



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).
Document Owner	Marine Environmental Data Information Network (MEDIN) Standards Working Group
Reviewed by	MEDIN Data Standards Group
Date reviewed	
Version	2.3.5
Date approved and published on MEDIN website	21 July 2011
Date last checked for accuracy	21 July 2011
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.oceannet.org) rather than storing a local copy. A log of changes will be available on the website.

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, annexes, or sub element details. 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 data set, service 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 dataset so that in the future other people can easily discover datasets 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 standard, and includes all core INSPIRE metadata elements. It also complies with the UK GEMINI 2.1 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 2.1 specification <http://www.gigateway.org.uk/metadata/standards.html> for additional information.

In writing this document reference has been made to the technical guidelines for metadata produced by INSPIRE (see guidelines at <http://inspire.jrc.ec.europa.eu/reports.cfm>)¹.

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

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 Source' 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 data set 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 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 data set usually constitutes a specifically-funded piece of work,
- the dataset should be easily extractable from a database for a 3rd party,

¹ 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.

- if you are searching for a data set 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.

Element 1 - Resource title (M)

Mandatory element. Only one resource example 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. The following format is recommended:

'Date' 'Originating organization/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).

Examples

Example 1: 1992 Centre for Environment, Fisheries and Aquaculture Science (Cefas) North Sea 2m beam trawl survey.

Element 2 - Alternative resource title (O)

Optional element. Multiple alternative resource titles allowed. Free text.

The alternative title is used to add the names by which the resource (e.g. dataset) may be known and may include short name, other name, acronym or alternative language title.

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). 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 10). It is recommended that acronyms and abbreviations are reproduced in full. e.g. Centre for Environment, Fisheries and Aquaculture Science (Cefas).

Examples

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 e.g. a dataset using the controlled vocabulary, MD_ScopeCode from ISO 19115. The resource type must 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 is available at

http://www.oceannet.org/marine_data_standards/other_marine_data_standards/consider_data_set.html

Example

series

Element 5 - Resource locator (C)

Conditional element (must be completed if known). Multiple resource locators are 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.

Example

Resource locator url:

<http://www.defra.gov.uk/marine/science/monitoring/merman.htm>

Resource locator name: The Marine Environment National Monitoring and Assessment Database

Resource locator function: download

Element 6 - Unique resource identifier (M)

Mandatory element (for datasets and series of datasets). One occurrence allowed. Free text.

A Unique Resource Identifier allows a dataset to be identified by a code. This code is generally assigned by the data owner and commonly consists of the organisation which 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 enquiries@oceannet.org and we will generate a code for you or endeavour to provide a tool to generate your own codes.

Example 1.

Code: 5639287
Codespace: <http://www.bodc.ac.uk>
Example 2:
Code: <http://www.bodc.ac.uk/5639287>

Element 7 - Coupled resource (C)

Conditional element. Mandatory if linkages to the datasets on which the service operates on are available. Multiple coupled resource occurrences allowed.

An INSPIRE element referring to data services such as a data download or mapping web services. It identifies the data resource(s) used by the service if these are available separately from the service. You should supply the Unique resource identifiers of the relevant datasets (See element 6).

Example

MRMLN0000345

Element 8 - Resource language (C)

Conditional element. Mandatory when the described resource contains textual information. Multiple resource languages allowed. This element is not required if a service¹ is being described rather than a dataset or series of datasets. Controlled vocabulary, ISO 639-2.

¹ See Element 4 resource type for definition of a 'service'

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 <http://www.loc.gov/standards/iso639-2>.

Examples

Example 1: eng (English)

Example 2: cym (Welsh)

Element 9 - Topic category (C)

Conditional element. Mandatory for datasets and series of datasets. Multiple topic categories are allowed. This element is not required if a service¹ is being described. Controlled vocabulary.

This element is mandatory for INSPIRE and must be included. This indicates the main theme(s) of the data resource. It is required for INSPIRE compliance. The relevant topic category should be selected from the ISO MD_TopicCategory list. Within MEDIN the parameter group keywords from the controlled vocabulary P021 available at <http://vocab.ndg.nerc.ac.uk/client/vocabServer.jsp> (included in element 11) are mapped to the ISO Topic Categories so it is possible to generate the topic categories automatically once

the keywords from BODC Parameter Discovery Vocabulary (P021) have been selected.

¹ See Element 4 resource type for definition of a 'service'

Examples

Example 1: biota

Example 2: oceans

Element 10- Spatial data service type (C)

Conditional element. Mandatory if the described resource is a service¹. One occurrence allowed.

An element required by INSPIRE for metadata about data services e.g. web services¹. If a service is being described (from Element 4) it must be assigned a service type from the INSPIRE Service type code list.

¹ See Element 4 resource type for definition of a 'service'

Example

Download

Element 11 - Keywords (M)

Mandatory element. Multiple keywords allowed. Controlled vocabularies.

The entry should consist of two sub-elements: the keywords and reference to the controlled vocabulary used as shown in the sub elements below. To allow searching of the dataset, keywords should be chosen from 3 code lists given below and the OAI harvesting keyword. In addition if a 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 should be given.

INSPIRE keywords

A list of the INSPIRE theme keywords is available at http://www.eionet.europa.eu/gemet/inspire_themes At least one INSPIRE theme keywords is required for INSPIRE compliance.

MEDIN Keywords

MEDIN strongly recommends the use of the BODC Parameter Discovery Vocabulary (P021) to provide further ability to search by terms that are more related to the marine domain. This list is available at <http://vocab.ndg.nerc.ac.uk/client/vocabServer.jsp> In particular the parameter groups and codes that are used may be searched through a more user friendly interface which has been built as part of the European funded SeaDataNet project at http://seadatanet.maris2.nl/v_bodc_vocab/vocabrelations.aspx

Vertical Extent Keywords

A vocabulary of keywords is available to describe the vertical extent of the resource (e.g. data set). The vocabulary can be downloaded as L131 (Vertical Coordinate Coverages) at <http://vocab.ndg.nerc.ac.uk/client/vocabServer.jsp>. These lists are also available through a more user friendly interface at

http://seadatanet.maris2.nl/v_bodc_vocab/welcome.aspx/http://vocab.ndg.nerc.ac.uk/client/vocabServer.jsp

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

OAI harvesting

If xml files are being collected using the MEDIN OAI harvesting process an additional keyword is required to allow the data discovery service to distinguish MEDIN records from other records such as NERC. The required term to use in the xml fragment is NDGO0001 from the N010 vocab at <http://vocab.ndg.nerc.ac.uk/client/vocabServer.jsp>.

Other Keywords

Other vocabularies may be used as required as long as they follow the format specified in 11.1 – 11.2.3

Keywords for services

Define the category or subcategory of the service using its language neutral name as defined in Part D 4 of the Metadata Implementing Rules. The Metadata Implementing Rules can be found at

<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:326:0012:01:EN:HTML> and the keyword vocabulary available at <http://inspire-registry.jrc.ec.europa.eu/registers/GLOSSARY/items/184.jsessionid=1C0A17A72079DC671E4AB9A07D7C66AB>

Examples

keywordValue: Fish taxonomy-related counts

keywordValue: Temperature of the water column

thesaurusName: BODC Parameter Discovery Vocabulary

dateType: revision

date: 2009-10-13

keywordValue: upper_epipelagic

thesaurusName: SeaDataNet vertical coverage

dateType: Creation

date: 2006-11-15

Element 12 - Geographic bounding box (C)

Mandatory element for datasets and conditional for services. One occurrence of each sub-element allowed. Numeric and controlled vocabulary.

These 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. A minimum of two and a maximum of four decimal places should be provided.

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.

Example

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

Element 13 - Extent (O)

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

Keywords selected from controlled vocabularies to describe the spatial extent of the resource MEDIN strongly recommends the use of the SeaVox Sea Areas salt and freshwater body gazetter available as vocabulary C191 at <http://vocab.ndg.nerc.ac.uk/client/vocabServer.jsp> which is a managed vocabulary and has a worldwide distribution.

Other vocabularies available including ICES areas and rectangles www.ices.dk , or Charting Progress 2 regions may be used as long as they follow the format specified in 13.1 – 13.2.3.

Example

This example includes multiple extents from different vocabularies.

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

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

Element 14 - Vertical extent information (O)

Optional element. This element should only be filled in if the vertical Coordinate Reference System is known. One occurrence allowed. Numeric free text and controlled vocabulary.

This element should only be filled in if the Coordinate Reference System (CRS) is registered in the 'European Petroleum Survey Group (EPSG) database. <http://info.ogp.org.uk/geodesy/> If you do not have the defined CRS you should 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 if it is filled in then its sub-elements are either mandatory or conditional.

Example

minimumValue: 42

maximumValue: 94

verticalCoordinateReferenceSystem: <urn:ogc:def:crs:EPSG::5701>

Element 15 - Spatial reference system (M)

Mandatory for datasets and series, conditional where relevant to services. One occurrence allowed. Controlled vocabulary.

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 (<http://www.epsg.org/Geodetic.html>). . To find a code click on the OGP Online Registry and if you know the title (eg WGS84) then type this in the 'Name' field and click search. The name, code and further information is 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 lat and longs then click search. The website also provides a database of the reference systems and web services to access the information.

Examples

Example 1: WGS84 – <urn:ogc:def:crs:EPSG::4326>

Example 2: National Grid of Great Britain – <urn:ogc:def:crs:EPSG::27700>

Element 16 - Temporal reference (M)

Mandatory element for data sets and series; optional for services. One occurrence allowed of each sub element. Date/Time format.

It is recommended that all known temporal references of the resource are included. 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, the date of last revision and creation date are conditional (i.e. must be completed if known) Only one occurrence for each sub-element is allowed.

Examples

Example 1:

dateType: creation

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

Example 2:

dateType: revision

date:2008-05-12 (full date provided)

Example 3:
dateType: publication
date:1952-06-00 (month and year provided, but no day)

Element 17 - Lineage (C)

Mandatory element for datasets or series of datasets. One occurrence allowed. This Element is not required if a service¹ is being described. Free text.

Lineage includes the background information, history of the sources of data used and can include data quality statements. The lineage element can include information about: source material; data collection methods used; data processing methods used; quality control processes. Please indicate any data collection standards used. Additional information source to record relevant references to the data e.g. reports, articles, website. 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 data set or series.

¹ See Element 4 Resource type for definition of a 'service'

Examples

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 0.1m² 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.

Element 18 - Spatial resolution (C)

Conditional for datasets and series where a resolution distance can be specified. Multiple occurrences allowed. Numeric (positive whole number) and free text.

Provides an indication of the spatial resolution of the data. The element has largely been derived for the mapping community and is currently poorly defined however MEDIN recommends that you provide the average distance (i.e. resolution) between sampling locations in metres. For example, if a dataset was composed of a grid of stations which have an average distance between stations of 2 km then 2000 metres should be recorded. In the case of a multibeam survey it 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 between the transect lines.

For single samples and observational data MEDIN recommends using 'not applicable' which may be en-coded as shown in the last example below.

MEDIN is in discussions with GEMINI and ISO to allow the use of scale for this element (e.g. pressure) and also to allow the use of 'unknown'. GEMINI accepts that in many cases only approximate values can be given.

Examples

Example 1:
distance:10
units: metres

Element 19 - Additional information source (O)

Optional element. Single occurrence allowed. Free text.

Any references to external information that are considered useful, e.g. project website, report, journal article may be recorded. It should not be used to record additional information about the resource.

Examples

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).

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

Element 20 - Limitations on public access (M)

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

This element describes any restrictions imposed on the resource for security and other reasons using the controlled ISO vocabulary RestrictionCode. If restricted or otherRestrictions is chosen please provide information on any limitations to access of resource and the reasons for them. If there are no limitations on public access, this must be indicated.

Examples

Example 1:
accessConstraints:
otherRestrictions: No restrictions to public access

Example 2:
accessConstraints:
otherRestrictions: Restricted public access due to sensitive species, only available at 10km resolution.

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

Mandatory element. Multiple occurrences allowed. Free text.

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.

Examples

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. Multiple occurrences are allowed for some responsible party roles. Must include minimum of person/organization name and email address. Free text and controlled vocabulary.

Provides a description of an organization or person who has a role for the dataset or resource. MEDIN mandates that the roles of ‘Originator’ and ‘Custodian’ (data holder) and the role of ‘Distributor’ should be entered if different to the Custodian. The ‘Metadata point of contact’ is also mandatory. Other types of responsible party may be specified from the controlled vocabulary if desired.

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

Examples

Data point of contact:

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

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

Originator:

IndividualName: Dr A. Smith,
OrganizationName: University of Swansea
ResponsiblePartyRole: Originator

Metadata point of contact:

IndividualName: Miss Hannah Freeman
EmailAddress: haee@bodc.ac.uk
TelephoneNumber: 01517954898

Element 23 - Data format (O)

Optional element. Multiple data formats are allowed. Controlled vocabulary.

Indicate the formats in which digital data can be provided for transfer. A controlled vocabulary has been defined for use by MEDIN which is M010 'MEDIN data format categories' available at <http://vocab.ndg.nerc.ac.uk/client/vocabServer.jsp>. The term from this controlled vocab should be used for the sub element 'name of format' and 'unknown' used for the sub element version

Example 1

Database

Unknown

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 (data set) is modified or updated and should be included if known. For example if the data set 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_FrequencyOfUpdate codelist).

Examples

Example 1: monthly

Example 2: annually

Element 25 - Conformity

This element relates specifies if the data set 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.

Conditional element. Multiple occurrences allowed. Required if the resource provider is claiming conformance to INSPIRE.

Element 26 - Metadata date (M)

Mandatory element. One occurrence allowed. Date format.

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. This should be provided as a date in the format:

yyyy-mm-dd

Example

2008-05-12

Element 27 - Metadata standard name (M)

Mandatory element. One occurrence allowed. Free text.

Identify the metadata standard used to create the metadata. It is recommended that the term below is used to comply with this MEDIN standard.

Example

MEDIN Discovery Metadata Standard

Element 28 - Metadata standard version (M)

Mandatory element. One occurrence allowed.

Identify the version of the metadata standard used to create the metadata. It is recommended that the term below is used to comply with this MEDIN standard.

Example

2.3.2

Element 29 - Metadata language (M)

Mandatory element. One occurrence allowed. Controlled vocabulary.

Describes the language(s) elements of the metadata.

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 UK language codes is available online <http://www.loc.gov/standards/iso639-2>.

Examples

Example 1: (English)

eng

Example 2: (Welsh)

cym

Element 30 – Parent ID (O)

Optional element. One occurrence 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.oceannet.org/marine_data_standards/other_marine_data_standards/consider_data_set.html 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.

Example

79557726-b60a-4cf3-a8fd-9799c603d4dc

File Identifier

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>