MEDIN help Defra towards their target of making 8,000 datasets open

In the last edition of this newsletter we reported the ambition of Elizabeth Truss MP, Secretary of State for the Environment, Food and Rural Affairs, to make Defra a more data-driven organisation, where data is open by default. The Open Data project has started with the Data Accelerator programme, challenging Defra to make at least 8,000 Defra datasets openly available by June 2016.

Of course, the principles of good data stewardship and data sharing are familiar to colleagues in the Marine Directorate at Defra, and the data collections agencies they work with, as they have been key supporters of the Marine Environmental Data and Information Network (MEDIN). Since its inception in 2008, MEDIN has been helping Defra publish its marine environmental data, and helping the public discover Defra's data. Even before MEDIN, the National Data Archive Centres (DACs) were publishing Defra data. For example, data from the National Tide Gauge Network (currently owned by the Environment Agency) have been published by the British Oceanographic Data Centre for several decades, helping predict where and when storm surges may occur and determining when to deploy the Thames Barrier.

So, how is MEDIN helping Defra reach its target of releasing 8,000 open datasets? The first step has been to identify all the Defra datasets described in the MEDIN metadata catalogue. Working with Defra, MEDIN is now working to ensure that as much of Defra data as possible are published as open data - which is defined by the Data Accelerator project as data available under Open Government License (OGL), and with direct access to online data from metadata on data.gov.uk.

MEDIN publishes metadata and, on request, forwards metadata records to the UK Government's data.gov.uk. MEDIN already provides this service for the Marine Management Organisation, and for some Cefas data. MEDIN continues to offer this service to all organisations participating in the Defra Data Accelerator project.

MEDIN has identified hundreds of datasets to count towards the Data Accelerator programme, as long as

the data owner is willing for the data to be published under OGL and/or the metadata to be published on data.gov.uk. Defra data archived at and published by the MEDIN Data Archive Centres are considered within scope of the programme. The seven MEDIN Data Archive Centres continue to provide the UK marine community with a long-term solution to publishing UK marine data, providing thematic specific expertise and ensuring that the data they publish has been quality controlled and is available for re-use in the future.

MEDIN welcomes this initiative to push ever more Defra data into the public domain and is pleased that there is increasing recognition of the opportunities and efficiencies that can be made by enabling the re-use of data.

SeaDataNet II project reaches a successful conclusion

The EU funded SeaDataNet II project finished at the end of September 2015 and held its final Plenary meeting at Ifremer, Brest, to showcase the main results.

This project, building upon and upgrading the activities of the earlier SeaDataNet and previous associated projects, has developed a pan-European infrastructure for managing, indexing and providing access to ocean and marine datasets and data products, acquired from research cruises and monitoring activities in European marine waters and the global ocean. The SeaDataNet II consortium comprised 58 organisations from 23 member states of the European Union and 11 other countries bordering the European seas. These organisations include National Oceanographic Data Centres and marine information services of major research institutes, together with experts in IT, data publishing, and modelling, as well as international organisations, namely IOC, ICES, and EU-JRC. *|END:IF|*

Participants of the meeting in Brest (image by M. Gouillou, IFREMER)

SeaDataNet has become one of the EU's largest marine data infrastructures and represents a model for other platforms dealing with marine data management. The developed standards, tools and services for data and metadata, the controlled vocabularies and the data quality control guidance documents and software have been adopted and adapted in many different EU data management projects. The number of data centres connected has increased over the SeaDataNet II project from 65 to 102 and the number of Common Data Index (CDI) entries has increased from circa 1.05 million to 1.8 million. The data sets cover physical oceanography, chemistry, geology, geophysics, bathymetry and biology, originating from more than 550 organisations. All the data sets are collected in research projects and monitoring programmes and after a process of quality control and formatting, archived in the local databases of the data centres

which are connected to the SeaDataNet CDI Data Discovery and Access service.

Coordination activities of SeaDataNet covered maintenance and operation of all the SeaDataNet discovery services and updating the metadata and data catalogues. Capacity building and training took place to ensure that all partners were familiar with the different SeaDataNet tools. Research and development activities included the specification and governance of standard metadata, data and data product formats, quality check methods, common vocabularies and new services, as well as the development and governance of software tools, services and interoperability solutions and the development and updating of standard data products. Major outcomes achieved include:

- **Standardisation:** Major achievements during the project include INSPIRE compliance, a new version of NERC-BODC vocabulary server, the NetCDF transport format and the extension of the Ocean Data View (ODV) format for handling biological data.
- **Interoperability** with other data management systems including global initiatives, such as the IOC-IODE Ocean Data Portal (ODP) and the GEOSS portal; machine to machine interfaces; operational updates of the SeaDataNet catalogues (metadata harvesting); introduction of Marine ID, to manage registration and authentication of data users; and a demonstration study using Sensor Web Enablement (SWE) technologies
- **Robustness:** Including (i) monitoring of the widely distributed infrastructure to ensure that the services remain highly available and accessible to users, to detect failures of systems components and fix them, and to identify critical components within the system in order to update them to improve their robustness, and (ii) Data duplicate checks which can arise when different data providers submit the same data sets. Therefore, a duplicate detection algorithm has been integrated into the CDI validation and import procedures and reports are forwarded back to data providers for checking and possible correction.
- **Data products:** Regional products have been generated for the Arctic Ocean, North Atlantic Ocean, Baltic Sea, North Sea, Mediterranean Sea and Black Sea. These are described in the SeaDataNet product catalogue (SEXTANT) and are freely available for Marine-ID users. Two products are available: (1) aggregated datasets of temperature and salinity (ODV collections of all SeaDataNet measurements of temperature and salinity) and (2) temperature and salinity climatologies (regional gridded field products based on the aggregated datasets).

Further information is available from the SeaDataNet web-site, and the project's final newsletter.

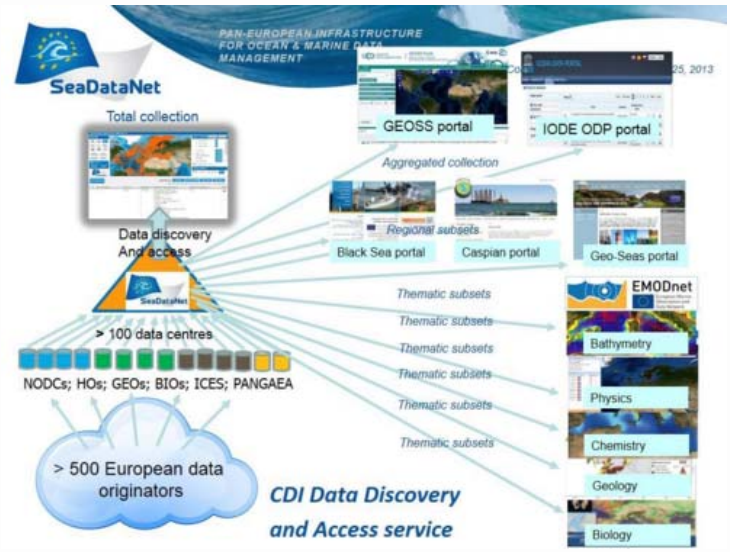


Illustration of adoption of CDI service in other projects and portals

Visualisation of T&S data collection for Atlantic

Marine Industries Positive about Sharing Data

"A Review of Access to Industry Marine Environmental Data", a new report available from the UK Productive Seas Evidence Group (PSEG) has just been published and may be downloaded from http://www.gov.scot/Topics/marine/science/MSCC/PSEG.

The UK Government, primarily through the Marine Science Coordination Committee (MSCC), has committed to enabling open access to as much marine data as possible. This newly published report, commissioned by PSEG, takes this commitment forward for marine environmental data collected by major UK marine industries.

Following a stakeholder engagement exercise, the report highlights the benefits to the UK which could result from greater use of industry data. More specifically it addresses:

- What data are being collected by 8 key UK marine industries?
- How these data could be used by the public sector?
- What barriers are preventing data from being shared?
- How can we overcome these barriers?

This report contributes to the overall objective of ensuring that data collected by industry for consenting, monitoring or other purposes that has a re-use value is stored in a secure, long term archive and is made as widely available as possible, as quickly as possible. The Marine Environmental Data and Information Network (MEDIN) infrastructure and the network of MEDIN Data Archive Centres (DACs) are acknowledged as a successful process for cataloguing, storing and accessing data. However, the report notes that there is a lack of awareness within the commercial sector of the range of options already available for sharing and safeguarding industry data at the MEDIN DACs. In response to recommendations within the report, MEDIN will be holding an open meeting on 9th February 2016 in London to share information about MEDIN with the commercial sector and to discuss how the barriers to sharing industry data can be overcome collaboratively. To be kept up to date with the latest open meeting news join the email distribution list.

The study was carried out by ABP Marine Environmental Research and Peter Barham Environmental and

was co-funded by the Marine Management Organisation, Marine Scotland, MEDIN and The Crown Estate.

The suggested citation for the report is as follows: ABPmer, (2015). 'A Review of Access to Industry Environmental Data'. Report to the UK Productive Seas Evidence Group (PSEG), November 2015, 67 pages.

Get involved in the redesign of the MEDIN portal

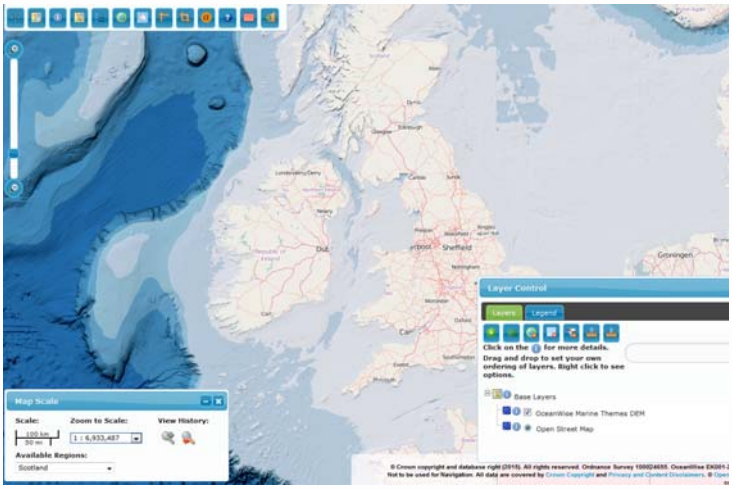
In 2016, we'll be looking at a complete overhaul of the MEDIN data discovery portal to make it more user and mobile friendly, and give you quicker and easier access to marine data.

As an important part of that, we'd like to invite you - the portal users - to directly input into the direction the portal will take.

Anyone who wants to get involved will have the opportunity to give us direct feedback on what they like about the current site, as well as what they'd like to see in the new version, and we'll be running a range of online and in-person questionnaires and workshops to ensure that no matter your location or availability, you can still get involved in and help us make MEDIN the go-to portal for marine data.

Click here to find out more.

Marine Scotland's NMPi continues to grow



National Marine Plan interactive (NMPi), Marine Scotland's on-line portal, continues to grow to facilitate national and regional marine planning and support future marine assessments.

In response to stakeholder feedback changes have recently been deployed. At first start up the *'Layer Control'* will be empty, except for base layers and any of the User's own 'My Information' layers if you are logged in. Users can now *'Add layers'* from the main layer list (and using the familiar search function), so creating a short list of layers on which the user is working. This makes it easier to see which layers are selected and turned on from the growing list of content which now approaches 600 layers. The user can now change the *order* in which these draw on screen, by 'dragging and dropping' the layer name, so, for example, ensuring that point and line data overlay polygon data. This increases the usability of any created maps especially when linked to the printing function.

Users are also able to *save a group of layers*, their 'work in progress' and re-open at their next session. In the 'Layer Control' box there is a 'Save Map Configuration' button which will store an xml file on your local drive of choice. You do not have to be logged in to do this. The xml file can be shared with other users by e-mail for example, or used on next start up to return to 'work in progress' by using the 'Load Map Configuration' button. Any number of such files can be kept and given a name meaningful to the user.

'Time Aware', a feature to assist with displaying data that will help inform future assessments, has also been introduced. This will allow a time series of data to been seen, for example year by year where a data set has multiple years of data. This allows NMPi to host a single layer of data to cover multiple time periods and the user can control the visibility of data and see change over time. This will usually be year on year but could be other time intervals. Once the layer is turned is, right click over the layer name and select 'Time Aware'. A 'Layer Over Time' control box appears. The data will scroll over a pre-determined cycle but the user can stop and start and drag from year to year. As data layers are updated more layers will have this functionality. Currently users can see this function on:

- Marine strandings data 2013-2014 (cetaceans, seals, sharks, turtles
- Keep Scotland Beautiful - Blue Flag and Seaside Awards since 2012
- Nephrops - TV-assessed burrow density 2007-2014

A 'Submit Fault Report / Comment' button is also now available in the main NMPi tool bar (2nd button from right hand end) that allows users to send comments or details of faults direct to Marine Scotland. Users

have the choice to include an e-mail address for a reply.

Other enhancements are in preparation for 2016, as well as the inclusion of new data sets. A future Marine Data News article will keep you posted. In the meantime, if there are any particular data sets that would be useful to display on NMPi, please contact Marine Scotland. Feedback is welcome at any time.

News

JNCC release first batch of datasets as open data

The Joint Nature Conservation Committee (JNCC) has released 493 marine survey datasets under an Open Government Licence as part of the #OpenDefra campaign launched by Elizabeth Truss MP, Secretary of State for the Environment, Food and Rural Affairs in June 2015.

This first batch of datasets were collected as part of the "Marine Nature Conservation Review"(MNCR), between 1990 and 1998.

JNCC anticipates that these data could be applied to areas such as developing predictive models, detecting change and potential causes of this and supporting more efficient regulation.

More information

Call for Papers: Advances in Data Modeling and Knowledge Representation for Research Data

A call for papers has been released for a special edition of the Data Science Journal on Advances in Data Modelling and Knowledge Representation.

The Data Science Journal is a peer-reviewed, open access, electronic journal dedicated to the advancement of data science and its application in policies, practices and management of Open Data.

More information.

Workshops and Events

MEDIN Open Meeting: 9 Feb 2016

MEDIN will be holding an open meeting on 9th February 2016 in London to share information about MEDIN with the commercial sector and to discuss how the barriers to sharing industry data can be overcome collaboratively.

This one day, free-to-attend open forum is an opportunity to discuss the barriers to sharing industry marine data and to learn about how MEDIN can help share and safeguard such data.

To register for the meeting please join the email distribution list.

MEDIN Workshops:
Mar 9 2016

The next MEDIN workshop will be on the 9th March 2016 (subject to change) at British Geological Survey in Keyworth.

The workshop is free to attend, with lunch provided. There is still time to register either by contacting Sean Gaffney, or signing up here.

For more information on this and future workshops please visit the MEDIN website.

Challenger Society 2016 Conference - Oceans and Climate:
Sept 5-8 2016

The 17th Biennial Conference of the Challenger Society for Marine Science is to be held in Liverpool at the University of Liverpool on the 5-8th September 2016.

The Challenger Conference usually attended by hundreds of researchers from the UK and world-wide provides a showcase of marine science and technology, covering all areas of ocean research along with early career events and education & outreach opportunities.

More information

International Conference on Marine Data and Information Systems - IMDIS:
Sept 12-14 2016

Gdansk (Poland) September 12-14, 2016

The IMDIS series of Conferences promotes the meeting of different communities working in informatics, data management, research, environmental protection, etc. It focuses on access to data, metadata and products, communication standards and adapted technology to ensure platform interoperability. IMDIS 2016 aims at providing an overview of the existing information on marine environmental data, and showing the progresses on development of efficient infrastructures for managing large and diverse data sets.

The Conference will be organised in four sessions:

- Marine information and data management
- Marine environmental data bases: infrastructures and data access systems
- Data Services in ocean science
- Services for users and education

Submission of extended abstracts (1 - 2 pages) is through the IMDIS web pages available from December 2015. The deadline for submission of papers and posters is April 15, 2016.

If you are intrested in supplying an article, or you wish to advertise an upcoming workshop or event in Marine Data News please contact:
MEDIN enquires