggregate extraction located off Great Yarmouth in the North Sea.

Element 4 - Resource type (M)

Mandatory element. One occurrence allowed. Controlled vocabulary.

Identify the type of resource using the controlled vocabulary, MD_ScopeCode from ISO 19115. The resource type shall be a dataset, a series (collection of datasets with a common specification) or a service. In the vast majority of cases for MEDIN the resource type will be a dataset or a series. Further information on the difference between a dataset and a series and the definition of a service is available at http://www.medin.org.uk/medin/data/faqs.

Example

series

Element 31 – Hierarchy level name (C)

Conditional element (shall be completed when Resource type is not “dataset”). Single occurrence allowed. Free text.

This is the name of the hierarchy level for which the metadata is provided. It should be used in conjunction with Resource type to provide users with information on the hierarchy of data within the resource.

Example 1

series

Example 2

Collection of observations over time

Example 3

map service

Element 5 - Resource locator (C)

Conditional element (shall be completed when online access is available). Multiple occurrences allowed. Free text.

Formerly named online resource. If the resource is available online you must provide a web address (URL) that links to the resource. If there is no online access to the resource but there is a publicly available online resource providing additional information about the described resource, a link to this information resource should be provided instead. This element should be used to provide the URL of any Digital Object Identifier (DOI) landing page(s) for the data resource.

Example 1

Resource locator URL:
https://doi.org/10.5285/481720C2-35BD-6C10-E053-6C86ABC06BB3
Resource locator name: An improved database of coastal flooding in the United Kingdom from 1915 to 2016
Resource locator function: information
Resource locator description: URL accesses a landing page (at the British Oceanographic Data Centre) for the UK database of coastal flooding from 1915 to 2016, allowing interested parties to download the data anonymously.

Element 6 - Unique resource identifier (M)

Mandatory element for datasets and series of datasets. Multiple occurrences allowed. Free text.

A Unique Resource Identifier allows a resource to be identified by a code. This code is generally assigned by the data owner and commonly consists of the organisation that manages the dataset and a number or code which is used to uniquely identify it within the databases of the organisation. If this code is unique then it is possible for an organisation to identify a dataset that a 3rd party may be referring to and also to quickly identify where dataset records may be duplicated in a portal.

The two parts to the element can either be provided separately as a code + a codespace or combined as 1 code. MEDIN recommends the use of code + a codespace as shown in example 1. Preferably the www address of the organisation should be given rather than the organisation acronym or name. *The code and the codespace should not include any spaces.* If you are unable to generate a Unique Identifier Code please contact medin.metadata@mba.ac.uk and we will generate a code for you or endeavour to provide a tool to generate your own codes.

Where present, a resource DOI should be recorded as a resource identifier, with the code reflecting the DOI and codespace being ‘doi’.

Example 1

Code: 5639287
Codespace: http://www.bodc.ac.uk

Example 2

Code: http://www.bodc.ac.uk/5639287

Example 3

Code: doi:10.5285/481720c2-35bd-6c10-e053-6c86abc06bb3
Codespace: doi

Element 7 - Coupled resource (C)

Conditional element. Not applicable to datasets or series. Mandatory for View and Download services, optional for other service types. Multiple occurrences allowed. Free text.

This identifies the data resource(s) on which the service operates. Each occurrence shall be a URL that points to the metadata record of the data on which the service operates

Example

http://portal.oceannet.org/portal/start.php#details?tpc=006_00806134608655879d57842c8138b1ff

Element 8 - Resource language (C)

Conditional element. Mandatory for datasets and series, not applicable to services[3]. Multiple occurrences allowed. Controlled vocabulary, ISO 639-2.

Describes the language(s) of any textual information contained within the resource.

Select the relevant 3-letter code(s) from the ISO 639-2 code list of languages. Additional languages may be added to this list if required. A full list of recognized languages is available online at http://www.loc.gov/standards/iso639-2/php/code_list.php.

For Welsh, ISO 639-2 allows either of ‘cym’ or ‘wel’, but MEDIN recommend that ‘cym’ is used as this is the abbreviation of the language’s own name for itself. This follows guidance from GEMINI.

If there is no textual information in the data resource, then the code value **zxx** from ISO 639-2/B for ‘no linguistic content; not applicable’ shall be used.

Example 1

eng (English)

[3] See Element 4 resource type for definition of a ‘service'

4. Elements classifying spatial data and services

Element 9 - Topic category (C)

Conditional element. Mandatory for datasets and series of datasets. This element is not required if a service[4] is being described. Multiple occurrences allowed. Controlled vocabulary.

This indicates the main theme(s) of the data resource. The purpose of this element is to provide a basic classification for the data resource, for use in initial searches. The relevant topic category/categories shall be selected from the ISO MD_TopicCategory list. The full list can be found in controlled vocabulary library P05 on the NVS2 Vocabulary Server https://www.bodc.ac.uk/resources/vocabularies/vocabulary_search/P05/.

MEDIN have mapped the MEDIN keywords (see element 11) to the ISO Topic Categories, so it is possible to generate the topic categories automatically once MEDIN keywords have been selected from the SeaDataNet Parameter Discovery Vocabulary (P02) https://www.bodc.ac.uk/data/codes_and_formats/vocabulary_search/P02/.

Example 1

biota

Element 10- Spatial data service type (C)

Conditional element. Mandatory if the described resource is a service[5]. Not applicable to datasets or series. One occurrence allowed. Controlled vocabulary, INSPIRE Service type code list

An element required by INSPIRE for metadata about data services e.g. web services. If a service is being described (from Element 4) it must be assigned a service type from the INSPIRE Service type code list.

Example

Download

Element 32 – Spatial representation type (C)

Conditional element. Mandatory for datasets and series of datasets. This element is not required if a service[8] is being described. Multiple occurrences allowed. Controlled vocabulary, MD_SpatialRepresentationTypeCode from ISO 19115

This element describes the method used to spatially represent geographic information. The type in which the spatial data is represented may be of importance when evaluating the fit for purpose of the datasets.

This element is regarded by the INSPIRE metadata technical guidance as interoperability

[4] See Element 4 resource type for definition of a ‘service'

[5] See Element 4 resource type for definition of a ‘service'

metadata for datasets and series. The element shall be populated with the code from MD_SpatialRepresentationTypeCode that most appropriately describes the resource.

Example

grid

Element 11 - Keywords (M)

Mandatory element. Multiple keywords allowed. Controlled vocabularies.

The purpose of this element is to indicate the general subject area(s) of the data resource using key words. This enables searches to eliminate resources that are of no interest to users.

Keywords should be chosen using the code list options given below. OAI harvesting keywords should be linked to the data resource as described below if the metadata record is being submitted to MEDIN and to data.gov.uk.

In addition, if a spatial data service is being described, then a keyword defining the category or subcategory of the service using its language neutral name as defined in Part D 4 of the Metadata Implementing Rules shall be given.

The entry shall consist of two sub-elements: the keywords and reference to the controlled vocabulary used as shown in the sub elements below.

INSPIRE keywords (M)

MEDIN require at least one INSPIRE theme keyword as this ensures INSPIRE compliance.
A list of the INSPIRE theme keywords is available at http://www.eionet.europa.eu/gemet/inspire_themes or library P22 in the NVS2 Vocabulary Server https://www.bodc.ac.uk/data/codes_and_formats/vocabulary_search/P22/.

MEDIN Keywords (C)

The contents of the dataset shall be described using the SeadataNet Parameter Discovery Vocabulary (P02), unless there are no applicable terms in the list. This improves the discoverability of datasets by using terms related to the marine domain.

The P02 terms are available at https://www.bodc.ac.uk/data/codes_and_formats/vocabulary_search/P02/. The parameter groups and codes that are used may also be searched hierarchically through a user friendly interface which has been built as part of the European funded SeaDataNet project at http://seadatanet.maris2.nl/v_bodc_vocab_v2/vocab_relations.asp?lib=P08.

Vertical Extent Keywords (C)

Element 11: ‘vertical extent keyword’ shall be populated only if Element 14: ‘Vertical extent information’ cannot be completed.

A vocabulary of keywords is available to describe the vertical extent of the resource (e.g. dataset). The vocabulary is available as library L13 (Vertical Coordinate Coverages) at https://www.bodc.ac.uk/data/codes_and_formats/vocabulary_search/L13/.

Other Keywords (O)

Keywords from other vocabularies may be used as required, as long as they follow the

format specified in 11.

Take care that selected keywords do not duplicate information that is used to populate other Elements in the Profile e.g. selection of sea area keywords, which should go into Element 13: ‘Extent’.

Keywords for services (C)

If a service is being described, the category or subcategory of the service shall be described using its language neutral name. This is defined in Part D 4 of the Metadata Implementing Rules which can be found at http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:326:0012:0030:EN:PDF and the keyword vocabulary available at http://inspire.ec.europa.eu/registry/.

Making Metadata Available to the MEDIN portal and data.gov.uk via OAI, CSW and WAF

If XML files are being collected using the MEDIN harvesting process, an additional keyword is required to allow the discovery web service to distinguish MEDIN records. The required term to use in the XML fragment is NDGO0001 (from the N01 controlled vocabulary at https://www.bodc.ac.uk/data/codes_and_formats/vocabulary_search/N01/). If you wish your discovery metadata records to also be made available to the UK Geoportal ‘data.gov.uk’ via MEDIN then you should include the additional term NDGO0005 i.e. Include both NDGO0001 and NDGO0005 in keywords to indicate a record will be published to both portals.

Example 1

keywordValue: Fish taxonomy-related counts
keywordValue: Temperature of the water column
thesaurusName: SeaDataNet Parameter Discovery Vocabulary
dateType: revision
date: 2009-10-13

Example 2
keywordValue: upper_epipelagic
thesaurusName: SeaDataNet vertical coverage
dateType: Creation
date: 2006-11-15

Element 12 - Geographic bounding box (C)

Mandatory element for datasets and series, conditional for services on their being a defined extent for the service. Multiple occurrences of each sub-element allowed. Numeric and controlled vocabulary.

The purpose of this element is to record the geographic extent that is covered by the metadata resource. This extent range is recorded as one or more bounding boxes that have positional information expressed as decimal degrees longitude and latitude. A minimum of two decimal places shall be provided for each coordinate.

Multiple bounding boxes are acceptable and can be used to describe resources that have a disparate geographic coverage; each bounding box must only have one occurrence of

the east, west, north and south sub-elements.

Latitudes between 0 and 90N, and longitudes between 0 and 180E should be expressed as positive numbers, and latitudes between 0 and 90S, and longitudes between 0 and 180W should be expressed as negative numbers. In the event that a single point is being described we recommend using the en-coding shown in the last example.

The latitude and longitude of the bounding box(es) is implicitly in WGS84.

Example

westBoundingLongitude: -4.351
eastBoundingLongitude: -1.348
northBoundingLatitude: 52.949
southBoundingLatitude: 52.117

Element 13 - Extent (O)

Optional element. Multiple occurrences allowed. Numeric and controlled vocabulary.

This element defines the geographic extent of coverage of the data resource relative to a defined authority. Keywords should be selected from controlled vocabularies to describe the spatial extent of the resource. MEDIN strongly recommends the use of the SeaVoX salt and freshwater body gazetteer available as vocabulary C19 at https://www.bodc.ac.uk/data/codes_and_formats/vocabulary_search/C19/, which is a managed vocabulary and has a worldwide distribution.

Other vocabularies available, including ICES areas and rectangles http://vocab.ices.dk/, or Charting Progress 2 regions, may be used as long as they follow the format specified in 13 (these are available as vocabulary C64 at https://www.bodc.ac.uk/data/codes_and_formats/vocabulary_search/C64/).

Example

This example includes multiple extents from different vocabularies.

extentName: Scotland
vocabularyName: ISO3166 Countries
dateType: Creation
date: 2005-04-29

extentName: ICES Area IVb
vocabularyName: ICES Regions
dateType: Revision
date: 2006-01-01

extentName: Northern North Sea
vocabularyName: Charting Progress 2 regions.
dateType: Revision
date: 2008-09-01

extentName: North Sea
thesaurusName: IHO Sea Areas 1952

dateType: creation
date: 1952-01-01

Element 14 - Vertical extent information (C)

Conditional element. This element shall be filled in if the vertical coordinate reference system is known. Multiple occurrences allowed. Numeric free text and controlled vocabulary.

This element shall be filled in if the vertical Coordinate Reference System (CRS) is registered in the 'European Petroleum Survey Group (EPSG) database. http://www.epsg-registry.org/.

If you do not have the defined CRS you shall complete the vertical extent vocabulary defined in Element 11 – Keywords, to describe the vertical extent of the resource.

One of the elements '11: vertical extent keyword' or '14: vertical extent information' must be completed.

The vertical extent element has three sub-elements; the minimum vertical extent value, the maximum vertical extent value, and the coordinate reference system. Depth below sea water surface should be a negative number. Depth taken in the intertidal zone above the sea level should be positive. If the dataset covers from the intertidal to the subtidal zone then the sub element 14.1 should be used to record the highest intertidal point and 14.2 the deepest subtidal depth. Although the element itself is optional its sub-elements are mandatory if the field is filled.

For services, this element should be used to record the maximum vertical boundaries of all resources covered by the service.

Example

minimumValue: 42
maximumValue: 94
verticalCoordinateReferenceSystem: urn:ogc:def:crs:EPSG::5701

Element 15 - Spatial reference system (M)

Mandatory element. Multiple occurrences allowed. Controlled vocabulary.

Describes the system of spatial referencing (typically a coordinate reference system (CRS)) used in the resource. This should be derived from the EPSG register of geodetic parameters (http://www.epsg-registry.org/).

To find a code click on the EPSG link above and if you know the title (e.g. WGS84) then type this in the 'Name' field and click search. The name, code and further information are displayed. If you are looking for a specific type of reference system such as 'vertical' then click in the 'Type' box, hover over coordinate reference system and click on vertical and then click the search button and all recorded vertical reference systems are shown. If you want to search for a reference system in a particular part of the world (e.g. Northern Ireland Grid) the you may do so by submitting a term to the 'Area' box or fill out the latitude and longitudes then click search. The website also provides a database of the reference systems and web services to access the information.

Example 1 (for WGS84)

Code: urn:ogc:def:crs:EPSG::4326

Example 2 (for National Grid of Great Britain)

Code: urn:ogc:def:crs:EPSG::4277
Thesaurus name: EPSG Geodetic Parameter Registry
Date type: revision
Date: 2016-09-29

Element 16 - Temporal reference (C)

Mandatory for datasets and series; conditional for services where a temporal extent is relevant to the service. Multiplicity as stated below. Controlled vocabulary and Date/Time format, yyyy-mm-dd or yyyy-mm-ddThh:mm:ss

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 for datasets and series of datasets and shall be provided. Temporal extent and data of publication should be provided for services where a temporal extent is relevant to the service. The date of last revision or date of creation for the resource may also be provided. One occurrence is allowed except Temporal extent where multiple temporal extents are allowed to describe datasets and series which are temporally irregular.

Example 1

dateType: creation
date: 2008-05-12T12:34:09 (date and time provided)

Example 2

dateType: publication
date:1952-06-00 (month and year provided, but no day)

Example 3

dateType: temporalExtent
date: begin: 1980-01-01 end: 1990-03-01

5. Elements describing data quality

Element 17 - Lineage (C)

Mandatory element for datasets or series of datasets. This Element is not required if a service[6] is being described. One occurrence allowed. Free text.

The purpose of this element is to record 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. The lineage element should include information about: source material; data collection methods used; data processing methods used; quality control processes. Please indicate any data collection standards used. Apart from describing the process history, the overall quality of the dataset or series should be included in the Lineage metadata element. This statement should contain any quality information required for interoperability and/or valuable for use and evaluation of the dataset or series. Acronyms should be expanded to their full text the first time they are mentioned in the Lineage element. The abbreviated version of the term can be used from then onwards.

Although not required for describing a service, MEDIN recommend that this element is populated with information about the creation of the service and the data used to generate the service.

Element 19. Additional information should be used to record relevant references to the data e.g. reports, articles, website.

Example 1

This dataset was collected by the Fisheries Research Services and provided to the British Oceanographic Data Centre for long term archive and management.

Example 2 (no protocols or standards used)

Forty 0.1m^2 Hamon grab samples were collected from across the region, both within and beyond the extraction area, and analyzed for macrofauna and sediment particle size distribution in order to produce a regional description of the status of the seabed environment. Samples were sieved over a 1mm mesh sieve. In addition, the data were analyzed in relation to the area of seabed impacted by dredging over the period 1993-1998. Areas subject to 'direct' impacts were determined through reference to annual electronic records of dredging activity and this information was then used to model the likely extent of areas potentially subject to 'indirect' ecological and geophysical impact.

Example 3: (collected using protocols and guidelines)

Data was collected using the National Marine Monitoring Programme (NMMP) data collection, processing and Quality Assurance Standard Operating Procedures (SOPs) and complies with MEDIN data guidelines.

Example 4

Survey data from Marine Nature Conservation Review (MNCR) lagoon surveys were used

[6] See Element 4 Resource type for definition of a 'service'

to create a GIS layer of the extent of saline lagoons in the UK that was ground-truthed using 2006-2008 aerial coastal photography obtained from the Environment Agency and site visits to selected locations.

Example 5

The Nutrients General Quality Assessment (GQA) described quality in terms of two nutrients: nitrates (mg NO3 /l) and phosphates (mg P/l) and graded from 1 to 6. Grades were allocated for both phosphate and nitrate; they were not combined into a single nutrients grade. There were no set 'good' or 'bad' concentrations for nutrients in the way that we describe chemical and biological quality. Locations in different parts of the country have naturally different concentrations of nutrients. ‘Very low’ nutrient concentrations, for example, are not necessarily good or bad; the classifications merely stated that concentrations in this location were very low relative to other sampling areas. Classification for phosphate Grade limit (mgP/l) Average Description: 0.02 to 0.06 Low >0.06 to 0.1 Moderate >0.1 to 0.2 High >0.2 to 1.0 Very high >1.0 Excessively high Classification for nitrate Grade limit (mg NO3/l) Average Description: 5 to 10 Low >10 to 20 Moderately low >20 to30 Moderate >30 to 40 High >40 Very high.

Element 18 - Spatial resolution (C)

Conditional for datasets and series where a resolution distance or scale can be specified. Conditional for services where there is a restriction on the spatial resolution of the service. Multiple occurrences allowed. Numeric (positive whole number).

Provides an indication of the spatial resolution of the data resource or the spatial limitations of the service. This element should only be populated if the distance or equivalent scale can be specified.

For services, spatial resolution cannot be encoded in the ISO 191939 XML Schema that this MEDIN Standard adheres to. Therefore, spatial resolution of services shall be recorded in the Abstract where necessary.

MEDIN requires that you shall provide the average distance (i.e. resolution) between sampling locations in metres, where this is possible. For example, if a dataset was composed of a grid of stations that have an average distance between stations of 2 km then 2000 metres should be recorded.

In the case of a multi-beam survey, the distance in metres should be the average distance between each sounding or ‘ping’ on the sea bed. For transect data such as an intertidal beach survey or a single beam echo sounder survey the resolution should be taken as the distance in metres between the transect lines.

For single samples and observational data MEDIN recommends using ‘not applicable’ which may be encoded as shown in the last example below.

Where the data being described in the resource is captured from a map, the scale of that map should be recorded. Spatial resolution should only be expressed by equivalent scale where a distance cannot be determined.

Example 1 (distance)

distance:10
units: metres

Example 2 (equivalent scale)

5,000

Element 19 - Additional information (O)

Optional element for datasets or series of datasets. This Element is not required if a service[7] is being described Single occurrence allowed. Free text.

The purpose of this element is to record relevant information that does not clearly belong in another element. This may be a reference, e.g. a URL.

Information about access to the resource should not be in this element, but should be provided in Element 5 ‘Resource Locator’.

Information about licencing or fees should be provided in Element 20 ‘Limitations on public access’.

Example 1

Malthus, T.J., Harries, D.B., Karpouzli, E., Moore, C.G., Lyndon, A.R., Mair, J.M., Foster-Smith, B.,Sotheran, I. and Foster-Smith, D. (2006). Biotope mapping of the Sound of Harris, Scotland. Scottish Natural Heritage Commissioned Report No. 212 (ROAME No. F01AC401/2).

Example 2

http://www.cefas.co.uk/publications/files/datarep42.pdf

[7] See Element 4 Resource type for definition of a 'service'

6. Elements relating to data usage

Element 20 - Limitations on public access (M)

Mandatory element. Multiple occurrences allowed. Controlled vocabulary and free text.

This element describes any restrictions imposed on accessing the resource for security and other reasons. Please provide information on any limitations to access of resource and the reasons for them. If different parts of the resource have different access constraints, generate occurrences for each.

Example 1

accessConstraints: otherRestrictions
otherConstraints: No restrictions to public access

Example 2

accessConstraints: otherRestrictions
otherConstraints: Restricted public access due to sensitive species, only available at 10km resolution.

Example 3

accessConstraints: otherRestrictions
otherConstraints: no limitations

Element 21 - Conditions applying for access and use (M)

Mandatory element. Multiple occurrences allowed. Free text.

This element provides information on any constraints on using the resource. Any known constraints such as licensing, fees, usage restrictions should be identified. If no conditions apply, then “no conditions apply” should be recorded.

If there is a formal licence title, that should be supplied along with, if available, a licence URL.

Example 1

Data is freely available for research or commercial use providing that the originators are acknowledged in any publications produced.

Example 2

Data is freely available for use in teaching and conservation but permission must be sought for use if the data will be reproduced in full or part or if used in any analyses.

Example 3

Not suitable for use in navigation.

Element 22 - Responsible party (M)

Mandatory element. This shall include a minimum of organisation name and email address. Multiple occurrences are allowed for some responsible party roles. Free text and controlled vocabulary.

Provides a description of an organisation or person who has a role for the resource. MEDIN mandates that the roles of 'Originator', 'Custodian' (data holder), 'Distributor', 'Metadata point of contact' and 'Owner' shall be entered. Other types of responsible party may be specified from the controlled vocabulary (INSPIRE registry[8] or ISO Codelist CI_RoleCode[9]) if desired.

If the data has been lodged with a MEDIN approved Data Archive Centre (DAC) then the DAC shall be specified as the Custodian.

Examples

Distributor:

JobPosition: DASSH Data officer
OrganisationName DASSH
PostalAddress: The Laboratory, Citadel Hill, Plymouth PL4 8SR
TelephoneNumber: 01752 633291
EmailAddress: dassh.enquiries@mba.ac.uk
WebAddress: http://www.dassh.ac.uk
ResponsiblePartyRole: distributor

Data point of contact:

JobPosition: Marine officer
OrganisationName Joint Nature Conservation Committee (JNCC)
PostalAddress:City Road, Peterborough, PE1 1JY,
TelephoneNumber: 01733 562626
FacsimileNumber: 01733 555948
EmailAddress: marine.teamexample@jncc.gov.uk
WebAddress: http://jncc.defra.gov.uk
ResponsiblePartyRole: pointOfContact

Originator:

OrganisationName: SeaZone Solutions
EmailAddress: info@seazone.com
ResponsiblePartyRole: Originator

Metadata point of contact:

JobPosition: BODC Enquiries Officer
EmailAddress: enquiries@bodc.ac.uk
TelephoneNumber: 01517954912
ResponsiblePartyRole: pointOfContact

[8] http://inspire.ec.europa.eu/metadata-codelist/ResponsiblePartyRole
[9] http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml

Owner:

JobPosition: Operations Director
OrganisationName: Oceanwise Ltd
EmailAddress: info@oceanwise.eu
TelephoneNumber: 01420768262
ResponsiblePartyRole: owner

Element 23 - Data format (C)

Mandatory for datasets and series, not applicable to services[10]. Multiple occurrences are allowed. Controlled vocabulary.

Indicate the formats in which digital data can be provided for transfer. MEDIN have defined a controlled vocabulary which is M01 'MEDIN data format categories' and is available at https://www.bodc.ac.uk/resources/vocabularies/vocabulary_search/M01/. One or more terms from this controlled vocabulary shall be used for the sub element 'name of format'. Sub element 'version' shall be populated with information about the version of the resource format(s) if known, and 'unknown' if no information is available.

Example 1

name:Database
version:Unknown

Example 2

name: Network Common Data Form
version: CF 1.6

Element 33 – Character encoding (C)

Conditional for datasets and series of datasets, not applicable to services[11]. Multiple occurrences are allowed. Controlled vocabulary.

This describes the character encoding used in the dataset. It shall be populated if an encoding is used that is not based on UTF-8, otherwise it is optional.

Select all applicable character encodings from ISO character set codelist (MD_CharacterSetCode). The full code list is presented in library G09 on the NVS2 Vocabulary Server
https://www.bodc.ac.uk/data/codes_and_formats/vocabulary_search/G09/

Example 1

8859part1

Example 2

utf8

[10] See Element 4 resource type for definition of a 'service'
[11] See Element 4 resource type for definition of a 'service'

Element 24 - Frequency of update (C)

Mandatory for datasets and series of datasets, Conditional for services where frequency of update is relevant to the service. One occurrence allowed. Controlled vocabulary.

This describes the frequency that the resource (dataset) is modified or updated and shall be included if known. For example if the dataset is from a monitoring programme which samples once per year then the frequency is annually. Select one option from ISO frequency of update codelist (MD_MaintenanceFrequencyCode codelist). The full code list can be found in library G17 on the NVS2 Vocabulary Server https://www.bodc.ac.uk/data/codes_and_formats/vocabulary_search/G17/

Example 1

monthly

Example 2

annually

7. Elements relating to Conformity

Element 25 – Conformity (M)

Mandatory element. Multiple occurrences allowed. Free text, controlled vocabulary and date.

This element specifies if the dataset or service being described is conformant with other specifications such as the INSPIRE data specifications or MEDIN data guidelines. There are 3 sub-elements which give the title of the specification, the degree of conformity (if it is or not conformant) and an explanation which gives further details of how conformant it is or any other useful information for the user. Conformity can be assessed with respect to more than one specification.

Example 1

D2.8.I.5 INSPIRE Data Specification on *Addresses* – Guidelines, publication, 2010-04-26
True
Only mandatory items included

Example 2

MEDIN Data Guideline for sediment sampling by grab or core for benthos, publication, 2009-07-29
True
All mandatory and conditional items were completed

8. Elements relating to metadata

Element 26 - Metadata date (M)

Mandatory element. One occurence allowed. Date format.

This describes the last date the metadata was updated on. If the metadata has not been updated it shall give the date on which it was created. This shall be provided as a date or date and time in the format:

yyyy-mm-dd or yyyy-mm-ddThh:mm:ss

Example 1

2008-05-12

Example 2

2008-05-12T09:09:09

Element 27 - Metadata standard name (M)

Mandatory element. One occurence allowed. Free text.

This element is to identify the metadata standard used to create the metadata. For MEDIN discovery metadata profiles, it shall be populated with the text ‘MEDIN Discovery Metadata Standard’.

Example

MEDIN Discovery Metadata Standard

Element 28 - Metadata standard version (M)

Mandatory element. One occurence allowed. Free text

This element shall be populated with the version of the MEDIN Discovery Metadata Standard used to create the metadata record for the resource.

Example

3.0

Element 29 - Metadata language (M)

Mandatory element. One occurrence allowed. Controlled vocabulary.

Describes the language used in documenting the metadata.

This element should be used to indicate the main language used in populating the metadata for the resource. If a second language is used in some elements e.g. Alternative

title, the main language should still be used to populate this element.

Select the relevant 3-letter code(s) from the ISO 639-2 code list of languages. Additional languages may be added to this list if required. A full list of recognized languages is available online at http://www.loc.gov/standards/iso639-2/php/code_list.php.

For Welsh, ISO 639-2 allows either of 'cym' or 'wel', but MEDIN recommend that 'cym' is used as this is the abbreviation of the language's own name for itself. This follows guidance from GEMINI.

Example 1 (English)

eng

Example 2 (Welsh)

cym

Element 30 – Parent ID (O)

Optional element. One occurence allowed. Free text.

This field holds the file identifier code of the series metadata record for which the dataset which is being described is part of. Therefore, this element allows links to be made between a dataset and a series (see http://www.medin.org.uk/data/faqs for MEDINs definition of these terms). This will then allow the MEDIN portal to be able to find related metadata records. For example, a large multidisciplinary project may be described as a 'series' and each of the themes of work will be described as 'datasets'. Using this field allows the user when viewing the series metadata to ask for the metadata records of all the datasets of each theme. Alternatively, a user may ask for all related records when viewing a dataset.

For services, this element should only be populated if the service that the metadata record is populated for consists of part of a larger set of services.

Example

79557726-b60a-4cf3-a8fd-9799c603d4dc

File Identifier (M)

The file identifier is a code that is encoded in XML that is globally unique and remains with the same metadata record even if the record is edited or transferred between portals or tools. It is not therefore an actual element but part of the xml record. The file identifier can be used to identify and remove duplication of records in a portal if it is harvesting records from a wide range of sources. As such it is not an element of the metadata but is used to uniquely identify the metadata xml record (as opposed to the element Unique Resource Identifier which refers to the dataset, series or service itself).

The file identifier should be created either by the organisation generating metadata or by the tools from which the metadata record is generated. Applications that are used subsequently to edit the metadata shall not change the file identifier. MEDIN recommends

the use of a 'Globally Unique Identifier' or GUID as the file identifier. It is a system generated 128-bit integer number used to identify resources (e.g. 79557726-b60a-4cf3-a8fd-9799c603d4dc). GUIDs can be generated from a variety of sources including internal PC systems and online resources such as http://www.guidgenerator.com/online-guid-generator.aspx