Marine Environmental Data and Information Network (MEDIN)

Annual DAC Network Report for 2016-17

'Measure once, use many times'

Summary highlights

MEDIN has established an operational network of seven linked marine data archive centres (DACs) covering bathymetry, fish and shellfish, fisheries, aquaculture and related samples, historic environment, marine geology and geophysics, marine meteorology, marine species and habitats, and water column oceanography. The DACs continue to archive data from MEDIN partner and third party organisations to agreed individual programmes. The 2016-17 DAC annual reports show that:

- The number of data sets held by the DACs is 11,013, an 18% increase on 2015-16, and almost 3 times more than in 2011-12
- 3,252 new datasets were archived in the DACs (twice the number in 2015-16 and over three times more than in 2011-12)
- 5,684,599 requests for data were received by the DACs (a twelvefold increase in the records downloaded from UKHO together with high numbers from HES and RCAHMW has led to a large increase overall in the number of requests. BODC, Cefas and ADS have also had significant increases)
- Royal Commission on the Ancient and Historical Monuments of Wales (RCAHMW) accredited as a component of the Historic Environment DAC

1 Introduction

MEDIN has established an operational network of linked marine data archive centres (DACs) to provide secure long-term storage for marine data. This network provides the capability to upload and retrieve data. Those organisations archiving data at a MEDIN DAC should have free access to their data, and DACs will 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 provide view and download services for data sets covered by INSPIRE.
- To form the first point of call of expertise for the management of marine data.

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

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 up and how do the DACs see their interaction with the portal.

This short document provides a report on the current status of DACs in terms of data sets held and recently uploaded, requests from users for data, and financial outlook. This is a summary of information from the individual DAC reports. These reports are available on request to enquiries@oceannet.org.

2 DAC Listing

There are currently seven DACs in the MEDIN DAC network, as listed in the table below. More details are available on each DAC through links on the DAC web page on the MEDIN website at http://www.oceannet.org/data_submission/index.html. 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. During the year RCAHMW was accredited as a component of the Historic Environment DAC, and the original four DACs (BODC, BGS, DASSH and UKHO) were re-accredited. The Met Office is currently undergoing re-accreditation; the FishDAC components (Cefas and Marine Scotland) are due for re-accreditation this year. Further discussions took place with AFBI, who are considering the process of seeking accreditation as the Northern Ireland component of the FishDAC, and a meeting Historic Environment DAC Working Group was held to discuss accreditation of the Northern Ireland and England components.

Name	Coverage	Contact Information	Web links	MEDIN Status
BODC	Water column Oceanography	enquiries@bodc.ac.uk 0151 795 4884	www.bodc.ac.uk	Accredited 2009; Re-accredited 2017; operational.
British Geological Survey	Marine geoscientific data	medin@bgs.ac.uk	www.bgs.ac.uk/services/ngdc/management/marine/home.html	Accredited 2009; Re-accredited 2017; operational.
DASSH	Marine Species and Habitats	Dassh.enquiries@mba.ac.uk 01752 633291	www.dassh.ac.uk	Accredited 2009; Re-accredited 2017; operational.
Met Office	Marine Meteorological Data	enquiries@metoffice.gov.uk	www.metoffice.gov.uk	Accredited (Dec 2011); Re-accreditation in progress; operational.
United Kingdom Hydrographic Office	Bathymetry	bathy.dac@ukho.gov.uk	www.gov.uk/guidance/inspire-portal-and-medin-bathymetry-data-archive-centre	Accredited 2009; Re-accredited 2017; operational.
FishDAC (Cefas, Marine Scotland Science)	Fisheries and Shellfish, Fisheries, Aquaculture and related samples	Cefas: data.manager@cefas.co.uk	http://www.cefas.defra.gov.uk/publications-and-data/fishdac.aspx	Accredited 2012, Re-accreditation due 2017; 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-accreditation due 2017; operational.
Historic Environment DAC	Marine Historic Environment fieldwork derived datasets	Archaeological Data Service: help@archaeologydataservice.ac.uk	http://archaeologydataservice.ac.uk	Accredited 2012; Re-accreditation due 2017; operational
		Historical Environment Scotland: peter.mckeague@hes.scot	www.canmore.org.uk	Accredited May 2014; 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 we ask the DACs to report on their performance based on a standard set of metrics which include the numbers of data sets held, the number of new data sets archived, the number of requests for data and the number of MSCC partners who have data archived in the DAC. The table below gives the figures from 2011-12 through to 2016-17.

Please note that it is difficult 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 5 data sets, which are augmented each year with that year's new data. Thus over 2 million observations were added to the Met Office's five data sets during 2016-17.

3.1 DAC Metrics

DAC	Oceano-graphy	Marine Geoscience	Species & habitats	Marine Met.	Bathymetry	FishDAC		Historic Environment DAC		
Year	BODC	BGS	DASSH	Met Office	UKHO	Cefas	Marine Scotland	ADS	HES	RCAHMW
No of data sets held										
2011-12	916	533	1592	4	650					
2012-13	983	675	1973	4	650			112		
2013-14	983	768	2438	4	1409	46	36	126	1	
2014-15	1008	864	2622	4	3815	58	56	139	1	
2015-16	1027	1021	2897	4	4098	63	91	140	1	
2016-17	1045	1051	3215	5[1]	4224	1245	162	141	3	1[2]
New data sets archived										
2011-12	237	16	378	0	128					
2012-13	240	77	20	0	12			9		
2013-14	218	75	70	0	63	5	16	3	609	
2014-15	254	4182	6	0	91	7	9	0	798	
2015-16	287	539	6	0	211	6	10	1	434	
2016-17	290	1500	55	1	158	1172	11	3	51	11
No. of Requests for Data										
2011-12	83,594	- ~100	111,490	NA	0					
2012-13	72,205	-	113,852[3]	NA	16			17,170		

[1] Over 2 million observations added to the data sets in 2016-17

[2] One data set with 7670 individual maritime site records (mainly wreck sites)

[3] Calculated differently from the first year – based on requests for data received through NBN. Statistics are provided for calendar year, 2011-12 values is annual total for 2012, later years are calculated pro-rata across the financial year.

		~100								
2013-14	115,626	6,600 ~100	272,862	NA	84,000	NA	NA	26,501	NA	
2014-15	85,041	7,200 70	581,212	NA	85,000	NA	NA	31,926	NA	
2015-16	129,398	19,980[4] 70	2,355,054	NA	85,000	NA	NA	61,711	NA	
2016-17	177,282[5]	30,876 41[6]	2,598,148[7]	NA	1,000,000	842	NA	81,356	684,477	1,095,625[8]
No of MSSC partners with data in DAC										
2011-12	4	8	8	1	3	0				
2012-13	14	8	8	1	3	0		1		
2013-14	15	8	8	1	9	1	1	1	0	
2014-15	15	11	8	1	9	1	1	1	0	
2015-16	14	11	8	1	9	1	1	1	0	
2016-17	14	11	15	1	9	1	1	1	0	0

Figures 1-3 below provide graphical representations of the changes in each of the DACs and DAC components for several metrics. All DACs showed an increase in data holdings, with the Met Office and BODC adding substantial numbers of observations to existing datasets, and the rest of the DACs reporting an increase in the overall number of datasets held. The number of data sets added to the DACs by year shows a complicated picture, which partly reflects variation in funding available for the DACs to archive new data sets and how the DAC operates (as noted above the Met Office continually adds data to existing databases/data sets; BODC receives data accessions from a variety of projects and programmes which are integrated into existing data sets/ databases).

[4] Web map / WMS requests – number of users per year; new web service introduced in 2015-16

[5] Excludes 7,000,413 requests to the NERC Vocabulary Server (NVS2.0)

[6] Based on manual / email enquiries

[7] Refers to number of records downloaded (number of downloads is 1685)

[8] Text and map-based searches; stats do not distinguish between maritime or terrestrial sites

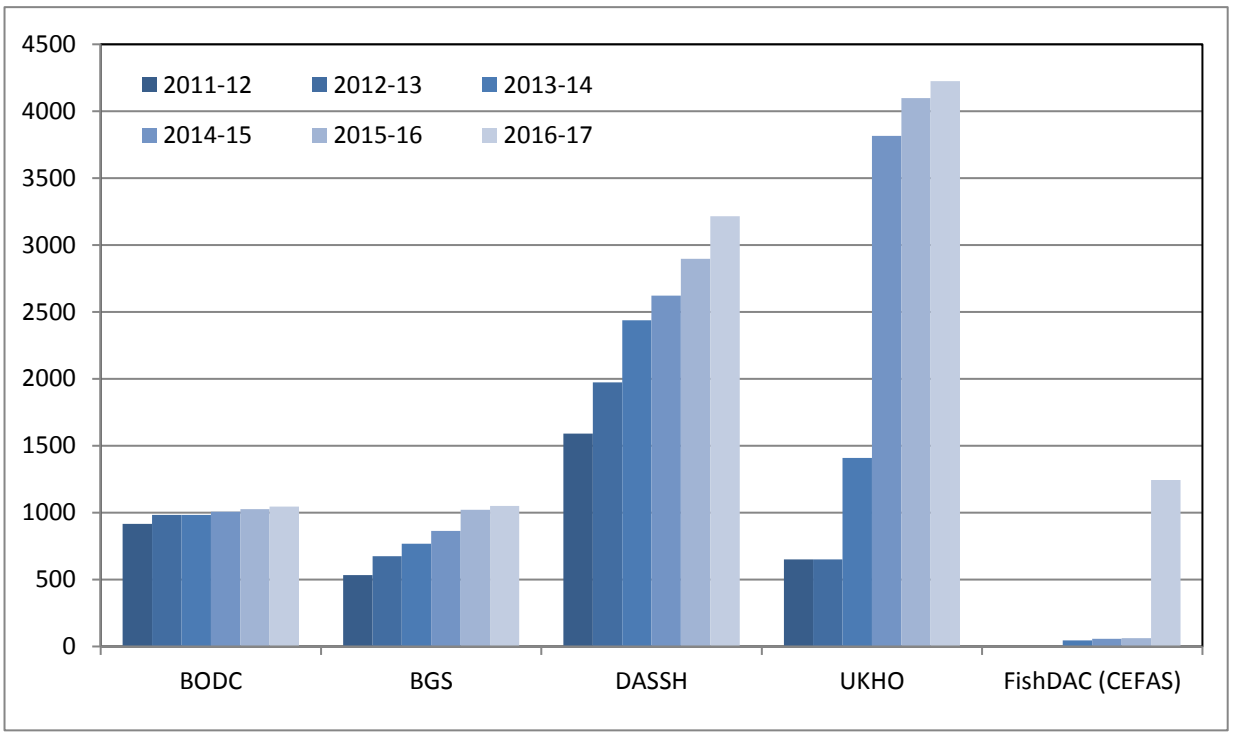


Figure 1a: No. of data sets held by DAC by year

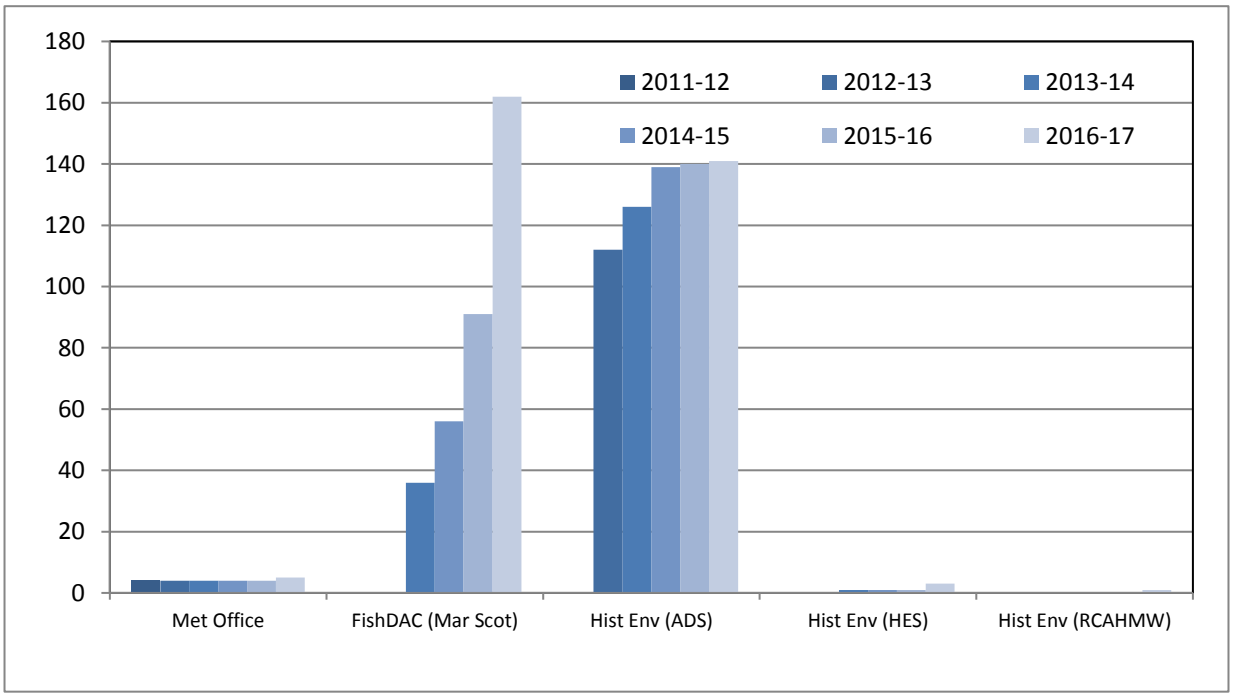


Figure 1b: No. of data sets held by DAC by year

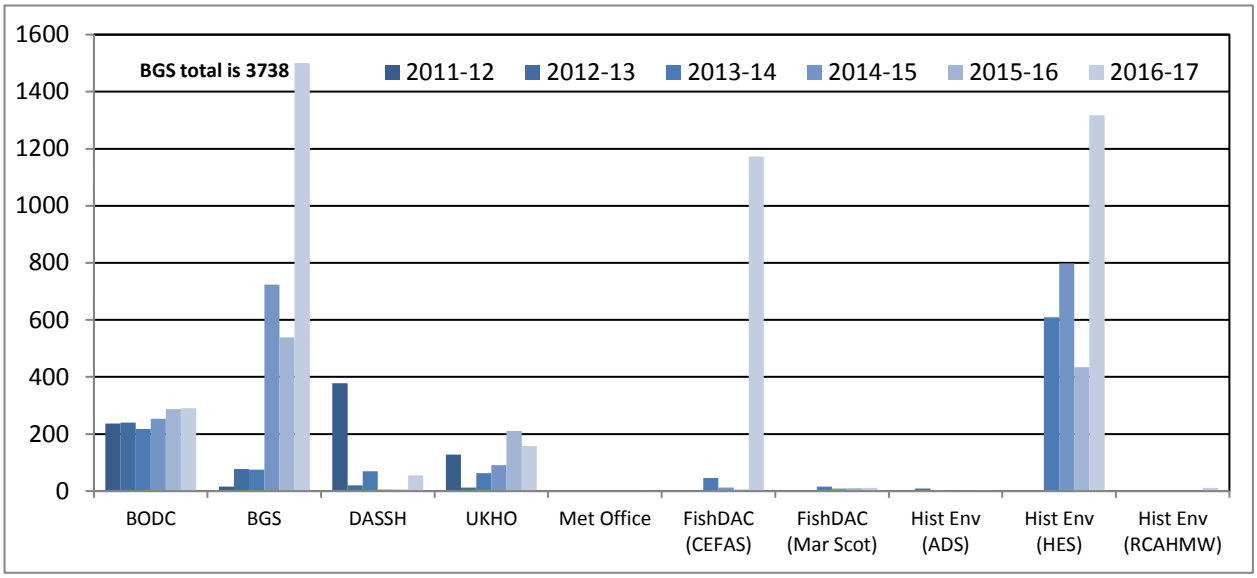


Figure 2: No. of new data sets archived by DAC and by year

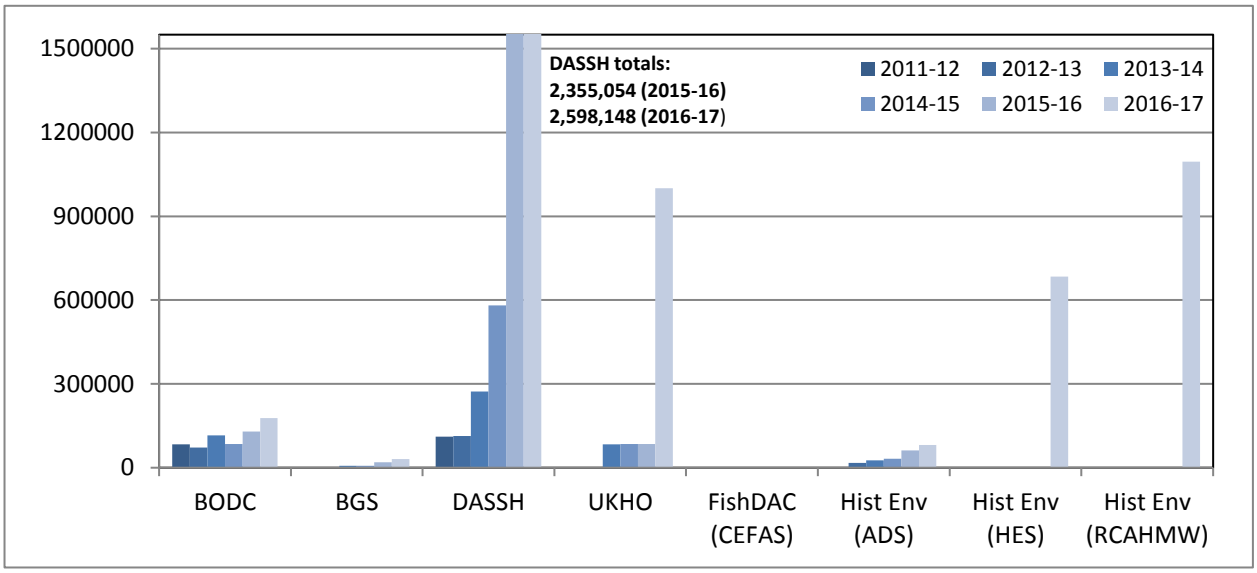
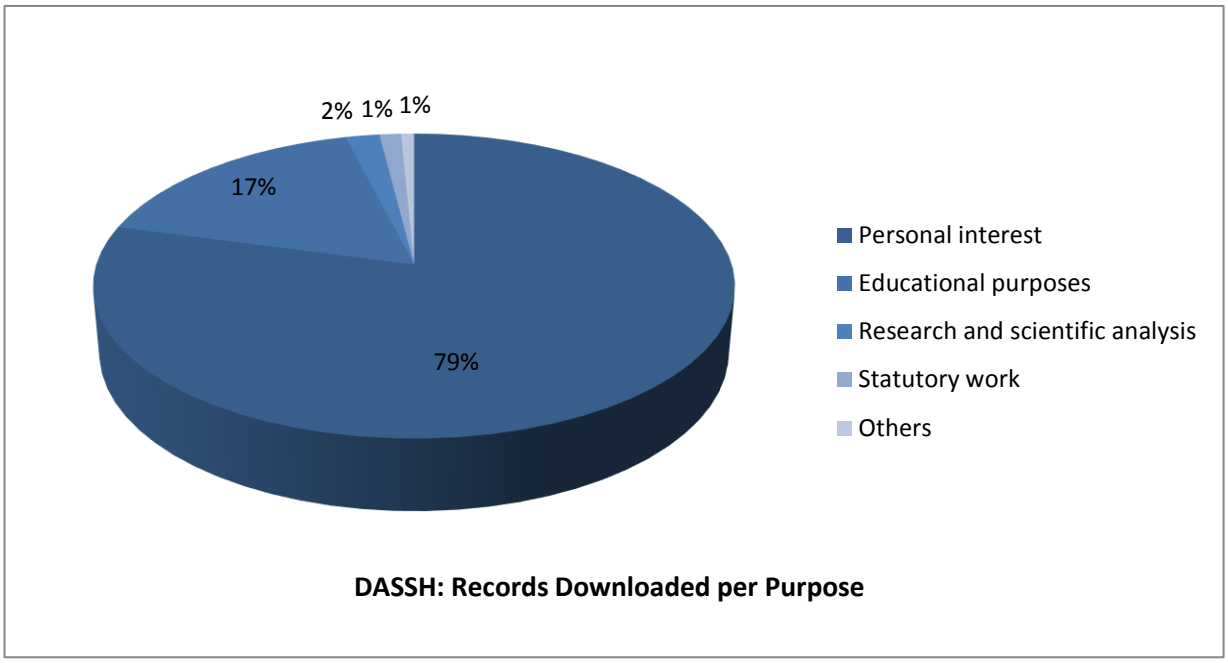


Figure 3: No. of requests for data (by year) from those DACs that record this information



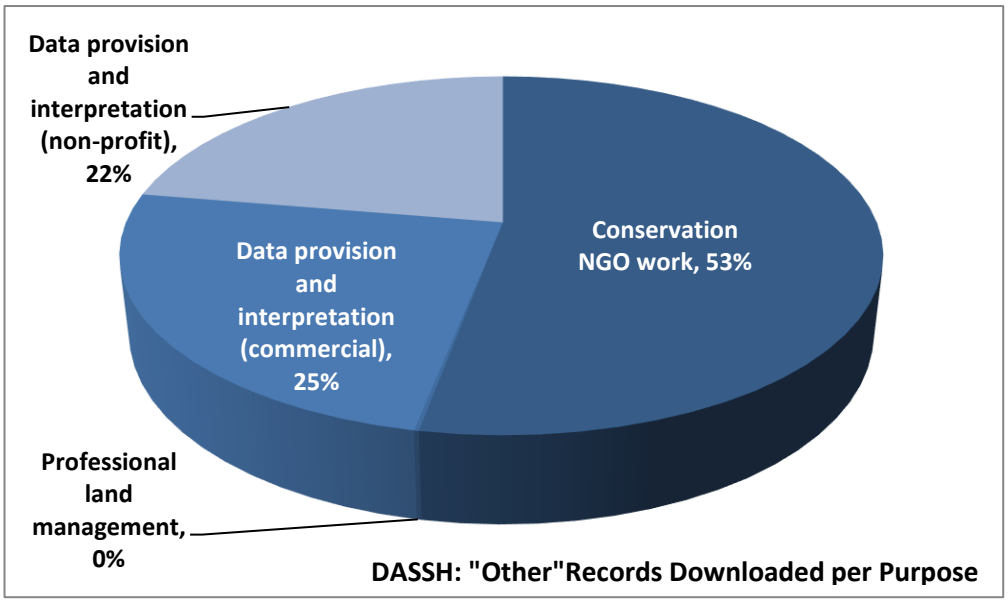


Figure 4: Charts showing distribution of purpose for records downloaded from DASSH

3.2 New datasets:

The past year has been very busy for the DACs in terms of major new datasets being archived. Table 1 summarises new datasets archived by each DAC during FY 2016/17.

BODC

During 2016/17, BODC received 290 accessions of data from 84 organisations in 18 countries as follows:

93 accessions from NERC laboratories (including collaborative centres & NOC)
51 accessions from UK universities
3 accession from UK Government funded laboratories
0 from commercial organisations
34 from charitable organisations
109 accessions from overseas laboratories

Additionally, 45 datasets were added to the Published Data Library and received a DOI.

BGS

- MCZ data – 20 surveys from 10 MCZ sites.
- MCA backscatter data - 122 HI surveys (these were received over the course of the past few years rather than just 2016-17)
- 541 sampling activities added to the BGS Coastal & Marine Database
- 211 geophysical lines added to the BGS Coastal & Marine Database
- BGS geophysical records scanned - 1625 (this includes parts, so it's larger than the number of lines these records represent)
- Tarmac records - 14 out of 44 boxes scanned (1219 geophysical lines and charts and reports)

UKHO

A total of 158 new datasets were made available

- 42 CHP surfaces including 3 by the Royal Navy
- 3 Royal Navy surfaces
- 21 MoU surfaces
- 92 Third Party surfaces

DASSH

55 new datasets since April 2016 with view and download services (including 5 geological datasets, EA Shad surveys 1990-2015, MCZ surveys, The Marine Fauna of Lundy and various Bioblitz surveys)

Met Office

- Following operational acceptance of the ship-borne automatic weather station (AMOS) network we were able to successfully create a new metadata record for this network
- Summary of data sets archived (in the last year) – All 5 Met Office datasets have been added to over the past year, adding in excess of 2 million observations.
- WMO standards and Guidelines are used.

ADS (Historic Environment DAC)

- Determining Potential: Onshore/Offshore Prehistory
  DOI: 10.5284/1041582
- Twentieth Century Naval Dockyards Devonport and Portsmouth: Characterisation Project
  DOI: 10.5284/1038393
- England's North Sea Ports
  DOI: 10.5284/1040801

RCAHMW (Historic Environment DAC)

447 items relating to maritime sites were catalogued, and 11 additional maritime sites were added.

A major archaeological survey collection, relating to the Resurgam designated submarine wreck, was archived during the year.

HES (Historic Environment DAC)

227 items, relating to 51 maritime records were catalogued. Several of the items archived are database tables or project reports linked to more than one site record in Canmore.

maritime summary		
PREFIX	**Count of archive**	**Description**
551	3	Fonds level record
B	2	Newscuttings
DP	170	Digital Images
DT	1	Database
MV	6	Digital tape

CEFAS (FishDAC)

- 1172 new data holdings
- Fishing survey and Fisheries Science Partnership datasets

Marine Scotland Science (FishDAC)

- North Sea International Bottom Trawl Survey Quarter 1 + 3
- West Coast International Bottom Trawl survey Quarter 1 +4
- Deep water survey, West Scotland
- Herring Acoustic survey data
- Mackerel Acoustic survey data
- Industry led survey for herring assessment data, ICES are VIA
- Monkfish surveys: 3 surveys in total
- Rockall Survey, quarter 3.
- Salmon and Sea trout catch statistics 1952 – 2016
- Effects of Lift bag on Nephrops selectivity research data
- Fish and Shellfish Stocks 2016

WP	44	Word Processed Documents / Pdfs

One major survey collection 'Clyde Built' undertaken by Cotswold Archaeology from April 2015 was archived.

Collection URL	No of items	Description
1524544	1	Archaeological archive relating to the project Clyde-Built: A study of the Clyde Shipbuilding Industry as Represented by Shipwrecks within the Clyde River and Estuary (Fonds level).
1525383	43	Diver event logs from the project Clyde-Built: A study of the Clyde Shipbuilding Industry as Represented by Shipwrecks within the Clyde River and Estuary
1525384	158	Diver photographs from the project Clyde-Built: A study of the Clyde Shipbuilding Industry as Represented by Shipwrecks within the Clyde River and Estuary

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 indicates how nationally and internationally integrated the DAC system is.

4.1 Partnerships:

The MEDIN DACs have established a wide range of national and international partnerships, with BODC increasingly making data available for searching from the SeaDataNet portal (www.seadatanet.org), and DASSH working closely with the National Biodiversity Network (NBN, data.nbn.org.uk) providing data to the NBN Gateway and onward to the Global Biodiversity Information Facility (GBIF) and the Ocean Biogeographic Information System (OBIS, http://www.iobis.org/). BGS data are available for download via the Geo-Seas portal (www.geo-seas.eu/), and Cefas and Marine Scotland Science use the ICES DATRAS portal (http://www.ices.dk/marine-data/data-portals/Pages/DATRAS.aspx) for a number of key surveys. In addition, a number of the DACs are partners in the EMODnet thematic portal projects. BGS are leading the Geology theme and participating in the Bathymetry theme; BODC is a partner in the Physics, Chemistry and Bathymetry themes and DASSH/MBA is a partner in the Biology portal. In addition, bathymetry data from the UKHO is included in the Bathymetry theme, near-real time data from the Met Office is included in the Physics theme and data from MERMAN is included in the Chemistry theme (see the EMODnet web site for further details of the data available and links to the thematic portals at: www.edmodnet.eu).

4.2 Data Access and Sharing:

Increasingly data from the MEDIN DACs are being made available under the Open Government Licence (OGL) for data. Data from NERC (e.g. BGS and BODC), UKHO, Met Office, and the bulk of data from Cefas are now made available under this licence.

Additionally, access to data held by Marine Scotland and ADS is free, open and online, and where possible data held by DASSH are made freely available.

INSPIRE compliance is a key component of MEDIN, and a core responsibility of the DACs. Work to ensure compliance continues at the DACs. The current status is shown in the Table 2.

4.3 Highlights from the DAC network

In addition to providing the metrics summarised in section 3 above, the DAC reports also detail highlights from the previous year, which together show levels of activity and innovation and provide examples of how the DACs are working together to improve data access. Selected highlights from each of the seven MEDIN DACs are given below:

Bathymetry DAC (UK Hydrographic Office)

- Approx. 1,000,000 surface downloads from the portal – almost a twelve-fold increase on the previous three years.

Fisheries DAC (Cefas and Marine Scotland Science (MSS))

- **MSS** is increasingly publishing open data on its own data portal (http://data.marine.gov.scot) with minted DOIs. 133 open datasets have been released with minted DOIs (24 of which overlap with FishDAC function). The data portal supports API access, JSON and RDF downloads of datasets (so generally speaking 4 star open data).
- **Cefas** participated in the Defra 8000 exercise where 1840 datasets were published of which 1000 were FishDAC;
- FishDAC datasets, including 1172 new data holdings, are currently served through the Cefas Data Hub with view and download capability;
- Fishing survey and Fisheries Science Partnership datasets made available (with MEDIN small data archiving project funding).

Historic Environment DAC (ADS, HES and RCAHMW)

- ADS, HES and RCAHMW have been working together to encourage further participation in the Historic Environment Federated DAC by establishing relationships with DoENI and Historic England, including a Historic Environment DAC meeting in Belfast this year.
- **ADS:** Number of referrals from MEDIN Portal: 21 visits. This is the first year these statistics have been provided as the software used only records top 50 referrers. MEDIN has now entered the top 50. Although referrals are modest, the time spent on the site and low bounce rate (the percentage of visits that only had a single page view, meaning that the visitor left the website directly from the entrance page) indicate that users are finding relevant content.
- **HES:** Work is ongoing in integrating the Preservica software which when testing is complete will help fulfil the long-term archiving of digital material and help HES achieve status as a Trusted Digital Repository.
- Major survey collection 'Clyde Built' undertaken by Cotswold Archaeology from April 2015 archived.
- **RCAHMW** accredited as a component of the DAC.
- A major archaeological survey collection, relating to the Resurgam designated submarine wreck, was archived during the year.
- Secured major European funding for archaeological recording on headlands and islands of the Irish Sea, including some maritime recording for the next 5 years (CHERISH).

Marine Geology and Geophysics DAC (BGS)

Direct access to data has been improved through upgrades to the BGS Offshore GeoIndex (http://www.bgs.ac.uk/GeoIndex/offshore.htm) which now includes survey polygons and a URL to go directly to the data for that survey or site survey and zip file data delivery. The MEDIN metadata will be updated to include this rather than the current general link to the Offshore GeoIndex. An Accession Delivery System is also being developed to provide access to data that have not yet been added to the Offshore GeoIndex. Some of this work has been undertaken with MEDIN "Direct Access to Data" funding.

Marine Meteorology DAC (Met Office)

Operational acceptance of the Autonomous Marine Observing System (AMOS) network – automatic weather stations designed for ships – has led to a new data set and associated discovery metadata record.

Marine Species and Habitats DAC (DASSH)

Development of an automated DOI minting tool to enable, at the click of a button, DOIs to be minted and assigned to datasets. Mapping the DataCite DOI metadata standard mandatory elements to the MEDIN discovery metadata standard enables a workflow with minimal human interaction, ensuring the rapid generation and linking of DOI's to the data available through the DASSH portal. Undertaken with MEDIN "Direct Access to Data" funding.

Water Column Oceanography DAC (BODC)

96% of BODC's discovery metadata records in the MEDIN Portal now provide direct access to the data and research cruise collections held. To achieve this, BODC has reviewed and conditioned ~150,000 legacy data series loaded in its National Oceanographic Database (NODB) to satisfy the technical requirements to allow direct access services. Undertaken with MEDIN "Direct Access to Data" funding.

Two development projects completed:

- Delivery of data *via* Linked Open Data (LOD). Data now available for machine-to-machine search and access via a SPARQL endpoint located at http://linked.bodc.ac.uk/ . User documentation accessible at http://linked.bodc.ac.uk/documentation/.
- Developed a prototype Sensor Observation Service (SOS) Sensor Web Enablement (SWE) approach for data delivery.

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.

- **Species and Habitats DAC** (DASSH) receives rolling year-on-year funding from Defra and Scottish Government, assessed against a series of Key Performance Indicators. Additional "project" based funding comes primarily from Statutory Nature Conservation Bodies, with strategically-aligned EU project funding also contributing to the long-term sustainability.

- Funding for the **Water-column Oceanography** (BODC) and **Marine Geosciences** (BGS) DACs appears secure in the short to medium term with no reductions (although this is not inflation proofed). Beyond that, it is dependent on the outcome of the NERC Data Centre National Capability evaluation and commissioning process that is continuing into 2017/18. Nevertheless, NERC remains committed to Data Management for the medium and long-term.

- There has been some reduction of general running cost budgets year on year within the two components of the **FishDAC** (Cefas and Marine Scotland Science) as part of government spending reductions. For Cefas, future data management will be included in project management processes for the contracts. For Marine Scotland Science MEDIN obligations are being met in terms of routine archiving of data as well as involvement in work streams and meetings, and this work is considered sustainable in the long-term.

- Historic Environment DAC: The **ADS** business model includes the continuous development of the ADS Preservation Legacy Fund. Current funding forecasts show reasonable levels of predicted income for the next 3-5 years i.e. enabling current levels of staffing to be maintained (and increased) over this period. The ADS finances and budget forecasts are reviewed by the management board annually. The ADS host institution, the University of York is committed to the long-term preservation and access of ADS holdings.

- 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 ensure that we are properly resourced in the current, short term and medium-long term.

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 organisations.

MEDIN coordination
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