



www.oceannet.org

Title	MEDIN data guideline for recording and archiving of digital photographs
MEDIN Discipline	Cross Cutting
Author(s)	A. Leadbetter, M. Charlesworth,
Document Owner	A. Leadbetter
Reviewed by	MEDIN Standards Group
Date reviewed	10-08-2009
Version	3.1
Date approved and published on MEDIN website	15-09-2009
Date last checked for accuracy	15-07-2010
Summary	This guideline defines good practice for archiving of digital photographs.
Keywords	Digital, Photographs

Change history		
Version	Date	Change
1.0	17/09/09	First draft of document
2.0	30/11/09	Revised and published following comments from DASSH, CCW and BGS
2.1.	16/12/09	Small amendment to allow original coordinates in northing/easting rather than lat long following BGS advice.
3.0	06/05/10	Small additions following feedback and put into new structure
3.1	15/07/10	Minor edits following changes to common tables and specification of .csv format for transfer of data

1.1. Background

The Marine Environmental Data and Information Network (MEDIN) is working towards creating a framework of consistent standards covering the major types of data collection undertaken in the marine environment around the UK. The principle benefits of this suite of standards are:

- Allows contracting organisation to easily specify a format that data should be returned in that can be readily used and includes all relevant attributes
- Provides a consistent format for contractors to work to (rather than a different format for each contract)
- Data can be readily exported to Data Archiving Centres and other users
- Instills good practice amongst users

Each standard defines the data and information that must be stored with a particular data type to ensure it can be readily used and reused. As this type of information is specific for different data types, guidelines are developed for each type. This document describes one such format. Other standards can be accessed through www.oceannet.org.

1.2. Scope

This guideline covers the archiving of digital photographs regardless of the reason for which they have been taken, or from what platform they have been taken (e.g. ship, beach, underwater vehicle).

1.3. Using this data guideline

This guideline is split into sections that refers to information that can be collated at different levels. The metadata to be held with an image can be broadly split into that which will be applicable to a group of images taken using common techniques for example within one project or survey (Project or Survey Information), spatial and temporal information relevant for each image (Sample Event Information) and quality metadata applicable for each image (Image Properties). Tables that are common across MEDIN data guidelines can be downloaded from www.oceannet.org/standards.

The tables below outline the data fields, a description and where available a term list and/or format given at the end of each field which should be used to store the data. Each field is either mandatory, conditional or optional as indicated by M, C or O respectively.

To submit this data to a Data Archive Centre the raw image files should be provided as .tiff files, and the remaining data must be saved and transferred in the .csv file format.

1.4. Further information on the SeaDataNet, ICES and EPSG term lists

The available catalogues of term lists used for this MEDIN data guideline are provided primarily by SeaDataNet, the International Council for the Sea (ICES) and EPSG. If a term is not available in a recommended list then please contact MEDIN to arrange for the term to be added.

The SeaDataNet list may be viewed at http://seadatanet.maris2.nl/v_bodc_vocab/welcome.aspx . By clicking on the list any term may be searched for by using the drop-down menus or all terms viewed by clicking search. The terms may be viewed in groups of 15 or may be downloaded into an excel file.

The ICES term lists are available at <http://www.ices.dk/datacentre/reco/> Select which list you require from the 'Reference Code List' drop-down box. The results are shown for the selected list and may be downloaded into MS Excel by selecting the inverted green arrow.

There are a number of ways of describing a spatial dataset. Common horizontal coordinate reference systems include WGS84 and British National Grid. Common vertical coordinate reference systems include Highest Astronomical Tide and Ordnance Datum Newlyn (ODN). It is important that which coordinate reference system used for a data set is recorded so conversions can be carried out between reference systems. The EPSG database of coordinate reference systems (<http://www.epsg.org/Geodetic.html>) provides a dictionary of reference systems. In brief, 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.

1.5. Relationship between MEDIN data guidelines and MEDIN discovery metadata

The MEDIN discovery metadata format is aimed at allowing the non-informed user to discover data sets and it is likely that one 'discovery' data set record will contain a large range of data types that are in turn covered by a range of data guidelines. To enable individuals to reuse data of a specific nature (e.g. benthic invertebrate data) then related information must be collected (e.g. data owner,

reference systems used etc). Some of the information which is collected at the Survey Level in a data guideline is also required to create a discovery metadata record. Who creates the MEDIN discovery record for a dataset is case specific and dependant on the organisation, and the relationship it has with a Data Archive Centre. However it is intended that the information collected at the 'Survey Information' level is reused for creating a MEDIN discovery metadata record.

2.1. Project Information.

If your collection of data forms part of a wider project or time series then the below details must be recorded.

M, C, O indicate which fields are mandatory, conditional or optional.

Heading	M, C, O	Description	Recommended Term List or Format
Project name	M	The nationally/internationally accepted version of the project name	Free text; (e.g. Rapid Climate Change)
Project website	C	If a Project website exists give the address	e.g. (http://www.noc.soton.ac.uk/rapid/rapid.php)
Project start date	M	The date that the project started	Date; yyyy-mm-dd; (e.g. 2001-01-24)
Project end date	C	The date that the project is due to finish	Date; yyyy-mm-dd; (e.g. 2007-01-24)
Project code	M	Provide a code to uniquely identify the project and allow links to be made between the tables.	Free text; (e.g. RCC)

2.2. Survey Information (Data Activity).

The survey information is a uniquely identifiable programme of data collection such as a research cruise or survey event. This information is likely to be the same for all sample events (e.g. stations) and subsamples in a given data set such as a cruise. Note that in the event that these are not common to all sample events then they should be specified for each one. These fields are common throughout many other MEDIN data guidelines and only need to be given once and referenced if your data set is composed of many data types and therefore conforms to a number of MEDIN Data Guidelines. Where data collection is undertaken on research vessels the data below can often be sourced in the Cruise Summary Report.

M, C, O indicate which fields are mandatory, optional or conditional respectively.

Heading	M, C, O	Description	Recommended Term List or Format
Survey name	M	Title of the survey	Free text; (e.g. Menai Straight Benthic Survey 2004)
Survey description	M	Brief description of the purpose of the survey and other types of measurements that were made for the survey.	Free Text
Survey code	M	A unique code for the survey to allow links to be built between this and sample event data (the cruise identifier code could be used)	Free text; (e.g. JCR3022)
Responsible organisation	M	Organisation who has funded the work	Term List; European Directory of Marine Organisations (e.g. 28: Centre for Environment, Fisheries and Aquaculture Science, Lowestoft Laboratory)
Survey start date	M	The date and time that the survey started.	Date or DateTime; yyyy-mm-dd or yyyy-mm-dd hh:mm:ss (e.g. 2009-01-24 12:33:00)
Survey end date	M	The date and time that the survey ended.	Date or DateTime; yyyy-mm-dd or yyyy-mm-dd hh:mm:ss (e.g. 2009-02-16 16:33:00)
Spatial coordinate reference system	M	Describes the system of spatial referencing. I.e. the datum used to provide details of latitude and longitude.	Term List; http://www.epsg.org/Geodetic.html (e.g. WGS84 is EPSG::7030)
Position fix	M	Give the method and source of the position	Free Text; (e.g. Differential GPS taken from

method and source		fix instrument.	the ships navigation equipment.
Horizontal positional accuracy	M	How accurate the spatial positions are likely to be	Number; units = meters (e.g. 15)
Depth coordinate reference system	C	Give the reference to which the depth has been calculated e.g. Highest astronomical tide. Mandatory if seabed depths are given for each sample.	Term List http://www.epsg.org/Geodetic.html (e.g. ODN is EPSG::5701)
Vertical positional accuracy	C	How accurate the vertical resolution is. Must be provided if seabed depths are given.	Number; units = meters (e.g. 0.5)
Platform type	O	The platform type (e.g. Research Vessel) from which the sampling device was deployed.	Term list SeadataNet Platform Classes (L061) (e.g. 31)
Ship name	O	The name of the ship from which the sampling device was deployed.	Term list SHIPC at http://www.ices.dk/datacentre/reco/ (e.g. 74LG Lough Foyle)
Cruise report reference	O	Cruise report reference if applicable.	Free text; in reference format. e.g. Litt, E.J. 2009. PHiXT 4. 30 July to 2 August 2009 <i>RV Prince Madog</i> POL Coastal Observatory Liverpool Bay Cruise Report. POL Coastal Observatory, Liverpool.
Project code	M	State the code of the project given in the project table to allow links to be made between the tables.	Free text; (e.g. RCC)

2.3. Fixed Station Information.

In many cases a fixed point, transect or area is returned to on a number of occasions to form a time series. The actual sample event may not be in exactly the same location each time however due to ship movements or sampling strategy, however it is useful to record both the position which is intended to be sampled (fixed) and the actual sampling position (sample).

Therefore, the information below must be included if a fixed point, transect or area is used as the basis for replicate profiles or for repeat monitoring surveys. Actual profile coordinates should be placed in the sample event table. A fixed station may be a point, transect, or an area. If the fixed station is a transect or an area then the secondary latitude and longitude fields must be completed.

Heading	M, C, O	Description	Recommended Term List or Format
Local station identifier	M	A unique identifier for the station	Free text. e.g. Stanton Bank site 4 (point) e.g. Liverpool/Dublin transect (transect) e.g. Lagan Estuary (area)
Primary Latitude (decimal degrees)	M	The primary latitude of the fixed station given in decimal degrees. For a point this field is set to the point latitude; for a transect it is set to the latitude of the start of the transect; for an area it is set to the southern edge of the box. Units are positive north.	Decimal degrees; minimum of two and a maximum of five decimal places. e.g. 54.5837
Primary Longitude (decimal degrees)	M	The primary longitude of the sample given in decimal degrees. For a point this field is set to the point longitude; for a transect it is set to the longitude of the start of the transect; for an area it is set to the western edge of the box. Units are positive east (West is negative, East is positive).	Decimal degrees; minimum of two and a maximum of five decimal places. e.g. -5.5837
Secondary Latitude (decimal degrees)	C	The secondary latitude of the fixed station given in decimal degrees. For a point this field is not required; for a transect it is set to the latitude of the end of the transect; for an area it is set to the northern edge of the box. Units are positive north.	Decimal degrees; minimum of two and a maximum of five decimal places. e.g. 55.7393

Secondary Longitude (decimal degrees)	C	The secondary longitude of the sample given in decimal degrees. For a point this field is not required; for a transect it is set to the longitude of the end of the transect; for an area it is set to the eastern edge of the box. Units are positive east (West is negative, East is positive).	Decimal degrees; minimum of two and a maximum of five decimal places. e.g. -3.7394
Description of fixed station spatial form	M	Describe if the fixed station is a point, transect or an area.	Term list; <u>SeadataNet Geospatial Feature Type (L021)</u> (e.g. point)

2.4. Image Context Information (Sample Event). This table holds information on the location and time of image capture.

Heading	M, C, O	Description	Recommended Term List or Format
Image identifier	M	A unique identifier for the image under consideration. The Exif file will give a unique identifier to that camera but possibly not unique for other cameras. Therefore, a prefix should be used to ensure that it is globally unique. MEDIN recommend that a combination of the organisation name and camera ID is used.	e.g. BODC_nicon590c_00058743 e.g. CCW_canon45i_PS74926
Sample Reference	C	If the image has been used as a photographic record of for example a biological or geological sample then an identifier for that sample should be given here to allow the image and the sample to be related.	Free Text e.g. CCW_Llyn_may10_Stn46a
X coordinate position of sample given in original recorded format	M	The X coordinate (latitude or northing) of the image given in whichever format was used to record at the time of sampling. Units are positive north.	Minimum of 10 figures for NGR e.g. 50°47'24" e.g. 233978.057
Y coordinate position of sample given in original recorded format	M	The Y coordinate (longitude or easting) of the image given in whichever format was used to record at the time of sampling. Units are positive east.	Minimum of 10 figures for NGR e.g. -4°21'53" e.g. 331445.535
Latitude of sample (decimal degrees)	M	The latitude of the image given in decimal degrees. Units are positive north. 5 decimal places is recommended.	Decimal degrees; minimum of four and a maximum of seven decimal places. e.g. 54.5837
Longitude of sample (decimal degrees)	M	The longitude of the image given in decimal degrees. Units are positive east. 5 decimal places is recommended.	Decimal degrees; minimum of four and a maximum of seven decimal places. e.g. -3.4764
Date and time	M	The date and time of image collection.	yyyy-mm-dd or yyyy-mm-dd hh:mm:ss (e.g. 2009-01-24

			13:33:00)
Depth	M	The depth or height that the image was taken	Number; units = meters (e.g. -15). Below water surface is negative.
Bearing	O	The approximate compass direction in which the image has been taken.	Free text. e.g. NorthEast
Data originator	O	The person who took or provides the images	Free text. e.g. Joe Bloggs

2.5. Image Collection Methods and Properties (Data Production Tools). The image metadata fields are intended to hold information regarding the device used to record the image and the properties of the image itself. Many digital cameras record these metadata automatically at the time of image capture in a metadata format known as the Exchangeable image file (Exif) standard. If you wish to view these metadata right click on an image in windows explorer, choose the summary tab and click advanced. There is also a FireFox plug in that allows you to view these metadata easily. The Exif standard also incorporates fields for a Global Positioning System feed to provide a geo-reference to the image. The TIFF file format holds metadata in the Exif standard. The fields below give the metadata that should be stored with every image and the mandatory or optional conditions using the Exif standard. MEDIN recommend that the data below should be extracted and held in a separate table. At a minimum the information in the mandatory fields below should be present and held in the EXIF file.

Heading	M, O, C	Description	
Identifier	M	A unique code or name used to identify the image	e.g. BODC_nicon590c_00058743
Version	O	The version of the Exif specification being used	e.g. 0220 for version 2.2
Manufacturer	O	The manufacturer of the capture device, as recorded in the Exif tags	e.g. FUJIFILM
Model	O	The model of the capture device, as recorded in the Exif tags	e.g. FinePix S3Pro
Software	O	The software version used on the capture device, as recorded in the Exif tags	e.g. FinePixViewer Ver3.1
FileSource	O	The file source of the image, as recorded in the Exif tags	e.g. 3 for a digital stills camera or 'other' for any other source of digital image.
PixelXDimension	M	The width of the image in pixels	Number; units = pixels (e.g. 2000)
PixelYDimension	M	The height of the image in pixels	Number; units = pixels (e.g. 1250)
xResolution	M	The image resolution (x-dimension - width) in pixels per field ResolutionUnit	e.g. 300
yResolution	M	The image resolution (y-dimension - height) in pixels per field ResolutionUnit	e.g. 300
ResolutionUnit	M	The units of xResolution and yResolution	e.g. DPI
ExposureTime	O	The exposure time of the image in seconds	Number; units = seconds (e.g. 0.02)

Heading	M, O, C	Description	
ExposureBias	O	The exposure bias (or exposure compensation) of the image, which controls the amount of light entering the camera lens	Number; units = step (e.g. 0)
ExposureProg	O	The exposure program used when taking a digital photo.	The term list is shown in Appendix 1.
fNumber	O	The f-number (focal ratio, f-ratio or relative aperture) used in capturing the image	e.g. F5.3
FocalLength	O	The actual focal length of the lens used to capture the image, given in millimetres	Number; units = mm (e.g. 78)
MeteringMode	O	The metering mode used in the capture of the image.	The term list is shown in Appendix 2.
Flash	O	The flash mode used in the capture of the image.	The term list is shown in Appendix 3. If the field is blank it indicates that a flash was not used.
ApertureValue	O	The smallest f-number of the lens used to capture the image, in the units of the Additive System of Photographic Exposure setting. This is normally, but not limited to, the range of 00.00 to 99.99.	e.g. 23.20

2.6 Sample Data

The recommended file format for storing digital images is the TIFF (Tagged Image File Format). TIFF is universally recommended for long-term archive and curation of digital images as it does not suffer from lossy (i.e. a loss of information) compression like other formats such as JPEG which can cause a lack of true representation of the image which was originally captured. Many digital cameras record image properties automatically at the time of image capture in a metadata format known as the Exchangeable image file (Exif) standard. The TIFF file format supports metadata in the Exif standard. In instances when image files are provided in other formats (e.g. jpeg) we recommend that these files are converted to the TIFF format, however appreciate this may be impractical. Regardless of the approach a consistent format for image fields should be sought and compression of files should be strictly avoided.

Many digital stills cameras record their images in a raw image format (also known as a digital negative format or RAW) containing minimally processed data as collected by the imaging sensor. There is not, however, an industry standard digital negative format, so the actual file format may be different between different image capture device manufacturers, or even between models from the same manufacturer. Due to the proliferation of these formats, they are not suitable for the archiving of digital images for any use.

Appendix 1. Term list for image metadata field 'ExposureProg'

- 1 - Manual
- 2 - Normal program
- 3 - Aperture priority
- 4 - Shutter priority
- 5 - Creative program
- 6 - Action program
- 7 - Portrait mode
- 8 - Landscape mode

Appendix 2. Term list for image metadata field 'MeteringMode'

- 0 – Unknown
- 1 – Average
- 2 – Center weighted average
- 3 – Spot
- 4 – Multi-spot
- 5 – Pattern
- 6 – Partial
- 255 - Other

Appendix 3. Term list for image metadata field 'Flash'

- 0000 = Flash did not fire
- 0001 = Flash fired
- 0005 = Strobe return light not detected
- 0007 = Strobe return light detected
- 0009 = Flash fired, compulsory flash mode
- 000D = Flash fired, compulsory flash mode, return light not detected
- 000F = Flash fired, compulsory flash mode, return light detected
- 0010 = Flash did not fire, compulsory flash mode
- 0018 = Flash did not fire, auto mode
- 0019 = Flash fired, auto mode
- 001D = Flash fired, auto mode, return light not detected
- 001F = Flash fired, auto mode, return light detected

0020 = No flash function
0041 = Flash fired, red-eye reduction mode
0045 = Flash fired, red-eye reduction mode, return light not detected
0047 = Flash fired, red-eye reduction mode, return light detected
0049 = Flash fired, compulsory flash mode, red-eye reduction mode
004D = Flash fired, compulsory flash mode, red-eye reduction mode, return light not detected
004F = Flash fired, compulsory flash mode, red-eye reduction mode, return light detected
0059 = Flash fired, auto mode, red-eye reduction mode
005D = Flash fired, auto mode, return light not detected, red-eye reduction mode
005F = Flash fired, auto mode, return light detected, red-eye reduction mode