Marine Environmental Data and Information Network (MEDIN)

DAC Network Annual Report for 2020-21

'Measure once, use many times'

Executive summary

This report represents the continuing efforts of the distributed network of Marine Environmental Data and Information Network (MEDIN) Data Archive Centres (DAC) to underpin the overall aims of MEDIN to facilitate access to UK marine data. Whilst the majority of the content of the report is derived from the portal it is also important to reflect on the achievements against the DAC Workstream (WS1).

Despite the challenges of the global COVID19 pandemic, all WS1 activities remain on track thanks to the efforts of the DACs, including the aspirations for a DAC-wide approach to provenance tracking, and the development of a common Application Programming Interface (API) for access to marine environmental data. These items will be further progressed in the next reporting period. In addition, work has been

initiated to improve the coordination and archiving of multi-disciplinary data through an initial triage process, and improved cross-DAC communication.

Many DACs have also taken the opportunity to transition to the CoreTrustSeal accreditation scheme, which provides a globally recognised framework for the accreditation of data repositories, fully aligned with the MEDIN DAC accreditation process.

Whilst there is still scope for improvement in the direct access to data ("2-clicks"), the direction of travel is positive, and significant improvements can be seen within the detail of the report. The DACs continue to provide the foundational infrastructure for the delivery of Findable, Accessible, Interoperable, Reusable (FAIR) and open access to UK marine data and in the promotion best practice in marine data management.

Summary highlights

The Marine Environmental Data and Information Network (MEDIN) coordinates an operational network of seven linked marine Data Archive Centres (DACs) covering bathymetry; fish and shellfish, fisheries, aquaculture and related samples; the historic environment; marine geology and geophysics; marine species and habitats; marine meteorology; and water column oceanography. The DACs continue to archive data from MEDIN partners and third-party organisations to agreed individual programmes.

This is the second DAC annual report falling within the 5-year period of the current MEDIN Business Plan. DAC metrics are now applied more consistently across the DAC network, being pulled directly from the MEDIN Portal, where possible.

The 2020-21 DAC annual reports show that:

- 65% of the datasets described in the MEDIN portal are available from the MEDIN DACs. That is 10,170 datasets managed, quality controlled and distributed by MEDIN's coordinated network of DACs.
- More than 95% of the datasets available from MEDIN DACs are accessible online and 40% are downloadable within 2-clicks of finding them on the MEDIN portal.
- 4% of the datasets in the MEDIN Portal available from MEDIN DACs have Digital Object Identifiers (DOIs).
- 26% of the datasets have been submitted to the MEDIN DACs from Marine Science Coordination Committee (MSCC) organisations.
- More than 1.6 million 'requests' for data were made to MEDIN DACs during the reporting year.

1 Introduction

MEDIN has established an operational network of linked marine Data Archive Centres (DACs) to provide secure long-term storage for and access to marine data. This network provides the capability for users to upload and retrieve data. Those organisations archiving data at a MEDIN DAC have free access to their data, and DACs manage third-party access to these data according to the data provider's specification.

The required capabilities of DACs within the MEDIN framework are:

- To ensure the secure, long-term curation of key marine data sets, according to best practice and to relevant national and international standards.
- To make available clear, searchable information on their data holdings by the generation and publication of metadata on the MEDIN portal.

- To form the first point of call for expertise in the management of marine data.

In addition, MEDIN will, on request from the data provider, publish metadata records to data.gov.uk and hence the INSPIRE geoportal.

As a condition of its accreditation, each MEDIN Data Archive Centre is required to provide a short annual report so that Sponsors can assess how well the DAC framework is operating.

The MEDIN Sponsors' Board has emphasised the following requirements:

- Provide a statement on funding and sustainability.
- Include Key Performance Indicators, specifically measures of use (numbers of enquiries, numbers of downloads).
- Further information on dissemination – how is access to data currently served and how do the DACs see their interaction with the MEDIN portal.

This document provides a report on the current status of DACs in terms of metadata records in the MEDIN Portal where the DAC is custodian of the data, requests from users for data, and financial outlook. This is a summary of information from the individual DAC reports, which are available on request to enquiries@medin.org.uk.

2 DAC Listing

There are currently seven DACs in the MEDIN DAC network, as listed in Table 1 below. More details are available on each DAC through links on the DAC web page on the MEDIN website at https://www.medin.org.uk/data-archive-centres. These pages include information on what types of data are held and top-level guidelines on how to submit data to, and to access data from, each DAC.

Table 1: MEDIN Data Archive Centres

Name	Coverage	Contact Information	Web links	MEDIN Status
British Oceanographic Data Centre (BODC)	Water Column Oceanography data	enquiries@bodc.ac.uk 0151 795 4884	www.bodc.ac.uk	Accredited 2009; Re-accredited 2017; operational.
British Geological Survey (BGS)	Marine Geology and Geophysics data	medin@bgs.ac.uk	www.bgs.ac.uk/geolo gical-data/nationalgeoscience-datacentre/ngdc-datamanagement/marinegeoscience-datamanagement/	Accredited 2009; Re-accredited 2017; Core Trust Seal accreditation 2018; operational.
The Archive for Marine Species and Habitats Data (DASSH)	Marine Species and Habitats data	Dassh.enquiries@mba.ac.uk 01752 633291	www.dassh.ac.uk	Accredited 2009; Re-accredited 2017; operational.
Met Office	Marine Meteorology data	enquiries@metoffice.gov.uk	www.metoffice.gov.uk	Accredited Dec 2011; Re-accredited 2018; operational.
United Kingdom Hydrographic Office (UKHO)	Bathymetry data	CustomerServices@ukho.gov.uk	www.gov.uk/guidanc e/inspire-portal-andmedin-bathymetrydata-archive-centre	Accredited 2009; Re-accredited 2017; operational.

FishDAC • Cefas • Marine Scotland Science (MSS) • DASSH	Fisheries data - Fish and Shellfish, Aquaculture and related samples and environmental data	Cefas: data.manager@cefas.co.uk	http://www.cefas.defra.gov.uk/publication s-and-data/fishdac.aspx	Accredited 2012, Reaccredited 2018; operational.
		Marine Scotland Science: jens.rasmussen@gov.scot	http://www.gov.scot/ Topics/marine marine.gov.scot maps.marine.gov.scot data.marine.gov.scot	Accredited 2012; Re-accredited 2018; operational.
Historic Environment DAC • Archaeology Data Service (ADS) • Historic Environment Scotland (HES) • Royal Commission on the Ancient and Historical Monuments of Wales (RCAHMW)	Marine Historic Environment fieldwork derived data	Archaeology Data Service: help@archaeologydataservice.ac.uk	http://archaeologyda taservice.ac.uk	Accredited 2012; Re-accredited 2018; operational; Core Trust Seal accreditation 2020
		Historic Environment Scotland: peter.mckeague@hes.scot	www.canmore.org.uk	Accredited May 2014; Re-accreditation due 2019; operational.
		Royal Commission on the Ancient and Historical Monuments of Wales gareth.edwards@rcahmw.gov.uk	www.coflein.gov.uk	Accredited June 2016, operational.

3 DAC Performance

Each year, MEDIN asks the DACs to report on their performance using a standard set of metrics.

The metrics are now pulled directly from the MEDIN Portal where possible. They therefore show the number of metadata records in the portal where a DAC is the custodian of the data. This is not always directly representative of the number of data sets held at a DAC because of some variability in the granularity of metadata records. This is the second year of reporting using metrics from the MEDIN Portal.

The key metrics are as follows:

- Total number of metadata records present in the MEDIN Portal where each DAC holds the data.
- Number of new or updated records in the MEDIN Portal in reporting year where each DAC holds the data.
- Number of records where DAC holds the data, with:
  - A URL leading to online access to data
  - A URL allowing direct access to data (i.e., within 2-clicks)
  - A URL containing a Digital Object Identifier
- Number of records in the MEDIN Portal for Marine Science Coordination Committee (MSCC) partners who have data archived in a MEDIN DAC.
- Number of requests for data for each DAC (using figures supplied by DACs as it is not possible to obtain this from the MEDIN portal).

3.1 DAC Metrics

The metrics for 2019-20 and 2020-21 are shown in Table 2.

Table 2: Annual metrics for the MEDIN DACs

Year	BODC	BGS	DASSH	Met Office	UKHO	Cefas	MSS	ADS	HES	RCAHMW
Total number of metadata records where DAC is custodian										
2019-20	1107	857	710	7	4736	2058	282	74	47	26
2020-21	1107	857	723	7	4736	2096	308	263	47	26
New/updated records in reporting year										
2019-20	45	22	496	1	0	536	54	74	25	13
2020-21	2	693	152	2	0	421	54	189	7	1
Records with online access to data										
2019-20	1056	855	631	2	4736	1914	240	74	17	0
2020-21	1056	855	635	2	4736	1958	254	153	17	0
Records with 2 clicks to data										
2019-20	997	694	165	1	0	1914	57	73	12	0
2020-21	997	694	124	1	0	1958	64	150	12	0
Records with DOI										
2019-20	54	0	1	0	0	101	49	73	12	0
2020-21	54	0	9	0	0	115	56	150	12	0
Total number of records for data from MSCC organisations										
2019-20	80	151	15	5	34	2058	213	0	0	0
2020-21	80	151	22	5	34	2096	225	0	0	0
New records in reporting year for data from MSCC organisations										
2019-20	4	0	9	0	0	536	21	0	0	0
2020-21	2	12	9	2	0	421	18	0	0	0

The metadata harvest pipeline for BODC is currently broken and this means no new records harvested into the MEDIN Portal during 2020-21. This is being rectified as part of BODC's programme of work in August 2021-22.

Please note that it is not advisable to compare absolute values between DACs, as the size of data sets can vary significantly between DACs (and even within DACs). For instance, all the data held in the Met Office MEDIN DAC for marine meteorology data are held within 7 data sets, which are augmented each year with that year's new data. Approximately 9.5 million observations were added to the Met Office's data sets during 2020-21.

Figure 1: Percentage of metadata records in the MEDIN portal per DAC.

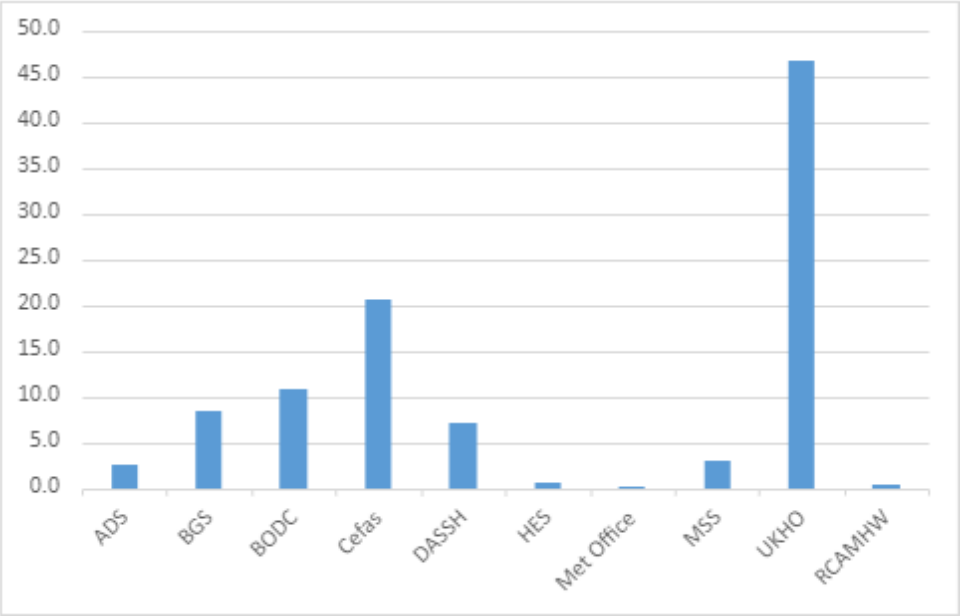


Figure 1 shows that the UKHO is the DAC with the most metadata records in the MEDIN Portal, with almost 50%. However, as noted earlier, the different DACs have different granularity in their metadata records so a direct comparison between DACs is not appropriate. Two of the DACs have updated or increased their metadata records in the MEDIN portal by over 70% in the past year (Figure 2). Some of this will relate to new datasets, or new data being added to existing datasets such as time series, and some to improving the quality of existing metadata. Note that making any updates to existing metadata records will count as changes in this metric.

Figure 3 shows there are six MEDIN DACs with over 80% of the metadata records for data they hold that have a Uniform Resource Locator (URL) leading to some form of online access to that data (not necessarily direct access in 2 clicks). MEDIN has been promoting direct access to data for several years and is pleased to record that three of the MEDIN DACs now provide direct access to data (within 2 clicks) from over 80% of their metadata records in the portal (Figure 4). One of the ways to provide direct access to data is using a Digital Object Identifier. Note that the percentage of ADS records offering access within 2 clicks and DOI has gone down from last year. This is because a number of older metadata records were harvested into the MEDIN portal, so the overall number of ADS records in the portal has increased, but the percentage offering direct access/DOI has dropped.

The breakdown of country of origin for the metadata records in the MEDIN portal where the DACs are custodian is shown in Figure 6. Figure 7 shows the percentage of records at each DAC where data originate from organisations that are involved with the Marine Science Coordination Committee (MSCC). MEDIN is a partnership initiative of the MSCC and MEDIN reports progress to MSCC. The number of MSCC organisations archiving data at MEDIN DACs is very variable across the DACs, as the data from each MSCC organisation is more relevant to some DACs than others.

Figure 2: Percentage of metadata records in the MEDIN portal per DAC that are new or were updated during 2020-21.

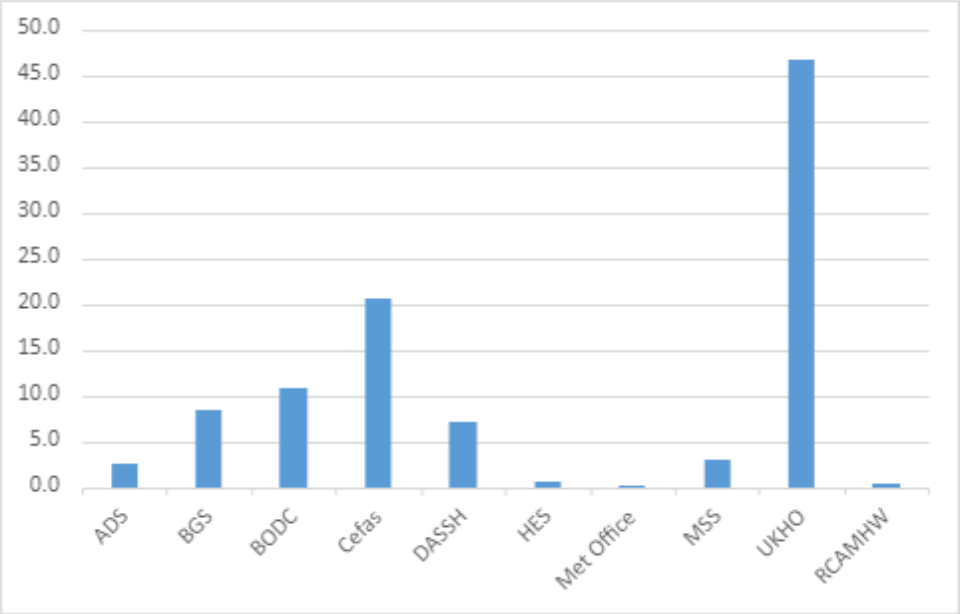


Figure 3: Percentage of metadata records per DAC in the MEDIN Portal with online access to data.

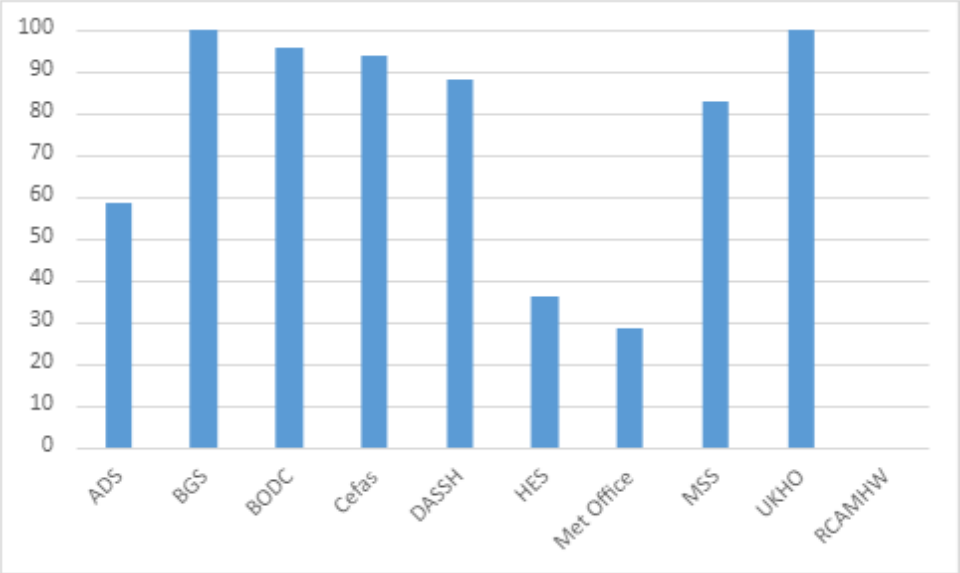


Figure 4: Percentage of metadata records per DAC in MEDIN Portal that offer 2-clicks to data.

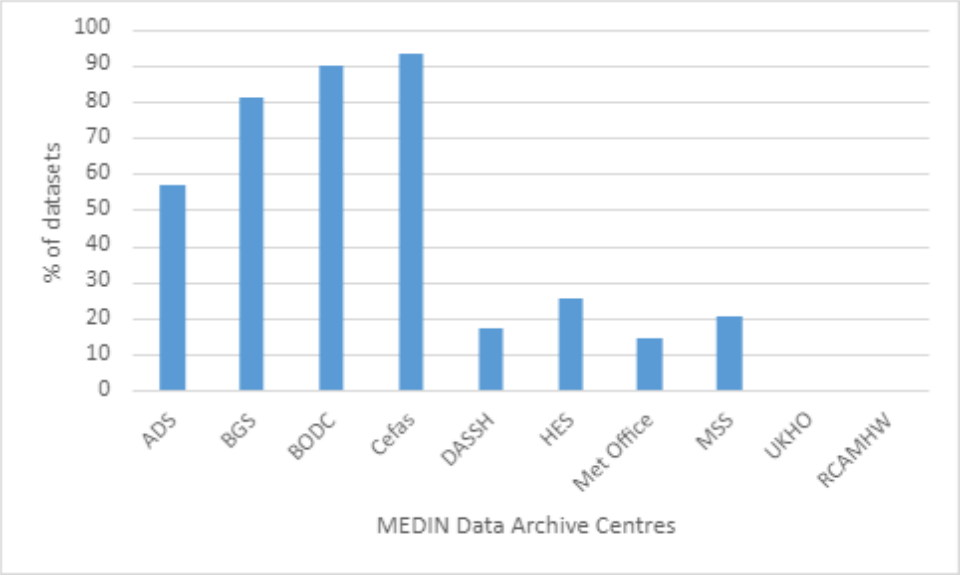


Figure 5: Percentage of metadata records per DAC in the MEDIN Portal with a Digital Object Identifier (DOI).

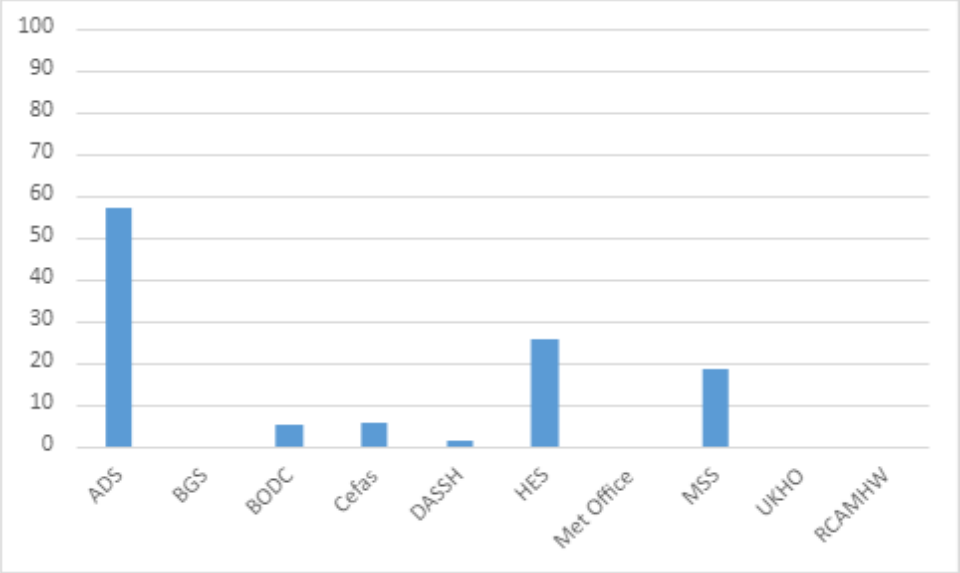


Figure 6: Percentage of metadata records per DAC in the MEDIN Portal by country of origin.

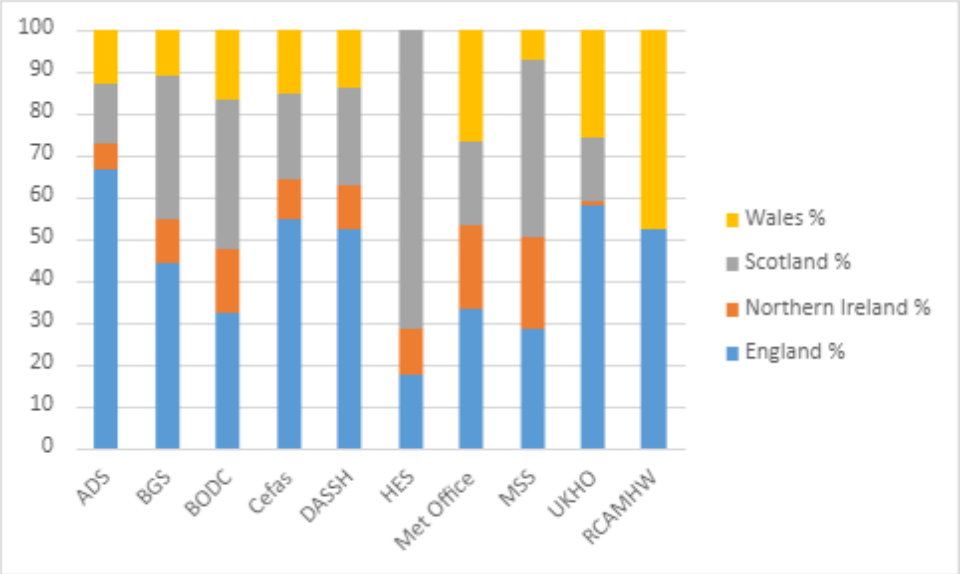
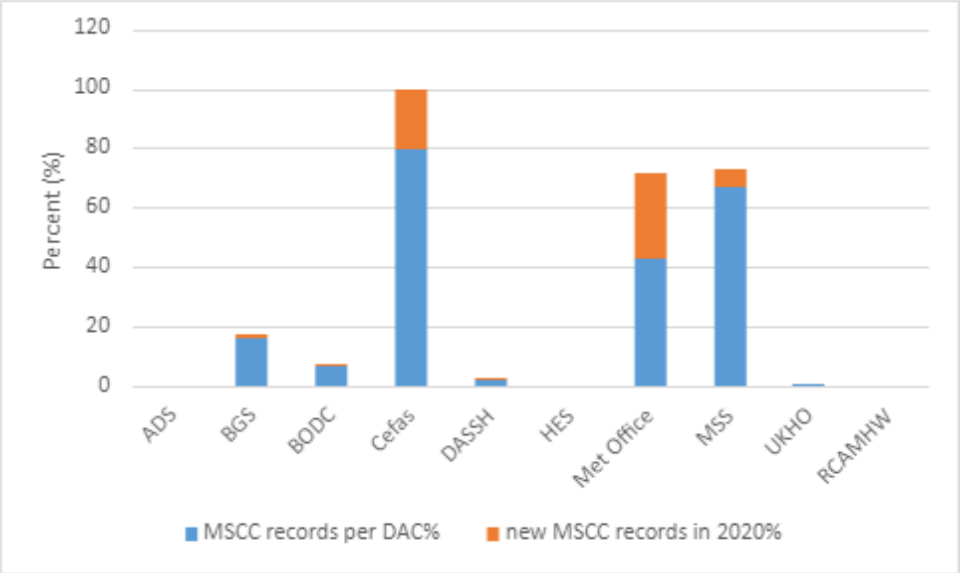


Figure 7: Percentage of all metadata records per DAC in the MEDIN Portal where data is archived in a DAC and was collected by MSCC organisations.



3.2 New datasets:

The MEDIN DACs continue to archive major new datasets. Table 3 summarises new datasets archived by each DAC during the financial year (FY) 2020-21.

Table 3: Summary of new datasets archived at MEDIN DACs during 2020-21.

DASSH	**RCAHMW (Historic Environment DAC)**

New datasets archived during 2020-21 include: • Marine Science Framework Directive (MSFD) datasets from pelagic work – nationally important data sets for policy making • Joint Nature Conservation Committee (JNCC) Habitat Data – to demonstrate downstream linkages between European Marine Observation and Data Network (EMODnet) Biology and EMODnet Seabed Habitats • 8 University of Plymouth EMODnet funded datasets - meeting one of our key performance indicators to have more academic interaction	No new archives have been added during the year. This is due to our 'Digital Delivery' project, to create a totally new platform for our data, including our archive catalogue, site data and online delivery system, Coflein. Consequently, our data systems have been ring-fenced to new records during the year. The redesign and build were completed in March 2021.
Met Office	**BGS**
There are 7 datasets on the MEDIN portal, which together describe most in-situ marine meteorological observations collected by the Met Office. The Met Office differs from other DACs in that it doesn't add new data sets, instead it adds new observations to its existing data sets. Overall, the number of unique observations contained within those datasets increased by over 9.5 million during 2020-21.	New datasets archived during 2020-21 include: • Wessex Archaeology London Gateway Geophysics dataset – related to ADS dataset • Legacy seismic data prepared for an enquiry ready for release
Cefas (FishDAC)	**Marine Scotland Science (FishDAC)**

We selected the following major datasets to highlight as they represent time series which may be used for multiple scientific purposes. **Phytoplankton abundance data from the Gibraltar coastline 2009-2019**. Phytoplankton Analysis Team, Cefas & HM Government of Gibraltar (2020). Cefas, UK. V1. doi: https://doi.org/10.14466/CefasDataHub.111 This dataset fills an apparent data gap of phytoplankton data from wider straits of Gibraltar region and facilitates the flow of data to international initiatives including Ocean Biodiversity Information System (OBIS) and Global Biodiversity Information Facaility (GBIF). **Zooplankton abundance data derived from the Plankton Imager system from the Western English Channel and Eastern Irish Sea from 2016 to 2019.** Pitois *et al* (2020). Cefas, UK. V1. doi: https://doi.org/10.14466/CefasDataHub.101 This is the first publication of this kind of Artificial Intelligence (AI) derived data, using brand new technology and methods bespoke to Cefas and our collaborators. **Trophic interaction data and taxonomic information for fish predators and their prey spanning the North East Atlantic and its marginal seas between 1836-2013.** Thompson *et al* (2020). Cefas, UK. V1. doi: https://doi.org/10.14466/CefasDataHub.102 This data product covers a very long time period and improves the understanding of how species interact to inform assessments of the health of marine food webs.	Most of the datasets archived reflect ongoing operations at Marine Scotland to collect fisheries and environmental data for management and monitoring. The majority of the surveys listed are recurring surveys that form part of internationally coordinated surveys. Marine Scotland also produces annual statistics on fisheries and aquaculture which now include the detailed data published. Some additional spatial data resources have been updated, and research reports have been published with the accompanying data. • 1 Deep water survey (North East Atlantic Shelf Slope) • 2 North Sea International Bottom Trawl Surveys (Quarter 3 2020 + Quarter 1 2021) • 2 West Coast Bottom Trawl Surveys (Quarter 4 2020 + Quarter 1 2021) • 1 Gear trials/gear comparison survey • Updated Girnock and Baddoch Fish trap counts • Scottish Fish Farm Production Survey data • Scottish Shellfish Farm Production Survey data • Scottish Sea Fisheries Statistics • Discard Survival and Condition in Orkney Brown Crabs (Cancer pagurus) – Report + data • Engaging the Fishing Industry in Marine Environmental Survey and Monitoring Report + data • Deepwater survey trawl manual • Spawning layers updated for whiting using ecological niche modelling • Haddock spawning updates 2017 (North Sea and West Coast) • Cod spawning updates 2016

	• Girnock and Baddoch: Ova to Ova StockRecruitment Dataset • 2020 MSS H2020: SMARTFISH WP8 Scotia Research Cruise • Deepwater Elasmobranch Species Data From MSS Trawling Surveys 1996 – 2019 • The National Electrofishing Programme for Scotland (NEPS) 2019 • Inshore Fisheries Pilots - Outer Hebrides Inshore Fisheries Pilot Area
UKHO	**HES (Historic Environment DAC)**

400 new bathymetric datasets were registered in the archive. We have received new permissions from the following organisations to supply bathymetry data through the Archive Centre: ABP Ayr, ABP Troon, ABP Fleetwood, ABP Garston, ABP Plymouth, ABP Silloth, Dart Harbour & Navigation Authority, Exeter City Council, Exeter Port Authority, Exmouth Marina, Isle of Man Harbours Division, Lynmouth Harbour Authority, Peel Ports London Medway, Port of Ramsgate (Thanet District Council), North Devon County Council, Sutton Harbour Holdings PLC, Mevagissey Harbour Trustees and Gloucester Harbour Trustees. An additional 204 datasets will be made available to cover the Irish Sea and the North East Atlantic.	One project report for the Sound of Mull Remote Sensing Project: interim report for Historic Scotland, 2004 was archived. The illustrated report describes Multibeam sonar surveys of eleven wreck locations, including the Historic Marine Protected Area at Duart Point in the narrow sound of Mull from 2004.
ADS (Historic Environment DAC)	**BODC**
London Gateway Port: Channel Clearance, Dredging and Capital Dredging Project Reporting https://doi.org/10.5284/1083494 Prior to the construction of the London Gateway (LG) Port, a major development on the north bank of the River Thames, extensive marine archaeological investigations were undertaken by Wessex Archaeology (WA). The project area investigated is just over 100 kilometres in length and ranges in width between 360 metres in the outer channel and almost a kilometre wide at the port location. This record relates to the marine archaeological element of the LG Port project, the results of a variety of investigation methods utilised over the span of the project which ran from 2006 until 2016. These included diver observation, geophysical survey and capital dredging watching brief and object reporting. As a result of the various phases of fieldwork over 40 archaeological sites of varying complexities were investigated, over 400 geophysical anomalies were identified, and over a thousand archaeological objects were recovered. Colossus Dive Trail Maintenance and Wheel Wreck Dating https://doi.org/10.5284/1079022 This project undertook maintenance of the Colossus dive trail and the collection of *further* dating evidence from the Wheel Wreck to aid the identification of this wreck. Both sites are protected historic wrecks and are only about six kilometres apart.	During 2020/21, BODC received 285 accessions of data from 99 organisations in 16 countries as follows: 8 accessions from Natural Environment Research Council (NERC) laboratories (not including collaborative centres & National Oceanography Centre (NOC)) 44 accessions from UK universities 4 accessions from UK Government funded laboratories 3 accessions from commercial organisations 93 accessions from charitable organisations (including NOC centres) 133 accessions from overseas laboratories The data comprise physical, chemical, biological and geophysical observations in a variety of forms including profiles, time series and discrete samples. Data sets are prepared using MEDIN guidelines and are loaded into the National Oceanographic Database (either the BODC Series or the BODC Samples database) after reformatting, usage metadata compilation, quality control (automatic tests and visual inspection), documentation and audit. During 2020 - 2021, an additional 79 datasets (14% increase) were added to the Published Data Library (PDL) and received a Digital Object Identifier (DOI). The PDL had 373 active downloads from 395 published datasets.

4 Highlights

In addition to providing metrics, the DAC reports also detail highlights from the previous year, which together show levels of activity, examples of usefulness of the DAC network and indicate how nationally and internationally integrated the DAC system is. A number of new developments and initiatives took place during 2020-21, enhancing the capability of the MEDIN DACs to the benefit of MEDIN's users.

4.1 New developments and capabilities:

Marine Species and Habitats DAC (DASSH): DASSH has undertaken total redesign of database structure to reflect international standards. The PostgreSQL database that underpins the DASSH infrastructure has been rebuilt to improve flexibility and scalability.

The MEDIN Automated Image Management System (MAIMS), although still a pilot study at this stage, has provided the basis for survey images to be submitted alongside a MEDIN guideline and directly accessible online via Resource Space with metadata.

There has also been development of the JNCC funded Quality Assurance Framework tool to provide automated checks and reports on MEDIN guideline-based proformas used for benthic imagery data. This supports the Big Picture Action Plan for harmonised access to marine benthic imagery.

Bathymetry DAC (UKHO): The Seabed Mapping App and Marine Data Portal now have advanced Google Analytics providing significantly improved analytics. This enables UKHO to better understand its users and to provide more robust metrics to MEDIN.

A beta version of the 100m Web Map Service was released, providing users who want a contextual view of the marine environment with an easy to use and integrated surface.

Fisheries DAC (MSS and Cefas)

MSS: A portal for Scotland's Marine Assessment 2020 was launched in December 2020, and includes significant amounts of assessments for Scottish fisheries in terms of productivity, economics, biodiversity, and regional pressures from activities. The full assessment is located on: https://marine.gov.scot/sma/

Cefas: This year Cefas was granted Associated Data Unit (ADU) status within the Intergovernmental Oceanographic Commissions (IOC) marine data programme, the International Oceanographic Data and Information Exchange (IODE). Cooperating with IODE and other ADU members will strengthen Cefas's international work on marine data exchange, enhance knowledge and improve the flows of Cefas data into global databases.

The Cefas Data Hub system, which includes the Cefas Data Portal, has undergone significant redevelopment this year and has reached the final testing phases prior to public re-release. Infrastructure has been migrated to the cloud and new functionality has been developed to improve the user experience for our depositors in particular. Highlights include faster, more accurate search capabilities, a more streamlined front end including more links to our external collaborators and improvements to both digital accessibility and data security.

Historical Environment DAC (ADS, RCAHMW and HES)

ADS: Work has begun on building a new ADS website that meets modern World Wide Web Consortium (W3C) Accessibility Guidelines, and a beta is currently in production. This incorporates a new template for datasets that is more intuitive, and utilises database stored metadata in a more helpful fashion.

Work also continues on evolving the ADS metadata standard. Primarily this is establishing a baseline core metadata framework that incorporates all ADS requirements for provenance, administration, and preservation, and also the requirements of partners such as MEDIN, Advanced Research Infrastructure for Archaeological Dataset Networking in Europe (ARIADNE), Keepers Registry and DataCite. The next stage of this undertaking is rebuilding historic web services and workflows to modernise the (meta)data

sharing capabilities of the ADS – this would include Open Geospatial Consortium (OGC) Web Map Services (WMS), basic download/export options (XML, CSV,JSON) and a series of Application Programme Interfaces (APIs).

RCAHMW: The "Digital Delivery" project was completed in March 2020. This was a major undertaking to replace the entire data platform, using a 'middleware' solution. This comprises an archive catalogue (ATOM) element, a geographical heritage site recording element (ARCHES) and a digital asset management system (iBase). These are brought together using a middleware solution supplied by Knowledge Integration. Selected elements of this platform are made available to the public via a completely new version of our online access system Coflein https://coflein.gov.uk/en/ which includes a fully searchable, Open Geospatial Consortium (OGC) compliant, map front end.

The new Coflein offers a greatly improved access to online content, including digital images, animations, PDF documents, etc. and employs more effective searching of linked data across related site/archive datasets. The change in technology allows us to make use of APIs in the future to link with different online resources, and to repurpose subsets of data for different outputs.

HES: Awareness of the Protected Military Remains Act has been raised through the creation of a Protected Military Remains Act 1986 layer. The 1986 Act lists Designated Vessels and Controlled Sites but, without a recognised data owner, there is no dedicated spatial dataset. HES created a layer from wreck locations recorded in Canmore for display in PastMap. Work commenced in scoping publishing the data as View and Download Services.

UK Research and Innovation (UKRI) Toward a National Collection (TaNC) application - Collaborative work focussed on contributing to the Historic England led UKRI Toward a National Collection Discovery Programme bid: Unpath: To unpath'd waters, undream'd shores. The application progressed through the initial sift in early 2021 and the outcome is expected by the end of September 2021. If successful, Unpath will provide a consistent approach to recording and presenting marine historic environment data across UK waters.

Water Column Oceanography DAC (BODC): A number of new developments have been taking place.

- Exposure of discrete samples data via ERDDAP, a technology that allows users to access and subset data online.
- General Bathymetric Chart of the Oceans (GEBCO) 2020 grid global release – the new platform has seen a large increase in downloads of GEBCO gridded datasets and an increase on sections of the grid.
- Implementation of a BODC archive on JASMIN computing infrastructure to enable archive of high-volume data that can't easily be stored in BODC's systems.
- Natural Environment Research Council (NERC) Vocabulary Service (NVS) – Enhancements to the NVS delivering a powerful and user-friendly User Interface (UI) for the NVS, which will improve user experience (human and machine-to-machine). Work has also taken place to deliver vocabulary services to the new French "Systeme d'Information sur le Milieu Marin", expanding the reputation of the NVS.
- Sea level Zooniverse project UK – over 2000 volunteers have helped to digitise 6 years of historic sea level data at 2 sites, covering more than 2 million data points and filling in gaps in the historic record.

Marine Geology and Geophysics DAC (BGS): BGS has been involved in the Geospatial Commission Coastal zone mapping 2 project, which is a collaborative project led by UKHO with BGS and other

government agencies involved in seabed and coastal mapping. BGS and UKHO were also in other Geospatial Commission projects such as Archive Data Capture, Licencing and Data Discoverability.

Marine Meteorology DAC (Met Office): Due to the coronavirus situation, major projects have been delayed due to different working conditions and restrictions on physical engineering works. Improvements have been made to metadata record keeping and sharing with international partners.

One major development has been a new piece of software which allows the extraction of data from our primary database in the International Maritime Meteorological Tape (IMMT-5 format). This has enhanced our ability to share our Voluntary Observing Ship data internationally and has unlocked a new ability to share our Automatic Observing Ship data through the same channels. This has resulted in an estimated 4 million extra observations being shared with the international community and should increase output by approximately 400,000 per year going forwards.

4.2 New funding streams:

Some of the MEDIN DACs received new funding streams during 2020-2021, in particular the **Historic Environment DAC**:

- **ADS:** has been awarded a contract worth over £1.3 million to secure the digital archives generated by archaeological works in advance of construction of the High Speed 2 (HS2) infrastructure programme.
- **RCAHMW:** is developing a potential bid, in partnership with the National Library of Wales to UK Research and Innovation (UKRI), for infrastructural funding for a national Trusted Digital Repository for Welsh humanities data, which will include historic environment data.
- **HES and ADS:** are involved, along with MEDIN, in a bid led by Historic England for the UKRI TaNC Discovery Programme to provide a consistent approach to recording and presenting marine historic environment data across UK waters.

Other DACs have also secured additional funding, for example **Marine Species and Habitats DAC (DASSH)** has a range of project-based funding and a continuation of EMODnet Biology. The **Water Column Oceanography DAC (BODC)** has secured EU funding in the European Union Horizons 2020 Integrated European Long-Term Ecosystem & socio-ecological Research Infrastructure (H2020 eLTER), and the European Commission European Open Science Cloud (EOSC), as well NERC capital funding for NERC Vocabulary Server (NVS) development work and glider data system developments.

4.3 International meetings

Many of the DACs have a strong presence internationally, reflected in the broad spectrum of international meetings attended. A few examples from this reporting year include:

A variety of EMODnet meetings (e.g. Biology, Chemistry, Geology, Ingestion, Partners meetings) involving **Marine Species and Habitats DAC (DASSH), Fisheries DAC (Cefas), Water Column Oceanography DAC (BODC), Marine Geology and Geophysics DAC (BGS).**

The **Fisheries DAC (MSS and Cefas)** has been involved in International Council for the Exploration of the Sea (ICES) meetings, such as MSS in the ICES Data and Information Group (DIG), and Cefas in the Assessment Working Groups, Data Governance, technical advisory and others.

Cefas and **BODC** attended the 26th Session of the United Nations Educational, Scientific and Cultural

Organisation (UNESCO) Intergovernmental Oceanographic Commission (IOC) International Oceanographic Data and Information Exchange (IODE). They also attended and presented at the International Conference on Marine Data and Information Systems (IMDIS).

The **Historic Environment DAC** has also been involved in international meetings. For example, **RCAHMW** held the annual Digital Past Conference in February, which is an international conference on the use of digital technology in recording, interpreting, and engaging with heritage. Due to Covid, this year the conference was entirely online but this helped to increase the international make-up of both delegates and contributors. **ADS** presented at the Digital Past conference - Introduction to Saving European Archaeology from the Digital Dark Age. ADS also presented at the EU Social Sciences and Humanities Open Cloud (SSHOC) workshop on Considerations for the Vocabulary Platforms - Vocabulary mapping tool for archaeology in ARIADNEplus; and the European Association of Archaeologists on Sustainability, Unsustainability and Opportunity for Archaeological Data.

Marine Meteorology DAC (Met Office) attended the Ships Observations Team (SOT) World Meteorological Organisation Integrated Global Observing System (WIGOS), meeting with international colleagues to discuss changes to unique identifiers and how this may affect the systems that use them, such as Volunteering Observing Ships (VOS).

Other meetings included Germany's National Meteorological Service - the Deutscher Wetterdienst (DWD) to discuss challenges and opportunities facing the Global Data Assembly Centre (GDAC), and discussed data formats at the International Maritime Meteorological Tape (IMMT) Task Team meeting.

4.4 Data Access and Sharing:

Data from most of the MEDIN DACs are being made available under open licences such as the UK Open Government Licence (OGL) for data. The majority of data from MSS, NERC (e.g. BGS and BODC), UKHO, Met Office, and Cefas are made available under this licence.

The ADS have moved from an older form of licence to Creative Commons Attribution (CC-BY) or Creative Commons Attribution non-Commercial (CC-BY-NC) for new datasets. ADS is slowly addressing updating historic licences where depositor/originator can be contacted. For RCAHMW, licensing of datasets is governed by the Re-use of Public Sector Information Policy. For HES, Canmore location data is released under an Open Government Licence. Individual archive items are subject to copyright terms agreed with depositor. DASSH licensing is assigned in consultation with the data provider. All data are made "as open as possible, as closed as necessary". DASSH promote CC-BY or OGL as the preferred license, but in some cases, data are deposited under CC-BY-NC.

The DACs' data access mechanisms are described below, along with improvements made during the reporting year:

Bathymetry DAC (UKHO): Data are available to download and be interrogated through a geospatial viewer and search query. The new portal enables access to more than just bathymetry data held at UKHO. The data are licenced under OGL and are also made available through third party portals such as data.gov.uk, EMODnet, and the International Hydrographic Office (IHO) Data Centre for Digital Bathymetry (DCDB).

UKHO have made over 4000 metadata amendments to its bathymetry datasets and are currently data cleansing their systems to enhance their metadata records. Work continues on updating the data platform so that UKHO can provide MEDIN v3.1 metadata records. This work and the data cleansing is estimated to complete early in 2022.

Marine Species and Habitats DAC (DASSH): Data are published via OGC web services, a map-based, web-accessible query tool and the Integrated Publishing Toolkit using Darwin Core Archive structured data.
The newly developed MEDIN Automated Image Management System (MAIMS), provides a mechanism for directly accessing survey images and metadata.

Data are made available through the following third-party portals National Biodiversity Network Atlas – atlas.nbn.org.uk, EMODnet Biology, EurOBIS, OBIS, GBIF.

Fisheries DAC (Cefas and MSS): MSS: For repeated annual surveys that are internationally coordinated through ICES working groups, the data are submitted to the ICES Database for Trawl Surveys (DATRAS), while metadata is sent to MEDIN with link directly to the DATRAS system. For nationally coordinated surveys or other datasets, data are uploaded to the Marine Scotland Open Data portal and made available with a DOI. Metadata containing the DOI are submitted to MEDIN.

For some aggregated data, spatial resources are made available on Marine Scotland Maps portal (http://maps.marine.gov.scot) and described on the Marine Scotland Information portal (http://marine.gov.scot). For spatial resources originating in or maintained by Marine Scotland and relating to fisheries, metadata are forwarded to MEDIN, with links directly to layers or information pages. Data are made available through third party portals such as ICES DATRAS - https://ices.dk/marine-data/data-portals/Pages/DATRAS.aspx
Scottish Spatial Data Infrastructure - https://www.spatialdata.gov.scot/

Cefas: Datasets are made available for public download from the Cefas Data Portal, with no registration or sign-in required. All data can be downloaded from the website in csv or shapefile format, additional spatial data is available WMS/WFS direct feeds. Public APIs are available to access all metadata and data.

Work continues to rescue and publish legacy datasets whilst maintaining the flow of all publicly funded non-sensitive data to the external portal. Major re-developments are underway to the external portal infrastructure and front end. These developments are currently in testing phase, to be BETA released during financial year 2021/22.

All metadata are automatically exported to MEDIN and data.gov.uk via Web Accessible Folders (WAFs). All metadata that include a DOI are also served to the Defra Shared Services platform. Such third-party data portals direct external users directly to the Cefas Data Portal in order to access the data.

Historical Environment DAC (ADS, HES and RCAHMW):
ADS: Datasets are available for download directly from ADS website. Datasets are normally grouped by collections that reflect the project carried out by the originator, such as a survey or monitoring project. Each collection has a DOI. A number of grey literature reports (155) are disseminated as individual records within an application known as the ADS Library. Each report has its own DOI.
Improvements include all archive and grey literature metadata are now in the new ARIADNE data portal https://ariadne-portal-staging.d4science.org/

HES: Public access is provided through an online portal (Canmore) and map-based search (PastMap). Users may select and download up to 1,000 records in csv or kml formats under an Open Government Licence. HES Public Services are the principal point of contact for bespoke data requests and downloads. They do not distinguish between terrestrial and marine data requests.

HES spatial datasets and services (Historic Marine Protected Areas, Scheduled Monuments (for Intertidal and wrecks), Listed Buildings (for Coastal built heritage) and Canmore – National Record of the Historic Environment), are available to download and access as View and download services from the HES spatial downloads page and through the Scottish Government Spatial.data.gov.scot metadata portal. Records are harvested to data.gov.uk but following Brexit no longer appear on the INSPIRE Geoportal.

Improvements to access include development of the Protected Military Remains Act dataset, and addition of the Canmore dataset to the HES spatial downloads page where users can download a zipped Shapefile or access the dataset through an ATOM feed.

RCAHMW: Our data is mainly accessed via Coflein our online database, and Historic Wales the collaborative historic environment portal for Wales. Data downloads are available on request to the RCAHMW enquiry service and archives can be accessed in a public reading room.

We are partners in the Arts and Humanities Research Council (AHRC) Towards a National Collection (TaNC) bid to develop interoperability of data in maritime heritage, along with ADS and HES and led by Historic England. This is very relevant to MEDIN's heritage DAC and includes all of its member organisations. One of the outputs is hoped to be improved interoperable discovery metadata for the MEDIN Portal.

Water Column Oceanography DAC (BODC): The BODC National Oceanographic Data Bank (NODB) delivery system, https://www.bodc.ac.uk/data/online_delivery/nodb/, now gives access to 136,000 data series, a 5% increase in the number of series available online from last year. Services offer a user choice of a one-click download (for publicly accessible data) or 'online shopping' with a basket and check out mechanism. Data are made available in various data formats under secure access control methodologies which includes user request tracking of auto-downloads.

BODC provides 231 data collection aggregations and 835 cruise collection aggregations. Where appropriate, these discovery metadata records carry a URL within the online resource metadata element that leads directly to the data. The URLs are of two types: 1) If a dataset has an associated Digital Object Identifier (DOI), the URL resolves to a landing page within the BODC Published Data Library that incorporates a one-click download service; 2) For non-DOI datasets, a URL resolves to a pre-filtered search result set in the BODC online web user interface (UI) specific to the data or cruise collection. The UI incorporates a one-click download service.

BODC have made further improvements to access arrangements through further development of ERDDAP instances, sensor web enablement schema and Google schema.org. All metadata records available through the Published Data Library (PDL) have been published with schema.org.

Data are made available through EU third party portals such as SeaDataNet and EMODnet.

Marine Geology and Geophysics DAC (BGS): Data are made available through the Offshore GeoIndex – https://www.bgs.ac.uk/GeoIndex/offshore.htm (also available as a Web Map Service)
SEA Data Portal - https://www.bgs.ac.uk/data/sea/home.html
BGS Deposited Data Search - https://www.bgs.ac.uk/services/ngdc/accessions/index.html?

Improvements to access include

- Ongoing improvements/additions to the Offshore GeoIndex.
- Improvements to data quality to meet Gemini 2.3/MEDIN 3.1 have been identified and partly actioned.
- Improvements to the data publishing process to generate Gemini 2.3/MEDIN 3.1 compliant xml metadata have been identified but not yet actioned due to lack of funding.

Geological maps created from data are incorporated into EMODnet map products and made available through the EMODnet Geology Portal.

Marine Meteorology DAC (Met Office): Datasets are requested through email/telephone enquiry. The request is passed to our Data Provisioning team which then provides a quote and then the information requested. In addition to this, there are data available on our public website for the previous 24 hours for our moored platform data. Furthermore, our Voluntary Observing Ship data (and shortly our shipborne automatic marine observations) are available through International Comprehensive OceanAtmosphere Data Set (ICOADS). Data are also available through the Centre for Environmental Data Analysis (CEDA). Data are shared in real time through the World Meteorological Organisation's (WMO) Global Telecommunications System (GTS).
Regarding improvements to access, plans have been made to improve communication in cases of data failure and platform movement. Additional metadata have been made available as well.

4.5 Data Standards and data quality

Use of MEDIN guidelines by depositors using the MEDIN DACs is variable across the DACs. DASSH and ADS promote the MEDIN guideline formats with their depositors to standardise the data being submitted.

Some DACs, such as Cefas and MSS only receive data from staff in their own organisations. The MEDIN metadata standard is followed, but the use of the guidelines is not widespread.

BODC and BGS receive data from a wide range of originators and in many formats and it is not known whether the MEDIN guidelines are used by their depositors. DACs such as UKHO, RCAMHW, Met Office and HES promote the use of specific standards relevant to their communities.

5 DAC Sustainability and Funding

An important aspect of the DAC network is the assurance of long-term sustainability and continuity of service provision. The MEDIN DAC network achieves this by requiring that the core capability of each DAC is underwritten by an organisation or group of organisations (usually the host organisation) that itself has a business requirement to manage data of a particular theme. This approach forms the backbone of the funding/cost model for the MEDIN DACs (see box below). Current status of the individual DACs is as follows:

- Funding for the **Bathymetry** (UKHO) and **Marine Meteorology** (Met Office) DAC activities have been incorporated into operational plans and are considered part of business as usual. Funding for the Bathymetry DAC is built into future plans at UKHO and more investment is being made each year to ensure continued and increased availability of bathymetry data.

- **Species and Habitats DAC** (DASSH) funding situation is relatively stable, with continuing support from Defra, Scottish Government and MEDIN for the operation of core DAC functions. This is augmented by secured funding from EMODnet Biology (DG-MARE) until 2023. Additional small project funding is sought each year to complement the overall DASSH work plan.

- Funding for the **Water Column Oceanography** (BODC) and **Marine Geology and Geophysics** (BGS) DACs appears secure in the short to medium term with no reductions (although this is not inflationproofed). These two data centres have been through the NERC Data Centre National Capability evaluation and commissioning process 5-year funding cycle (2018-2023). NERC remains committed to data management for the medium and long term.

- **FishDAC:** Cefas operates under a yearly funding cycle and funding is approved to support data management activities in FY2021. For Marine Scotland Science, the funding situation is stable, but with a growing programme of work and often competing demands.

- **Historic Environment DAC:** The ADS 5-year plan currently runs to 2021 and is reviewed by the Management Committee, on which MEDIN is represented. Although the ADS operating environment is likely to become more difficult in the next 3-5 years, the plan to 2021 is robust and commits to the furtherance of ADS aims and objectives and continuance of relationships with existing external partners such as MEDIN. No significant variation to the basic business model is expected for the next 5-year plan (2021-2026), although opportunities for exploring core funding for infrastructure is being investigated via UK Research and Innovation (UKRI).
  The other two components of the **Historic Environment DAC** (Historic Environment Scotland and RCAHMW) are funded through the Scottish and Welsh Governments respectively, which are committed to ensuring that they are properly resourced in the current, short term and mediumlong term. Historic Environment Scotland receives additional revenue from its Commercial and Tourism arm. The impact of Coronavirus pandemic has had a significant impact on income from HES Properties in Care. This is likely to impact on the wider organisation over the next couple of years.

MEDIN DAC Cost Model

The DAC cost model proposed and adopted in November 2010 identifies four aspects of the DAC function: Core Capability, MEDIN Coordination, Additional Archiving, and Data retrieval / distribution, as described below:

Core DAC Capability
"Core" DAC capability includes infrastructure costs and some routine data archiving. It is expected that core DAC funding is provided by organisations with a strategic interest in a national DAC capability for specific data types. MEDIN acts to provide an overview and to consider whether funding of this core capability is secure or at risk.
Funded by the organisation hosting the DAC, or in the case of DASSH by a consortium of organisations.

MEDIN co-ordination
MEDIN acts to ensure common standards and service provision across the MEDIN DAC network. The cost of MEDIN coordination activities is shared between MEDIN Sponsorship funds and the DACs themselves.
Funded by MEDIN Sponsor funds and DACs through in-kind effort

Additional Archive Costs
In the general case, the costs of archiving newly collected data should be funded by the data providers, in the form of one-off fees to the DACs in return for the services provided. This data archiving cost is not currently included in the overall budget of many monitoring and research programmes. *Funded by data suppliers*

Data retrieval / distribution

MEDIN DACs will provide data access to the original data provider at no cost and will manage third party access to data sets according to terms agreed with the data provider. If no constraints are required by the owner, data will be made available to third parties at no cost, beyond any necessary to cover costs of retrieval / provision. *No cost*