MEDIN DAC WG Meeting 8

10:00-15:00 Wednesday 25th April 2018, Uni. Liv. in London, London

|  |  |  |  |
| --- | --- | --- | --- |
| **Attending** | | | |
| Jon Parr – Chair (JP) | MBA | Louisa Matthews (LM) | ADS |
| James Ayliffe (JA) | BODC | Charlotte Miskin-Hymas (CMH) – minutes | MEDIN |
| Mary Brown (MB) | Cefas | Robin McCandliss – (RM) | MEDIN |
| James Cooke (JC) – phone | UKHO | Mary Mowat (MM) | BGS |
| Gaynor Evans (GE) – SKYPE | MEDIN | Clare Postlethwaite (CP) | MEDIN |
| Mike Jackson (MJ) - phone | UKHO | Jens Rasmussen (JR) | MSS |
| Dan Lear (DL) - SKYPE | DASSH | Callum Stone (CS) - phone | Met Office |
| **Apologies** | | | |
| Chelsea Bradbury (CB) | The Crown Estate | Peter McKeague | HES |
| Gareth Burns | AFBI | Kay Thorne | BODC |
| Gareth Edwards | RCAHMW | Oliver Williams | Cefas |
| Paul Henni | BGS |  |  |

1. **Round table introductions**

Welcome to new member Callum Stone, replacing Martyn Sunter to represent the Met Office.

1. **Review actions last meeting**

**Table 1:** Actions from DAC WG meeting on 7th November 2017

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | **Action** | | **Progress** | | 7.1 | **Clare Postlethwaite** to issue another call for small scale projects to be completed by end of this FY. | Completed. | | 7.2 | **Amended action: DASSH and MSS** to check their status for INSPIRE compliance is portrayed correctly in the INSPIRE table and report back to RM. | Ongoing: BGS, BODC, Cefas, UKHO have reported back. DASSH and MSS need to check | | 7.3 | **All DACs** to feedback to Lesley Rickards by 22nd December 2017 whether they can provide a breakdown each year of who is using their data. | Closed. DACs will report in annual report, where possible. | | 7.4 | **Lesley Rickards/Robin McCandliss** to send ADS the DAC reaccreditation form. | Done. ADS have returned it, review in progress. | | 7.5 | **Charlotte Miskin-Hymas** to include recent DAC reaccreditations in MDN. | Ongoing: Delayed to next edition of MDN in 2018 (end of June/July) | | 7.6 | **Jens Rasmussen** to form a small, technical focus group to explore increasing volumes and complexity, support for machine-to-machine access, unstructured data, semantic linking and feedback to DAC WG. | Closed: Superseded as it overlaps with the new business plan. There is sufficient technical representation in the writing group. | | 7.7 | **DACs** to provide comment and feedback on proposed changes to MEDIN web pages. | Done. | | 7.8 | **Peter Thijsse** to provide a validation tool log in for each organisation | Done. | | 7.9 | **Louisa Matthews** to discuss OASIS metadata validation issue with Gaynor Evans and Peter Thijsse | Ongoing: Remapping is needed to MEDIN metadata schema. Issue with getting developer resource and is not likely to happen soon. | | 7.10 | **Robin McCandliss** to arrange a DAC WG teleconference in early 2018 for DACs to report progress in fixing metadata validation issues. | Closed: Postponed due to portal bug needing resolution in early 2018. Progress to be reported instead as part of this meeting. | | 7.11 | **Jens Rasmussen** to validate metadata records using both Metadata Maestro and MARIS tool and report back to DAC WG at teleconference in early 2018. | Ongoing: Jens re-exported and validated 180 xml records through the Maestro, however all records failed validation in the MARIS tool. Bugs in the MARIS tool are the likely cause. | | 7.12 | **DACs** to provide Clare Postlethwaite (by 29 Nov) with statement for Sponsors and Executive Boards on usefulness of metadata checks, and outline changes made and not made. | Done. | | 7.13 | **Robin McCandliss and Charlotte Miskin-Hymas** to identify where to include a statement on licenses on the MEDIN website. | Done. | | **Actions outstanding from earlier DAC WG meetings** | | | | 6.1 | **Clare Postlethwaite** to ensure UKDMOS/MEDIN document is made available on MEDIN website once finalised. | Ongoing. | | 6.6 | **Mary Mowat** to consult with UKHO about where resource locators in the MEDIN catalogue should point to for data that are held both at BGS and UKHO. | Ongoing. Checking backscatter data going into UKHO records and Mary to follow up outside the meeting. | | 6.13 | **Mary Mowat** and **Chelsea Bradbury** to liaise to identify data sets that are present via TCE and BGS and ensure that it does not appear in MEDIN catalogue twice. | Ongoing. Mary has provided a list of the data held by BGS but TCE still need to carry out a review of data that they already hold. Hopefully this should be done in the next three months. | | 5.1 | **Amended action: Clare Postlethwaite** to a) liaise with Data Task Group to finalise MSFD indicator assessment metadata template, which will require a link to a MEDIN metadata record and a DOI for each dataset used in the assessment; b) set realistic expectations for data archive at the DACs to minimise the risk that following the monitoring programmes update, the DACs will be expected to archive the data with little notice. | Ongoing: The Data Task group have drafted a metadata template for MSFD which includes fields to link to MEDIN catalogue and to DOI’s where data is archived. | | 5.2 | **DACs** to report whether they could archive snapshots of the MSFD data sets with DOIs | Closed. DACs said they can mint DOIs for datasets. Combined with amended action 5.1. | | 5.3 | Add re-accreditation timetable to web-site; **Lesley** to pass to **Charlotte** | Done. Accreditation timetable sent to Charlotte. | | 4.1 | **Clare Postlethwaite** to make the Archive data URL field mandatory in UKDMOS.  *As of 11/10/16 14/276 have real time URL, 76 have archive URLs. This will be checked again following content update.* | Closed. No additional progress. Requests to update UKDMOS will need a request for the link to the data so field can be populated. CP will carry on reporting if information is useful. | | 3.1 | Following Clare/Lesley review,send a template to DACs that lists UKDMOS monitoring programmes; DACs to check list and identify the programmes where DACs should be/are publishing and where other services are publishing. | Closed: This is for discussion between UKDMOS, MSFD and the Data Task group. | | 3.8 | DACs provide input to Hannah Williams to indicating which INSPIRE themes are applicable to their data to include on MEDIN website. | Closed. | | 2.13 | Hannah to request 4 images from each DAC and give information on the format required. Then to mock up some pages and circulate them before the next meeting. | Closed. | | 2.15 | Clare to arrange an autumn workshop in Scotland with MPA organisations to ensure suitable data management plans are in place. Should involve BGS, DASSH, JNCC, SNH, Marine Scotland, Natural England and any other DACs that are interested. | Done. | | 2.18 | Clare to speak to JNCC about possibility of them becoming a Bird DAC. | Ongoing. Will be revisited this financial year as part of work plan. | |

A question was raised about other data types that MEDIN DACs do not currently handle. BODC is tentatively looking at taking on underwater noise data, and DASSH has discussed the possibility of being a backup archive for litter data for the Marine Conservation Society. Strategically, it is a decision for the Sponsors and the Executive team as to whether the DACs should expand to cover these areas and this should be addressed in the new business plan.

**3. Review of DAC Work Plan 2017/18**

3(a) Direct access to data

MEDIN supported development work during 2017/18 at BGS, BODC and DASSH to improve the ‘2-click to data’ progress. There were some outstanding tasks at the end of the financial year. Update from the three DACS:

BGS: All work has now been completed and a MEDIN metadata link is in place.

BODC: All work now completed.

DASSH: Not fully complete. A DOI online tool was developed to automate minting of DOIs. This tool needs to be integrated with the metadata creation tool, which is currently awaiting sign off by the Standards Working Group.

3(b) Small-scale projects

Four small-scale projects were funded by MEDIN in 2017/18, looking to increase access to industry data and increase access to monitoring data.

ADS was funded to develop a photogrammetry interface, allowing online access to monitoring data. Photogrammetry files are large and difficult to provide online. The solution developed at ADS allows streaming of a low-resolution version of the file so that a user can decide whether to download a high-resolution version.

***Action 8.1: Louisa Matthews to circulate a copy of the photogrammetry interface report to the DACs, as they may be interested in the solution for high volume data types.***

***Action 8.2: Louisa Matthews to send Dan Lear the ADS wiki process for streaming video files.***

Cefas was funded to sort and catalogue paper records of herring data from 1905-2000, a very important historic data set. These records (approximately 1200) were at risk of being lost, due to the building replacement work that is happening at Cefas.

DASSH had two projects funded – one internal and the other in collaboration with a consultancy (Ocean-Ecology).

The internal project was to enable better interaction with the Healthy and Biologically Diverse Seas Evidence Group (HBDSEG) Pelagic Indicator group, so that the data can be managed better. They have developed a method for extracting data in a format that can be easily used. It is currently being tested by the indicator group and a report is expected imminently.

The second project, in collaboration with Ocean Ecology Ltd (OEL), was to enhance OEL’s database ABACUS so that it exports data in the MEDIN data guideline format, enabling easier ingestion of data into DASSH. This is almost completed.

3(c) Work Plan 2017-18 Tasks carried forward

The following tasks from the 2017/18 work plan have been carried forward into the 2018/19 work plan for the DACS: identifying MSFD datasets to archive at the DACs, completion of AFBI FishDAC accreditation, and updating the DAC components on the MEDIN website.

**4. DAC network status**

4(a) Feedback from MEDIN Open meeting 24th April 2018 (CP)

The MEDIN Open Meeting was held to celebrate 10 years of MEDIN and gather input for the new MEDIN Business Plan. Overall, the feedback from attendees was positive. The morning session highlighted MEDIN achievements and challenges. Sean gave a well-received presentation showing the benefits of MEDIN partnerships, using examples from the University of Bangor, The Crown Estate and Natural Resources Wales. The ‘World Café discussion session’, structured around the FAIR principles, worked very well, with a good mix of attendees and lots of interesting discussion. The DACs also had a poster session, but more notice is required when requesting posters for future meetings.

4(b) Brief report from DACs

The DACs gave a verbal report on progress. Items of note:

* UKHO – Judith Thomas has now retired and Mike Jackson is taking over her role.
* Met Office – Callum Stone has taken over as the Met Office MEDIN representative, replacing Martyn Sunter.
* Marine Scotland - Taking part in INTERREG COMPASS project and how to incorporate MEDIN standard compliance.
* ADS – Louisa is leaving the ADS in September. ADS is experiencing a significant increase in deposits of marine data.
* BODC and BGS – business as usual.
* CEFAS – Laura Hanley is the new CEFAS Data Manager.
* DASSH – new role as an accredited IODE data centre and are the UK node for OBIS.

**5. DAC Work Plan 2018/2019**

5(a) Outline work planned

There is a new key target for DACs regarding metadata quality – achieving validation of records in the new portal. For 2018/19, it was noted that small data archiving project proposals may be at risk due to funding uncertainties – this is still the case, but it may be possible to use carry–over money from previous years to plug the gaps in the finances. CP will confirm this in the next month and hopes to launch them during Q1.

Another MEDIN codefest is on the work plan. The budget is limited to covering a room and staff time and previous codefests have been heavily dependent on the efforts of Jens and Dan. The MEDIN Executive were keen to broaden the next codefest out beyond DACs, but this may not be realistic with such a limited budget.

After some discussion, it was agreed that the realistic option for another codefest would be to follow the same format as previously, i.e. limiting involvement to the MEDIN DACs. While this is less ambitious than opening it up more widely, there were good results the last time that would make it worth doing.

***Action 8.3: Clare Postlethwaite to feedback to MEDIN Executive that opening the codefest out to a wider audience is not possible with the available budget, but there would be benefit in holding an internal one to share expertise.***

There was significant effort required to prepare for the codefest - harvesting MEDIN data for developers to access and use was laborious and difficult. If UKDMOS is to be the focus for the next hackathon, how feasible will it be to extract the data required? It was agreed that someone with technical knowledge of UKDMOS should be co-hosting (e.g. MARIS or Gaynor). Dan and Jens are happy to be involved, but the organisation of the event will need to be coordinated by a member of the MEDIN core team.

***Action 8.4: Clare Postlethwaite to clarify what needs to be done to prepare UKDMOS data for hackathon use and to find out what would be useful for users.***

***Action 8.5: Charlotte Miskin-Hymas to work with DL, JR, GE and MARIS and coordinate the MEDIN hackathon, to happen in Q3 alongside the mega meeting.***

5(b) Development of the next MEDIN 5 year Business Plan, DAC input (CP)

A brainstorming exercise was carried out at the last mega meeting to capture DAC priorities for the next 5 years, contributing to the next business plan. The outcome of that exercise was taken to the Executive and Sponsors board, where they agreed all the priorities identified.

A small writing team has been assembled, including representatives from the DAC Working Group. The plan is that the writing group meet over the next couple of months to draft the business plan and will send it to the DACs for review, probably by late summer.

**6. Promoting the DAC Network**

6(a) New DAC web pages demonstration

CMH gave a demonstration of the new MEDIN website. Feedback from the DACs was very positive. There are still edits to be made to the DAC pages, but the DACs felt that the overall layout, look and feel is a big improvement.

***Action 8.6: Charlotte Miskin-Hymas to contact Peter McKeague for interesting shipwreck images for the website.***

***Action 8.7: Charlotte Miskin-Hymas to send link to website to DACs and DACs to feedback comments on the text for their DAC pages.***

***Action 8.8: Robin McCandliss to include checking DAC webpages as a standing item on the DAC agenda to ensure the text is up to date and remind DACs to check before the DAC meetings.***

**7. Coordinated data archiving**

7(a) Check if there is an update on pilot of coordinated approach to archiving multidisciplinary datasets (DASSH)

DASSH have had no enquiries to deposit multi-disciplinary data as yet.

CP has had discussions with Cefas recently about MCZ data archiving and what data DASSH and BGS could expect in the coming year. BGS received less data than expected and the data DASSH received was not in standardised format during the last FY. In the next month, it will be agreed how the £30k funding that Cefas provides for archiving the data should be split between BGS and DASSH.

***Action 8.9: Clare Postlethwaite to organise a meeting to discuss MCZ data archiving at the MEDIN DACs for 2018-19. Including Cefas, BGS, DASSH, and possibly UKHO and JNCC.***

8(a) update on validation requirements for metadata

The metadata validation tool is returning many false errors due to several bugs that have not yet been fixed by MARIS. Error number 61 is particularly prevalent for all the providers and this is the priority bug for MARIS to address. It is important to address these issues as soon as possible, as the DACs are willing and engaged with the new validation tool and will lose interest if it is not working.

***Action 8.10: Gaynor Evans to reiterate with MARIS the need to fix the bugs and to provide a prioritised list.***

***Action 8.11: Gaynor Evans to circulate the metadata validation errors for each provider to the DACs***

8(b) Metadata Standards UUIDs should be unique and persistent.

There have been occasions where changes have happened to the unique identifiers in the new portal and this is of concern, as they should be persistent. According to the MEDIN Standard, ‘The file identifier is a code that is globally unique and remains with the same metadata record even if the record is edited or transferred between tools’.

After some discussion, it was felt that this may be a historic issue. There are also plans for GEMINI to include the UUID in the metadata schema rather than the XML, which should also reduce the risk. DASSH noted that there are some duplicate records in the catalogue, which will need to be manually resolved. Out of 1600 metadata records, ~80 are duplicates.

It was agreed that this should be reviewed in 6 months to ensure it has not happened again.

***Action 8.12: Gaynor Evans to check whether UUIDs have changed and report back at next DAC WG meeting.***

**9. AOB**

9(a) Duplication of data

JA raised a question about MSS Scottish Coastal Lab data (which includes water column data) that are published on the MSS portal, have a MEDIN metadata records and may have a DOI and the fact that BODC receive the water column data for ingestion – should BODC ingest these data (and push them into wider portals such as SeaDataNet), even though they are already published via MSS?

There will be an agenda item at the next DAC WG meeting to discuss the archive of data that have already been published (MCZ data, duplication of data, multiple portals).

9(b) Standardisation of metrics

As it is the time of year when DACs prepare their metrics for the annual reports, the question was raised again about consistency of metrics reported by the DACs. Is there a way that all DACs can submit meaningful metrics to quantify the use of data from DACs that can be compared across the DACs? Some DACs report numbers of rows as data sets, while others report numbers of files. Do we want to show perpetual growth or do we want to show quality, level of reuse etc? A suggestion to use DataCites figures for DOI resolutions was put forward, for those DACs that use DOIs.

CP requested that the DACs report this year as previously so that a 5-year cycle is captured. The graphs will be normalised in the overall annual report. The new business plan should address deriving a new baseline for DAC metrics for the next 5 years.

9(c) Date of next meeting

RM will poll for meeting dates in November to coincide with mega meeting and hackathon.

**ACTION LIST**

**New actions from DAC WG meeting 25th April 2018**

|  |  |  |
| --- | --- | --- |
| **Action** | | **Progress** |
| 8.1 | **Louisa Matthews** to circulate a copy of the photogrammetry interface report to the DACs, as they may be interested in the solution for high volume data types. | Done |
| 8.2 | **Louisa Matthews** to send Dan Lear the ADS wiki process for streaming video files. | Done |
| 8.3 | **Clare Postlethwaite** to feedback to MEDIN Executive that opening the codefest out to a wider audience is not possible with the available budget, but there would be benefit in holding an internal one to share expertise. |  |
| 8.4 | **Clare Postlethwaite** to clarify what needs to be done to prepare UKDMOS data for hackathon use and to find out what would be useful for users. |  |
| 8.5 | **Charlotte Miskin-Hymas** to work with DL, JR, GE and MARIS and coordinate the MEDIN hackathon, to happen in Q3 alongside the mega meeting. |  |
| 8.6 | **Charlotte Miskin-Hymas** to contact Peter McKeague for interesting shipwreck images for the website. |  |
| 8.7 | **Charlotte Miskin-Hymas** to send link to website to DACs and DACs to feedback comments on the text for their DAC pages. |  |
| 8.8 | **Robin McCandliss** to include checking DAC webpages as a standing item on the DAC agenda to ensure the text is up to date and remind DACs to check before the DAC meetings. |  |
| 8.9 | **Clare Postlethwaite** to organise a meeting to discuss archiving MCZ data at the MEDIN DACs for 2018-19. Including Cefas, BGS, DASSH, and possibly UKHO and JNCC. |  |
| 8.10 | **Gaynor Evans** to reiterate with MARIS the need to fix the bugs and to provide a prioritised list. |  |
| 8.11 | **Gaynor Evans** to circulate the metadata validation errors for each provider to the DACs |  |
| 8.12 | **Gaynor Evans** to check whether UUIDs have changed and report back at next DAC WG meeting. |  |
| **Actions outstanding from earlier DAC WG meetings** | | |
| 7.2 | **Amended action: DASSH and MSS** to check their status for INSPIRE compliance is portrayed correctly in the INSPIRE table and report back to RM. |  |
| 7.5 | **Charlotte Miskin-Hymas** to include recent DAC reaccreditations in MDN. | Delayed to next edition of MDN in 2018 |
| 7.9 | **Louisa Matthews** to discuss OASIS metadata validation issue with Gaynor Evans and Peter Thijsse | Ongoing: Remapping is needed to MEDIN metadata schema. Issue with getting developer resource and is not likely to happen soon. |
| 7.11 | **Jens Rasmussen** to validate metadata records using both Metadata Maestro and MARIS tool and report back to DAC WG at teleconference in early 2018. | Ongoing: Jens re-exported and validated 180 xml records through the Maestro, however all records failed validation in the MARIS tool. Bugs in the MARIS tool are the likely cause. |
| 6.1 | **Clare Postlethwaite** to ensure UKDMOS/MEDIN document is made available on MEDIN website once finalised. | Ongoing |
| 6.6 | **Mary Mowat** to consult with UKHO about where resource locators in the MEDIN catalogue should point to for data that are held both at BGS and UKHO. | Ongoing. Checking backscatter data going into UKHO records and Mary to follow up outside the meeting. |
| 6.13 | **Mary Mowat** and **Chelsea Bradbury** to liaise to identify data sets that are present via TCE and BGS and ensure that it does not appear in MEDIN catalogue twice. | Ongoing. Mary has provided a list of the data held by BGS but TCE still need to carry out a review of data that they already hold. Hopefully this should be done in the next three months. |
| 5.1 | **Amended action: Clare Postlethwaite** to a) liaise with Data Task Group to finalise MSFD indicator assessment metadata template, which will require a link to a MEDIN metadata record and a DOI for each dataset used in the assessment; b) set realistic expectations for data archive at the DACs to minimise the risk that following the monitoring programmes update, the DACs will be expected to archive the data with little notice. | Ongoing: The Data Task group have drafted a metadata template for MSFD which includes fields to link to MEDIN catalogue and to DOI’s where data is archived. |
| 2.18 | **Clare Postlethwaite** to speak to JNCC about possibility of them becoming a Bird DAC. | Ongoing. Will be revisited this financial year as part of work plan. |