

Process for becoming a Marine Environmental Data and Information Network (MEDIN) Data Archiving Centre

1 Introduction

A key objective of MEDIN is to coordinate 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. Data contributors have free access to their data managed within the DAC framework.

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 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 of expertise for the management of marine data.
- Once accredited, DACs must provide annual reports for the MEDIN Sponsors.

In 2019, MEDIN Sponsors and existing DACs agreed to adopt Core Trust Seal (CTS) accreditation in place of the previously used MEDIN accreditation, recognising the organisational and reputational benefits such international and cross-discipline accreditation will give them. From 2020 MEDIN initiated a phased transfer from MEDIN accreditation to CTS accreditation so that, within the timescale of the new MEDIN Business Plan (2019-24), all MEDIN DACs will move across to CTS accreditation.

This document describes MEDIN's established procedure to govern the process by which new Data Archive Centres are included in the network. When an organisation wishes to apply to become a MEDIN DAC, or an existing DAC wishes to maintain its MEDIN DAC status, the organisation must:

- achieve CTS accreditation
- publish metadata of their data holdings on the MEDIN portal
- be active members of the MEDIN DAC Working Group, supporting delivery of MEDIN's annual Work Plan
- provide annual reports for the MEDIN Sponsors (once accredited)

Accreditation Process

There are four stages to the accreditation process, finishing with formal approval by the MEDIN Executive Team:

- *Initiation / Preparation.*
- *Application for Core Trust Seal (note that the timescale for reviewing CTS applications is outside MEDIN's control and can take several months).*
- *Publication of metadata on the MEDIN portal.*

- *Approval by MEDIN Executive Team.*

The first, preparation, stage can take up to several years. The second and third stages can be completed concurrently and may take up to 12-18 months before the final approval by the MEDIN Executive Team.

Once accredited, the status and performance of DACs will be reviewed annually as part of MEDIN's annual review process.

1. Initiation / Preparation

Involvement: MEDIN DAC working group, expert stakeholders and organisation proposing to host a DAC.

Description: The MEDIN DAC Working Group identifies the need for a further Data Archive Centre within the MEDIN network. Working with interested parties the DAC Working Group proposes an outline scope, remit and *modus operandi* for the new DAC.

Duration: This part of the process can take between 6 months and 2 years, as it requires a consensus to be established between interested parties, and perhaps business plans to be developed.

2. Application for Core Trust Seal accreditation

Involvement: DAC host organisation

Description: The organisation applying to become a MEDIN DAC provides a detailed response to the list of sixteen [Core Trust Seal requirements](#) and applies for CTS accreditation. The MEDIN DAC Working Group can advise and support the organisation applying for CTS accreditation by sharing experiences, reviewing documentation and providing a sounding board. Core Trust Seal provides a peer review process for the applications. Note that Core Trust Seal charges an administration fee upon submission of accreditation application. This should be paid by the organisation applying for accreditation, but MEDIN can reimburse this charge on request. The DAC host organisation should inform the MEDIN DAC Working Group when they successfully achieve CTS (re)accreditation.

Duration: 12-18 months

3. Publication of metadata on MEDIN portal

Involvement: DAC host organisation

Description: The organisation applying to become a MEDIN DAC publishes its marine metadata holdings on the MEDIN portal. The MEDIN helpdesk can support this process and MEDIN runs regular workshops to train individuals or organisations to create MEDIN-compliant metadata to feed into the portal.

Duration: 12-18 months

4. Approval by MEDIN Executive Team

Involvement: MEDIN Executive Team

Description: The Executive Team considers the organisation's application to become (or renew its status as) a MEDIN DAC and:

- (a) Confirms approval of the DAC.
- (b) Confirms approval of the DAC but recommends specific actions to be taken by the DAC to meet requirements.
- (c) Postpones approval of the DAC until specific actions are taken.
- (d) Recommends that an alternative solution be found to provide data archiving facilities for the data categories under consideration.



Duration: At MEDIN Executive Team Quarterly Meeting

Annex 1: Marine Environmental Data and Information Network: Previous requirements for accreditation as a MEDIN Data Archive Centre prior to adoption of Core Trust Seal accreditation

This document lists the previous requirements for an organisation to become a Data Archive Centre (DAC) under the now obsolete Marine Environmental Data & Information Network (MEDIN) accreditation. It provides further explanatory information for each of these requirements. This information is included here as reference material in case it provides useful prompts for organisations applying for Core Trust Seal accreditation.

Requirement	To be accredited DACs must provide
ORGANISATIONAL FRAMEWORK	
Generally exhibiting evidence of expertise and a track record in the scientific area of the data	<p>DACs should describe the range and length of expertise of both the organisation and their staff.</p> <p>In addition, details of data sets or products available can also be provided.</p> <p>Any appropriate affiliations (e.g. national or international bodies, etc.) should also be noted.</p>
<i>DAC Response</i>	
Committed to provide sufficient resources for defined period and plans for transition if and when it ends	<p>In order to be accredited, a DAC must provide evidence that it is hosted by a recognised institution (ensuring long-term stability and sustainability) and that it has sufficient funding, including staff resources, IT resources and a budget for attending meetings, ideally for a 3 to 5-year period, and this information should be updated regularly.</p>
<i>DAC Response</i>	
Committed to return of data holdings to originators, or lodging with an alternative and suitable repository, if the DAC becomes unsustainable	<p>A long-term stewardship plan should be available including:</p> <ul style="list-style-type: none"> • A statement on how the DAC is financed and for how long. • Action that will be taken if the DAC becomes unsustainable.
<i>DAC Response</i>	
Provide annual report as specified by MEDIN	<p>Accredited DACs should provide an annual report to MEDIN according to the pro forma provided by MEDIN. The report comprises 4 sections as follows:</p> <ul style="list-style-type: none"> • A short summary of the remit and status of the DAC • An overview of activities and developments in the reporting year • Key Targets for the next reporting year • Report any changes against the specific MEDIN DAC requirements (in particular referring to any requirements placed as a condition of accreditation) <p>Other suggestions for future reports might include:</p> <ul style="list-style-type: none"> • Key Performance Indicators • Statement on readiness for INSPIRE compliance
<i>DAC Response</i>	

QUALITY CONTROL AND MAINTENANCE	
<p>Adherence to MEDIN Discovery Metadata Standard and appropriate international principles</p>	<p>MEDIN DACs need to provide evidence of adherence to these principles. Further information and links are given below.</p> <p>The MEDIN Metadata Discovery Standard must be used to record details of data sets. The fields used in the standard are compliant with other international conventions (INSPIRE, ISO19115), which means that the details can be transferred easily between organisations and queried by the MEDIN portal. The Metadata Discovery Standard also conforms to the GEMINI2 profile. Publication of metadata in the MEDIN Metadata Discovery Standard and availability to the MEDIN Discovery Portal meets both INSPIRE compliance and UK Location Programme requirements for discovery services.</p> <p>ISO 19115 (Geographic Information - Metadata) is an international standard that sets out metadata fields for describing spatial information datasets. ISO 19139 (Geographic Information - Metadata - XML schema implementation) is the standard that aims to define an XML encoding for the metadata elements defined in ISO 19115.</p> <p>The UK GEMINI Discovery Metadata Standard is a defined element set for describing geo-spatial, discovery-level metadata within the United Kingdom. It is derived from, and therefore compliant with, ISO 19115 Geographic Information – Metadata and the UK eGovernment Metadata Standard (eGMS). GEMINI was originated by the Association for Geographic Information and is currently being revised to produce GEMINI 2.</p> <p>A number of tools and documents to assist in creating MEDIN-compliant metadata are available from the Standards section of the MEDIN website.</p>
<p><i>DAC Response</i></p>	
<p>Data collection according to defined quality principles and accepted procedures</p>	<p>MEDIN DACs need to provide evidence of defined quality principles and procedures.</p> <p>DACs may also be able to advise on data collection procedures and should be able to direct data collecting organisations to appropriate standards, where these exist.</p> <p>MEDIN is also in the process of deriving data guidelines comprising requirements as to what must be recorded when data of a certain theme is being collected. This allows easier reuse of the data in the future. For example, if benthic invertebrate samples are collected, the instrument used to sample, the sieve size and taxonomic list used to record species should also be stated and use common lists of terms. MEDIN-approved data guidelines are available from the standards pages of the MEDIN website. Where MEDIN data guidelines do not already exist, it is recommended that the resources available on the other marine data standards web pages should be used.</p> <p>Provision of advice and feedback to the original data collectors is valuable, covering information to be recorded alongside data, established quality assurance procedures to be used, etc.</p>

<i>DAC Response</i>	
<p>Quality assurance of the collected data</p>	<p>MEDIN DACs should provide summaries of any quality assurance processes and algorithms that are in place. This should not be a detailed description of how the algorithms work but a broad summary of the checks that are run and, for example, whether data are visually inspected. The summary should include details of how any issues are resolved (e.g. are they returned to the data provider for rectification, fixed by the DAC, noted by quality flags in the data file and/or included in the accompanying metadata).</p> <p>In addition, details of any Quality Management System (QMS) or accreditation schemes implemented by the DAC should be provided. Where data have been collected in line with nationally or internationally agreed standards this should be indicated. For example:</p> <ul style="list-style-type: none"> • Quality Assurance of Information for Marine Environmental Monitoring in Europe (QUASIMEME) • Biological Effects Quality Assurance in Monitoring Programmes (BEQUALM) • National Marine Biological Analytical Quality Control Scheme (NMBAQC) • ISO9000 accreditation • Data collected to internationally agreed standards within major scientific projects (e.g. JGOFS, WOCE protocols and standards) <p>Where guidelines and standards are in use these should be mentioned. For example, the ICES Working Group on Marine Data Management has developed a series of "Data Type" guidelines, which have been designed to describe the elements of data and metadata important to the ocean research community. These guidelines are targeted toward physical-chemical-biological data types collected on oceanographic research vessel cruises.</p>
<i>DAC Response</i>	
<p>Committed to advising third party organisations collecting similar types of data on procedures, and providing data-banking (warehousing) and curation facilities for such similar data from other sources</p>	<p>Short description of DAC</p> <ul style="list-style-type: none"> • Short description of the remit of the DAC including the data types held and those accepted from external parties for archiving. • Licensing terms • Standard agreements covering: <ul style="list-style-type: none"> • Transfer of a copy of data to a DAC • Transfer of ownership to DAC • Use of the data held by DAC by external users <p>Format requirements</p> <ul style="list-style-type: none"> • Note that these are aspirational for new data being collected that needs to be submitted to a DAC. It is not intended that all historical data would need to be converted to these formats before acceptance by the DAC. Historical data needs to be addressed on a case-by-case basis. • At least one, but potentially more, format(s) that data can be submitted to the DAC. • Details of the process for establishing or agreeing alternative formats.

	<ul style="list-style-type: none"> The format description would need to cover both format and syntax. <p>It may be advantageous for the provider to submit data in their own format provided this is properly documented perhaps along with some sort of index of the data.</p>
<i>DAC Response</i>	
TECHNICAL INFRASTRUCTURE	
Databasing and banking with appropriate metadata standards	<p>MEDIN DACs should provide documentation of their working practice and procedures. This should include:</p> <p>Information on the technical metadata for all holdings.</p> <ul style="list-style-type: none"> Descriptions of the data structures (both entities and attributes) within which the data are stored Explanations of any lookups not obvious from the data holdings directly Locations of data holdings on the network or other physical locations Information on metadata schemes Editorial advice on the content expected in each mandatory field of ISO xxx List of any topic-specific additional fields and accompanying editorial guidance Information on georeferencing standards in use
<i>DAC Response</i>	
Auditable process for long-term custodianship and updating of data sets, with appropriate disaster planning	<p>MEDIN DACs should have a security policy describing how the data holdings are protected from both malicious and accidental loss. Note that the security policy should exist but should not be made public as it potentially exposes vulnerabilities. A policy should include the following:</p> <ul style="list-style-type: none"> How the holdings are physically protected (e.g. how access to the building is controlled, how secure the building is, who has access) Access to the network (if the holdings are accessible from the network) – what is the access policy, how is user access limited and by who, whether there is an internet link and details of how the firewall is configured and altered, how machines are patched, which users can log on to particular machines, policy on passwords (e.g. how often they are changed and how secure they need to be) Policy when staff leave the organisation Database policy – how users are established, what rights they have, how often administrator passwords are changed, what control is there over allowable passwords <p>How the data holdings are backed up – how often, where are the backups stored and how long for, how protected are the backups (e.g. fireproof safe, stored securely off site, who has access)</p>
<i>DAC Response</i>	
USER ACCESS AND COMMUNICATION	
Committed to, and focus on, customer service	<p>DACs should provide information on:</p> <ul style="list-style-type: none"> Response times to enquiries for data and information <ul style="list-style-type: none"> Description of aimed service level for responding to user requests (where these cannot be met online) Whether an Enquiries or Help Desk is available

	Details of surveys of customer satisfaction undertaken
<i>DAC Response</i>	
Committed to raising awareness of the holdings and promoting the use of the data	<p>Describe facilities available at the DAC to discovery data holdings:</p> <ul style="list-style-type: none"> • Details of how the data can be searched or interrogated by interested users (e.g. online metadata search, physical access on site, etc) • Short summary of any online search functionality <p>Describe other search facilities used, e.g.</p> <ul style="list-style-type: none"> • Discovery metadata available through the GI Gateway, National Biodiversity Network, UK MED Directory/EDMED, etc. <p>The DAC should provide an indication of participation in conferences and exhibitions; production of promotional leaflets; flyers and articles</p> <p>In addition to the activities above the DAC should provide information on:</p> <ul style="list-style-type: none"> • Data products available • Linkages with other organisations who use the data for generation of products • Current projects aiming to increase and promote data use • Statistics/metrics indicating data usage
<i>DAC Response</i>	
Making datasets freely available wherever possible (not necessarily at zero cost)	<p>MEDIN DACs should have a policy on data access. In general DACs should aim to make data sets freely available, although it is recognised there may be restrictions on access to data for a number of reasons including national security, commercial confidentiality, for scientific research to allow the principal investigators and their co-workers to exploit the data in the first instance. However, release of data to the wider community after a period of 1-3 years from data collection should be strongly encouraged. Metadata should be made available at zero cost and data should be made available at zero cost wherever possible.</p> <p>The data access policy should include the following:</p> <ul style="list-style-type: none"> • Details of what can / cannot be obtained on-line (e.g. metadata only, full dataset download) • Licensing arrangements • The format(s) that data can be provided in • The media used for providing data (if data are not online) • Costs associated with data provision (or cost scales) – including cost of media as well as staff time <p>Wherever possible, data policies should be in accordance with internationally-agreed data policies (e.g. IOC Oceanographic Data Exchange Policy, GOOS Data Policy, WMO Resolution 40, ICES Data Access Policy, etc.)</p>
<i>DAC Response</i>	

Assumptions

1. It is accepted that there may be instances where there is more than one copy of a dataset within the MEDIN structures but that there will be one MASTER (original) version, held by the originator or transferred to a DAC.
2. It is accepted that there may be instances where datasets of similar type are held in separate DACs.
3. It is accepted that there will be a range of different levels of value added and commercial activity with the MEDIN DACs.
4. There are Funders of Data Collection, Contributors of Data, Holders of Data and Users of Data in MEDIN (all subject to relevant sets of requirements) as well as DACs; these roles are not mutually exclusive.