

Title:	The MESH Sample Data Exchange Format (DEF)
Author(s):	Natalie Coltman (JNCC)
Document owner:	Natalie Coltman (natalie.coltman@jncc.gov.uk)
Reviewed by:	Jon Davies (JNCC)
Workgroup:	n/a
MESH action:	1
Version:	9
Date published:	08/07/2005
File name:	MESH Sample DEF.pdf
Language:	English
Number of pages:	7
Summary:	The MESH Sample Data Exchange Format (DEF) describes the format required for point data files that are supplied to the MESH partnership. Specifically it relates to fields in text files representing sample point data.
Reference:	JNCC 2005 The MESH Data Exchange Format (DEF)
Keywords:	Data Format, DEF, GIS
Bookmarks:	
Related information:	MESH online metadata catalogue: www.searchmesh.net/metadata

MESH Sample Data Exchange Format

This document is derived from the National Biodiversity Network (NBN) Exchange Format, the principal way of supplying datasets for automated upload to the NBN Gateway. The format has been adapted for exchanging data within the MESH project, known as the MESH Sample DEF.

The MESH Sample DEF differs from the other DEFs because it is text-based and has been designed to be straightforward to produce from a variety of applications. The Study Area DEF, Original Habitat DEF and Translated Habitat DEF all define the attributes of ESRI shapefiles, whereas the Sample DEF does not specify an ESRI shapefile format. This decision was made because point sample data are more widespread among potential data providers than polygon map data, but not all these providers have access to GIS packages. For these reasons MESH decided to specify a Sample DEF which will be open to a wider audience who may not have access to GIS packages.

In its simplest form, the Sample DEF encapsulates the basic components of a species or biotope occurrence record (*what* was recorded, *where* it was recorded, *when* it was recorded, and *who* recorded it). However, it is extensible and can include any additional data associated with each record.

The data file comprises one record per row with values separated by tabs (ASCII character 9). The first row of the data section must contain field names, selected from the list of reserved names below, plus any additional fields you want to include. There is no need to include optional fields that do not contain any data in your dataset, and unlike the other MESH DEFs the fields can be in any sequence. Each record within the exchange format file must occupy one line only. Tab and end-of-line characters must not appear anywhere else in the file.

Below is a full list of the reserved field names with brief descriptions. The maximum length of a value in each of the fields is given in brackets where applicable. Some of these fields must be supplied with every dataset (mandatory), others are optional. In some cases, whether the field is mandatory or not will depend on the type of data supplied and on the information provided in other fields; such fields are marked as conditional. The requirements are explained in the detailed notes that follow the summary tables.

Key fields

Field name	Format (length)	Description
RecordKey	Text (8)	Unique key for each occurrence record
SampleKey	Text (8)	Sample key (unique within a survey event or survey). Examples of samples are a section of a video tow or a grab sample.
SurveyEventKey	Text (8)	Survey event key (unique within a survey). A video tow or a group of grab samples are examples of survey events.
SurveyKey	Text (8)	Unique survey key

Date fields

StartDate	DD/MM/YYYY or	Start date of occurrence record
-----------	------------------	---------------------------------

	YYYY-MM-DD	
EndDate	DD/MM/YYYY or YYYY-MM-DD	End date of occurrence record
DateType	Text (2)	Vague date type

Biological fields

TaxonName	Text (16) Set list (dictionary)	Each biological occurrence record should be a taxon occurrence OR a biotope occurrence but not both; for biological data a value must be present in this column OR in BiotopeKey.
Abundance	Text (1)	The abundance of individuals seen as measured on the SACFOR scale.
Count	Number	Number of individuals of the taxon seen per m ² .
Presence	Text (1)	Presence or absence of the taxon at this location (P or A).
Percentage	Number	The percentage cover of a taxon.
Score	Number	An abundance score of 0-5 to indicate the biomass or quantity of flora and fauna present.
Habitat	Text (255)	Code or text description of the habitat recorded at this location. Each occurrence record should be a taxon occurrence OR a biotope occurrence but not both; a value must be present in this column or in TaxonName. Where [ImageFileName] is populated, [Habitat] is a mandatory field.
Version	Text (16) Set list (dictionary)	The name and publication date of the classification system used in [Habitat].
ImageFileName	Text (255)	A biotope occurrence record can be associated with a seabed photograph taken <i>at the same location</i> by including an image file name in this field.

Physical fields

Depth	Text (50)	The depth in metres below chart datum at which the occurrence record was recorded.
SeabedDescription	Text (255)	Description of the seabed found at the location of the occurrence record.

Location fields

StartEast	Number	Position of the start sample of the sample location in an east/west direction as a longitude.
-----------	--------	---

StartNorth	Number	Position of the start sample location in a north/south direction as a latitude.
EndEast	Number	Position of the end sample location in an east/west direction as a longitude.
EndNorth	Number	Position of the end sample location in a north/south direction as a latitude.
Datum	Text (8)	The datum for long/lat coordinates (e.g. "WGS84" or "OSGB36")
Precision	Number	Spatial precision of longitude and latitude (metres)

Key fields

RecordKey

Mandatory This is the primary key associated with an occurrence record. Each RecordKey should be unique.

SampleKey

Conditional A key to group occurrence records into discrete samples within a survey. Records belonging to the same sample will also share the same location, for example they are collected from the same *section* of video tow or the same grab sample.

SurveyEventKey

Conditional A key to sort occurrence records into discrete survey events. A video tow is an example of a survey event and for occurrence records derived from video tows, this field is mandatory. For other types of occurrence records it is optional.

SurveyKey

Mandatory This key can be used to divide the dataset into separate surveys, which represent logical subsets of the data within the exchange file. You can use survey keys to divide up the dataset in any way you wish (e.g. representing real separate surveys, or records from different sources, field trips, museum collections, recorders etc.) 'Survey level' metadata corresponding to the SurveyKeys in the exchange file should be supplied as records in the MESH metadata catalogue. A metadata template can be downloaded from the MESH website, www.searchmesh.net.

Date fields

Dates should be supplied in the 'vague date' format, described briefly below.

Vague dates are created by specifying the start and end dates of a date range together with a one or two character code (DateType), which identifies the type of vague date. In most cases the DateType can be inferred from the values of the start and end dates but explicitly stating the code avoids any ambiguity, which might lead to subtly different interpretations. The table below gives the DateTypes allowed in the exchange format.

StartDate	EndDate	DateType	Description
16/06/2000	16/06/2000	D	Date specified to the nearest day.
16/06/2000	18/06/2000	DD	Date specified to a number of days.
01/06/2000	30/06/2000	O	Date specified to the nearest month (first day)

			of the month to the last day of the month)
01/06/2000	31/07/2000	OO	Date specified to a range of months (first day of the start month to the last day of the end month)
01/01/2000	31/12/2000	Y	Date specified to the nearest year (first day of the year to the last day of the year)
01/01/2000	31/12/2001	YY	Date specified to a range of years.
	31/12/2000	-Y	Only the end date to the nearest year known.
	31/12/2000	ND or U	'No date' or 'unknown'. If the date is not known this can be used with EndDate set to the date the dataset was compiled.

All date values must be supplied in the format DD/MM/YYYY (typically used in the UK) or as YYYY-MM-DD (an ISO date). **Dates with 2-digit years will be interpreted literally.** For example, the date 21/09/97 will be interpreted as the 21st September AD97! Ensure you supply the full year.

StartDate and EndDate

Mandatory The start or end date of the record. For example, if the start date is 16th June, 2000 this field is 16/06/2000. One or two digits can be supplied for day and month but take care to specify the year properly (see above). For UK dates the backslash character ('\') must be used as the separator. Use a hyphen ('-') for ISO dates.

DateType

Mandatory The vague date type of the dates used in the record. For example, if the start date of the record is 16th June, 2000 and the end date of the record is 18th June, 2000, then the DateType = DD, since it represents a number of days.

Biological fields

Biological data in the MESH Sample DEF may be either species (taxon) occurrence records, or biotope (habitat) occurrence records. There are numerous ways to record species data, and five options are available within this data exchange format: Abundance, Count, Presence, Percentage, and Score. Each occurrence record should be a taxon occurrence (and entry in TaxonVersionKey) OR a biotope occurrence (an entry in BiotopeKey) but NOT both. Populating one of these recording method columns for species occurrence records is *mandatory*.

TaxonName

Mandatory for a species record This is the name of the taxon recorded at this location. An entry must be present either in this field or in the Habitat field.

Abundance

Optional for a species occurrence record, but one recording method for each species record is mandatory This is the abundance of a taxon as measured by the SACFOR scale of **S**uperabundant, **A**bundant, **C**ommon, **F**requent, **O**ccasional, **R**are. Data should be entered as a single letter from this scale. See also the field survey guidance at www.jncc.gov.uk/page-2683 and www.jncc.gov.uk/page-2684 for more information.

Count

Optional for a species occurrence record, but one recording method for each species record is mandatory This is the count of the taxon as measured by individuals per m².

Presence

Optional for a species occurrence record, but one recording method for each species record is mandatory This is the presence or absence of a taxon, recording using "P" or "N" (not present).

Percentage

Optional for a species occurrence record, but one recording method for each species record is mandatory This is the percentage cover of a taxon, recorded as a number between 0 and 100.

Score

Optional for a species occurrence record, but one recording method for each species record is mandatory This is the abundance of the taxon as measured by a numeric score of 0 to 5. It indicates the biomass or quantity of flora and fauna present:

5 = Very high

4 = High

3 = Moderate

2 = Low

1 = Very low

0 = Not present

See also the field survey guidance at www.incc.gov.uk/page-2683 for more information.

Habitat

Mandatory for a biotope occurrence record Code or text description of the habitat recorded at this location. Each occurrence record should be a taxon occurrence OR a biotope occurrence but not both; a value must be present in this column or in TaxonName. If the [ImageFileName] field is completed, the [Habitat] field is mandatory.

Version

Mandatory for a biotope occurrence record The name and publication date of the classification system used in [Habitat].

ImageFileName

Optional for a biotope occurrence record The file name of a seabed image taken at the same location as a biotope occurrence record: the [Habitat] field is mandatory if [ImageFileName] is populated. If there is more than one image *per sample* then each image must be taken at a distinct location and therefore have different positional information. Include the file extension in this field (e.g. .jpg). The image files must be supplied with the data in the Sample DEF format.

Physical fields

Data which does not contain biological information may still be of use to the MESH project. The physical data fields will describe the seabed and record the depth for the description.

Depth

Mandatory The depth at which the occurrence record was recorded. This may be given either in metres or as a depth range if the exact depth is not known.

SeabedDescription

Mandatory Description of the seabed found at the location of the occurrence record. This is a free text field to describe the seabed encountered at the location of this record, for example the substrate type and texture.

Location fields

Each row of the dataset must contain location information as a valid georeference.

StartEast and StartNorth

Mandatory This should be provided as longitude and latitude according to WGS84. Positive longitude values indicate a position east of the Greenwich median, negative values positions to the west. Positive values of latitude indicate a position north of the equator.

EndEast and EndNorth

Conditional These fields should be provided for occurrence records relating to a section of video tow. In other cases the fields can be left blank. This should be provided as longitude and latitude according to WGS84. Positive longitude values indicate a position east of the Greenwich median, negative values positions to the west. Positive values of latitude indicate a position north of the equator.

Datum

Mandatory The datum of the long/lat coordinates, i.e. "WGS84" or "OSGB36".

Precision

Mandatory The spatial precision of the georeference in metres. We appreciate this might have to be estimated in some cases.