Data management

Good data management is as important as good data collection. Following good data management guidelines will ensure the data you collect today will be useful in the future. Your survey findings risk being lost or a challenge to access or use if data is not managed properly.

Data management recommendations

1. Ensure quality control guidelines are in place (see getting started - quality control, page 16).
2. When possible use waterproof paper or a waterproof recording slate in the field to collect data.
3. Survey recording forms must be fully completed. This should be checked before finishing a survey.
4. Recording forms completed in the field should be retained for filing – or recorded onto a clean recording form where necessary (hard copies should be kept of all data sheets).
5. Survey recording forms should be entered into a spreadsheet at the end of each day.
6. Data entry should be checked by a second person to minimise transcribing errors.
7. Electronic copies should be stored in relevant (survey type or location) folders. Use the naming convention guidance in the appendix for data and photo storage.
8. Back-up copies of all electronic data should be made.

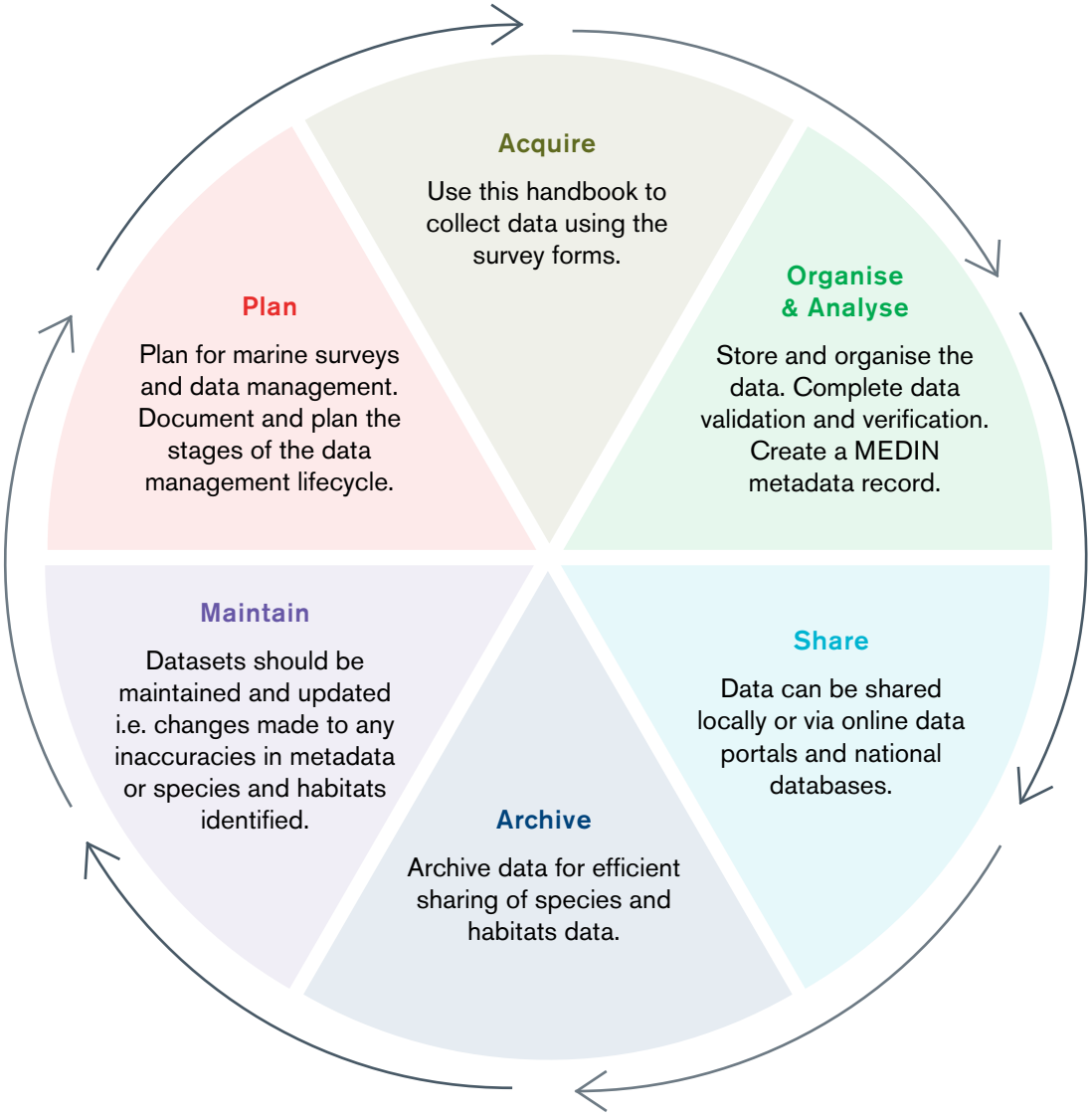
Data ownership

Data ownership and data sharing permissions should be agreed and documented at the time of the survey and/or at the time of organising data. We suggest that sharing and re-use be permitted under the terms of a Creative Commons Attribution Licence: CC BY (https://creativecommons.org/licenses/by/4.0/). We have provided space to record the data owner on the survey recording forms, ensure this is completed; the data owner is likely to be the community group organising the survey.

How to manage your survey data

As part of the community-led marine biodiversity monitoring project, NatureScot are working with community groups to find opportunities and solutions to allow data collected using the handbook to flow into national data portals, following the data lifecycle steps. Following the guidance on collecting and recording data within this handbook will help with establishing a data flow. After a pilot has been undertaken and a data flow established, we will provide an update and further guidance. In the interim, we have provided some information on citizen science portals, websites to share data and the appropriate data archiving centre (DAC) that are available. Ensure you check the copyright permissions and privacy policies of all portals, websites and DACs before use. Please note that you should only submit records via one data storage or sharing website/ app/portal to avoid duplication of records.

Data lifecycle



▲ **The data lifecycle illustrates the elements of data management. This will improve the overall quality of your data and should be used for all community-led survey projects.**

Quality assurance

The quality of the data you collect will be influenced by the survey design. You should follow the guidance within this handbook and follow the methods and survey forms fully to maximise the quality of your data. It's important that the quality of the data is known as it will be more scientifically useful and more likely to be used as evidence by policy makers.

There are two steps you can take to help increase the accuracy and overall quality of your data:

Validation

- This is the process of checking all survey data is complete and that the data satisfies set formats or values (e.g. checking the data is organised correctly with all data required - who collected the data, dates, times, GPS coordinates, etc.). Please see the 'Metadata – the data behind the data' poster on page 6.
- You should create a Marine Environmental Data and Information Network (MEDIN) Discovery Metadata record (see MEDIN compliance below). This is an automated process to help validate your data using the MEDIN Discovery Metadata Editor.

Verification

- This is a manual process where you check the species and/or habitat identification is correct. For example, you can check the marine species or habitat information by visually checking videos or photographs associated with the data. Verification can also be undertaken by accompanying a few participants and observing the measurements that they take or crowd sourced by asking people to assess each other's photographs or asking different people to take measurements at the same time. Sometimes it may be necessary to contact experts, to determine species or habitats identification. Don't be afraid to flag a record as uncertain.
- It is important to keep a record of species and/or habitat assignment history (sometimes referred to as taxon determination history). Practically this means you should keep a record of who assigned what species or habitat identification and when. This is important to track and maintain the quality of the data.
- The survey methods have been designed to ensure that photographic or video evidence is captured which will allow verification to be completed.

  Available citizen science portals for verification - *can be used for species records recorded in the species image library survey method.*

  #### iSpot

  iSpot is a website aimed at helping anyone identify anything in nature. Users can upload their observations of wildlife, help each other identify it, and share and discuss what they've seen. Records submitted to iSpot will be made available through the NBN Atlas once verified.

  Visit www.ispotnature.org.

  #### iRecord

  iRecord is a UK based website for sharing wildlife observations, storing biodiversity records and getting them reviewed by experts for species identification. You can also explore maps and reports of your data. The data can also be shared via the platform to other data repositories, such as NBN Atlas.

  Visit www.brc.ac.uk/irecord.

Taxon dictionaries

A taxon dictionary is a register of names of organisms that are taxonomically correct, controlled and kept up to date by experts who are responsible for controlling the quality of the information. In practical terms, this means a list of species names that are correctly accepted and deemed correct, their common names (if any), their classification within the taxonomic hierarchy and the background such as the original description (e.g. common starfish – *Asterias rubens* Linnaeus, 1758).

Marine Dictionaries to use for Scotland:

- **Marine Species of the British Isles and Adjacent Seas (MSBIAS)**

  This is a regional subregister of World Register of Marine Species (WoRMS) for marine species around the UK. This taxon dictionary is specifically used for Scottish and UK national databases. When identifying marine species or verifying records, you should use MSBIAS. This is the accepted taxon dictionary to ensure your data meets MEDIN compliance for marine data standards (see below).

  Visit http://www.marinespecies.org/msbias/

Marine habitat classification

The classification lists all seafloor habitats currently known to occur in UK waters. These habitats are organised in a hierarchy whereby each level introduces more detail. In any place on the shore or seabed, a suite of species will occur, each adapted to the particular environmental conditions of that place. Where such a suite of species occurs in other locations under similar environmental conditions, it can be defined as a community (or association or assemblage) of species which is occurring within a particular habitat type. The collective term biotope is now in common usage to encompass both biodiversity and environmental elements.

The classification is widely used by Government bodies, academic institutions, the private and charity sectors as a system for the consistent description of habitat types. If you wish to assign a habitat identification to a survey record, the biotope system should be used, however you should only assign classification to the appropriate level within your ability and the biotope description must be in agreement with your survey results.

Visit https://mhc.jncc.gov.uk/.

MEDIN compliance

Survey methods within this handbook have been created following the Marine Environmental Data and Information Network (MEDIN) data guidelines. The survey forms contain all mandatory MEDIN metadata fields required. This ensures your survey dataset, if completed fully, can be registered with MEDIN.

We recommend at the end of a survey or survey season to create a both a MEDIN guideline formatted record for the survey and a MEDIN discovery metadata record (one per survey). This discovery metadata record will ensure that your survey data is described in a standard format and publishing the discovery metadata to the MEDIN portal will mean others will know your data exists. The MEDIN Discovery Metadata Editor is an online tool that will help you create a record.

You can create an account at: www.dassh.ac.uk/medin_metadata/login. Details of how to create MEDIN compliant content can be accessed at https://www.medin.org.uk/medin-discovery-metadata-standard, and further assistance is available from the MEDIN Metadata Helpdesk (medin.metadata@mba.ac.uk - 01752 426237). MEDIN periodically run workshops on how to structure data into MEDIN guideline format and generate a discovery metadata record.

Visit www.medin.org.uk/.

Data storage

Your data should be well organised and stored securely. Consider how you are going to store the data securely and make them available in the long-term – be aware of data protection legislation if storing personal data, it is common practice with biological records to record who surveyed them and who identified what was found. Names are treated as personal data, community groups should seek advice on compliance with legislation regarding drafting appropriate privacy notices including identification of the appropriate legal basis for processing individual's names.

Sharing your data

Data must be quality assured before it can be submitted to national databases or data archiving centres.

MEDIN Portal

To share your discovery metadata records with others, the MEDIN Discovery Metadata Editor (see above) is a tool which allows you to publish your discovery metadata (one record per survey) to the MEDIN Portal. Ensure you tick the correct box to allow the publication of each discovery metadata record to the MEDIN portal. This will make your data more discoverable by others, documenting that a survey has taken place and is a part of good data management practice. This metadata record will not provide others with the results of your survey, but it does let others know the data exists.

Visit https://portal.medin.org.uk/portal/start.php

The NBN Atlas

The National Biodiversity Network (NBN) Atlas is a free online tool hosting the UK's largest collection of biodiversity information enabling data to be shared, analysed and researched. NBN Atlas is a tool to share the results of your survey.

Visit https://scotland.nbnatlas.org/.

Local Records Centres

If there is a Local Records Centre covering the area for your wildlife record then send the details to them. Contact them first to make sure you get it in the right format. They may be able to help with all the hard work from then on and submit the records to appropriate national databases and recording schemes and societies.

Visit www.brisc.org.uk/.

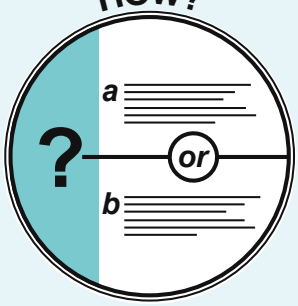
DASSH UK Data Archive Centre - *the archive for marine species and habitats data*

The benefit of submitting your data to a data archiving centre in MEDIN data guideline format is that the data can be easily shared with others and in a format others can easily use. While the MEDIN Portal (see above) will publish the associated metadata, if you would like to share the data results (i.e. the species and habitats information) with others, then data archiving centres can make this process straight forward. To archive your data, the data will need to be entered into MEDIN data guideline spreadsheets. Data guidelines relevant to your marine dataset can be found and downloaded for use here: www.medin.org.uk/data-standards/medin-data-guidelines. The DASSH UK Data Archive Centre is the suggested location for seabed species and habitats data.

Visit www.dassh.ac.uk.

METADATA
THE DATA BEHIND THE DATA

HOW?



Can you describe how the data was collected?
For example, our scientists collect data on board the research ships using water sample collectors called CTD'S

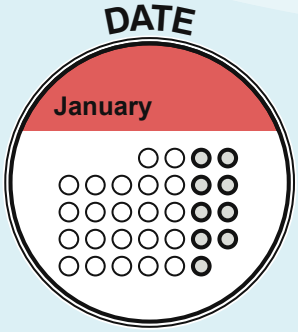
WHO?

Who collected and recorded the scientific data?
Data is collected by all kinds of scientists from different fields and organisations, this information is recorded with the data

WHY?

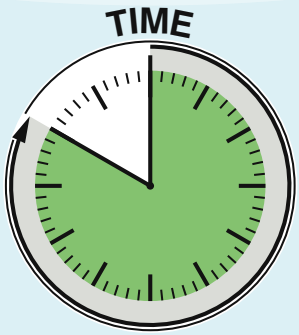
Can you explain why the data was collected?
For example, long-term data collection can be for evidence surrounding climate change

DATE



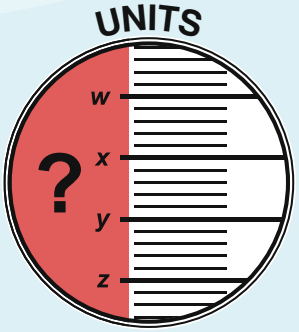
What was the precise date of the data collection?
Knowing which day, of which month and in which year the data were collected is very important to track changes in data

TIME



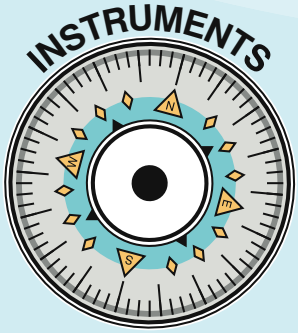
What time precisely was the data collected?
We record the exact hour, minute and second using the standard 24 hour notation method of hh:mm:ss, for example 13:18:54

UNITS



Which units will you use to measure the data?
Any data that does not fit the trend or looks unusual will be flagged as suspect and investigated to ensure quality

INSTRUMENTS

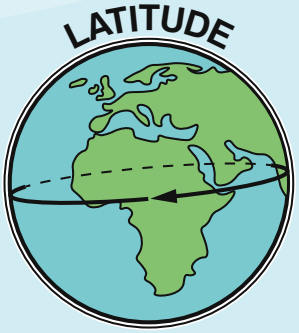


Which instruments are used to measure data?
Scientists use instruments to collect and analyse data, for example we could use Calibrated Verification Thermometers

LONGITUDE

What is the Longitude location of the data?
The distance east to west from the prime meridian in Greenwich, when used with Latitude a single point can be found

LATITUDE



What is the Latitude location of the data?
We measure the distance north or south in degrees along a meridian from the centre of the Earth's equator

METHOD

Can you explain the exact method step by step?
In the laboratory we state and record the processes completed by the scientist during data collection and analysis

POLICY

What is the policy on releasing the data?
Some data can be released immediately whilst other data is legally held by the scientist for 2 years as per the data policy

QUALITY CHECK

How have you checked the quality of the data?
All data is analysed and checked, any that does not fit the trend or looks unusual should be flagged and investigated

National Oceanography Centre

BODC British Oceanographic Data Centre
National Oceanography Centre

MEDIN