

MEDIN

marine environmental
data & information network

Data quality, management and access

Sean Gaffney
sgaf@bodc.ac.uk

*Measure once,
use many times*



Importance of data (examples)



- In context of post-spill monitoring, data can be needed for (amongst other examples):
 - Tracking the movement of contaminants: current fields and circulation models needed for this
 - Determining environmental impact of the spill. Previous surveys / measurements needed for baseline, records of other similar pollution are useful comparison



Cautionary tale

- How do monitoring agencies / other stakeholders know they can trust the data they need?
- This video shows some of the problems that can occur when trying to use other peoples data:

<https://www.youtube.com/watch?v=N2zK3s>

Atr-4

- Issues of opinion of usability of data
- Issues locating data
- Issues of reading files
- Issues of inadequate documentation and labelling



- Adhere to **F.A.I.R** data principles (Data Fairport Initiative, Netherlands, Jan. 2014)
- **Findable**
 - Meta(data) are registered in a searchable resource, have unique persistent identifiers, data have rich metadata, metadata specify a data identifier
- **Accessible**
 - Meta(data) can be retrieved by identifier using a standard open communication protocol, protocol allows for authentication/ authorization, metadata are accessible even when data are not



FAIRness (cont.)

- **Interoperable**

- Meta(data) use a formal, accessible, shared and widely applicable language, meta(data) use vocabularies that follow FAIR principles, meta(data) include qualified references to other meta(data)

- **Re-usable**

- Meta(data) have many accurate and relevant attributes, meta(data) are released with clear data usage conditions, meta(data) are associated with their provenance, meet domain relevant community standards.

<http://www.nature.com/articles/sdata201618>



BODC being F.A.I.R (examples)



**British Oceanographic
Data Centre**

NATURAL ENVIRONMENT RESEARCH COUNCIL

- Development of the NVS2 vocabulary server
(https://www.bodc.ac.uk/data/codes_and_formats/vocabulary_search/)
- The Published Data Library
(https://www.bodc.ac.uk/data/published_data_library/) which allows unrestricted access to data with DOI citation



MEDIN and F.A.I.R



In operation since 2008

Funded by 16 sponsors

Open partnership

Budget ~£500K



Work with academia, government and industry

Easier data sharing



MEDIN Marine Discovery Metadata Standard (Findability)

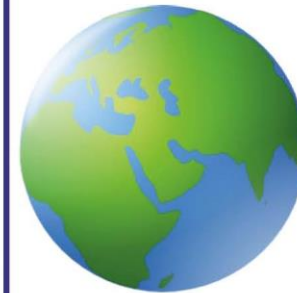
UK Marine Standard:



UK Standard:
GEMINI2



European Standard:
INSPIRE Directive



International Standard:
ISO19115

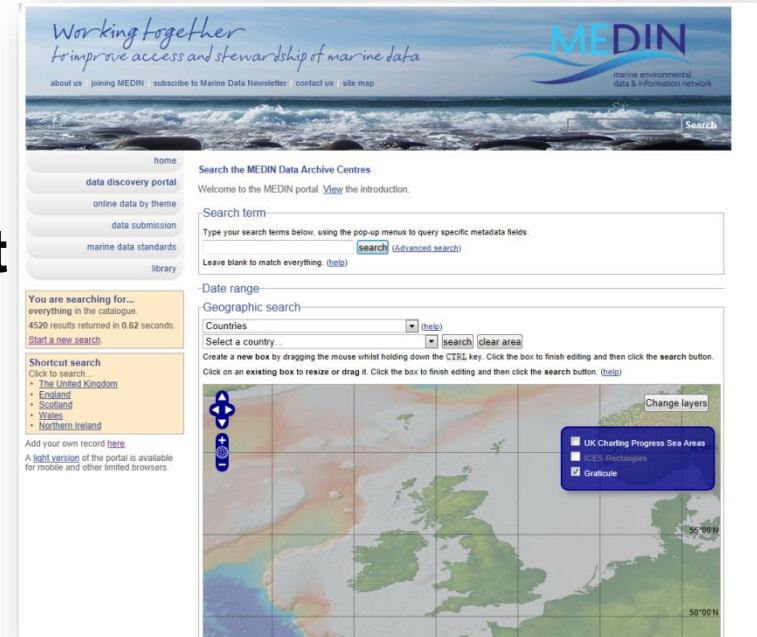


The MEDIN portal (Findability)

- portal.oceannet.org
- “... provides a single point of access from which to find out about marine data from all the different UK organisations...”

MEDIN 2014-2019 Business Plan

- MEDIN portal went online in June 2010
- Aim is to have all public funded data available under Open Government Licence



The screenshot shows the MEDIN portal homepage. At the top, it features the tagline "Working together to improve access and stewardship of marine data" and the MEDIN logo. Below this is a navigation menu with options like "home", "data discovery portal", "online data by theme", "data submission", "marine data standards", and "library". A search bar is prominently displayed, with a "Search" button. Below the search bar, there are sections for "You are searching for..." (showing 4520 results), "Shortcut search" (listing countries like The United Kingdom, England, Scotland, Wales, and Northern Ireland), and "Geographic search" (with a map of the UK and options for "UK Charting Progress Sea Areas", "ICES Rectangles", and "Graticule").

- Portal being upgraded and modernised. Tender process to start summer 2016/17.



Accessibility and inter-operability



Data in 7 Data
Archive Centres
(DACs)



MEDIN Data Guidelines (re-usability)

data guidelines

**The MEDIN data guidelines are currently being updated.
Please check back regularly to ensure you are using the latest version.**

Data Guidelines provide a list of information that should be collected with your data to ensure they can be re-used in the future. The guidelines are tailored to different methods and are arranged below by theme.

An Introduction to MEDIN Data Guidelines is available to [download](#) (977 KB).

An independent pilot study on the usability of MEDIN data guidelines is available to [download](#).

An example dataset that complies with MEDIN data guidelines can be downloaded [here](#) (376 KB).

bathymetry

human impact

marine archaeology

marine biodiversity

marine chemistry

marine geology and geophysics

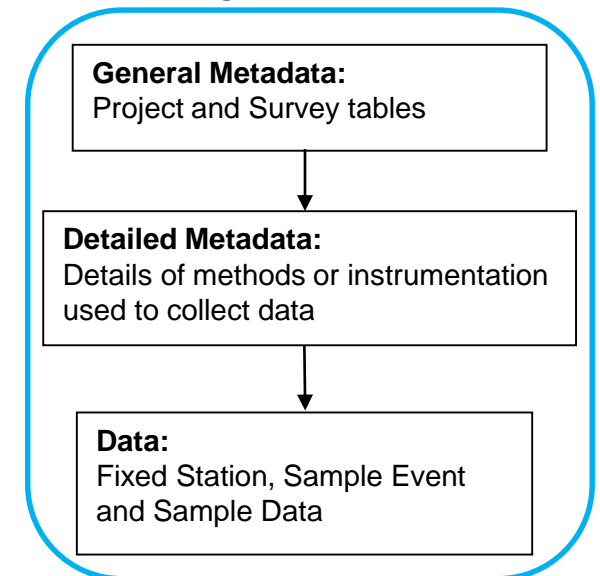
- MEDIN data guideline for archiving of digital images (sep11)
- MEDIN data guideline for marine gravimeter/ gravity data (mar12)
- MEDIN data guideline for magnetometer/ magnetic gradiometer data (mar12)
- MEDIN data guideline for seismic data (may 14)
- MEDIN data guideline for bathymetry data (dec 13)
- MEDIN data guideline for species and benthos data by grab or core (mar14)
- MEDIN data guideline for sampling sediment and rock characteristics (apr12)
- MEDIN data guideline for side scan sonar data (mar15)

Seabed Survey Data Products

physical oceanography

Provide guidance on what metadata needs to be collated to allow data to be re-used – not guidance on how to collect data

Drafted by DACs and other expert bodies. Help to speed up data ingestion into DACs and subsequently makes re-use and data sharing easier



Summary

- Getting hold of the right data quickly is vital when reacting to spill events
- Using correctly designed search portals which allow users to ID and locate the right data enables this to happen
- The data must be downloadable in a useful file format and be clearly described so no time is lost in interpretation of the contents
- Adherence to the **F.A.I.R data principles** ensures these requirements can be met
- MEDIN (including BODC and the other DACs) adheres to the **F.A.I.R principles** and acts as a central focus for marine data in the UK.
- Please deposit your data and search for data on MEDIN!



Thank you



“Data is a precious thing and will last longer than the systems themselves.”

Tim Berners-Lee

British computer scientist, best known as the inventor of the World Wide Web

- Any questions?

