

**MEDIN Standards Meeting  
Edinburgh BGS 29th April 2009  
10.30 – 16.00**

Mark Charlesworth (MEDIN) – WS Lead  
Steve Gontarek (SAMS)  
Dave Cotton (MEDIN)  
Roy Lowry (BODC)  
James Dargie (CCW)  
Steve Wilkinson (JNCC) - Chair  
Daniel Lear (MBA)  
Becky Seely (MBA)  
Roger Coggan (CEFAS) - VC  
Mary Mowat (BGS)  
Olivia Merritt (SeaZone)  
Rachel Heaven (BGS)

1) Introduction, welcome agenda and minutes of the last meeting

All were welcomed to BGS. Apologies were received from James Rapaport, Colin Graham, Chris Hill and Kieran Millard. Discussion points on metadata tools and granularity were added to the agenda. Minutes of the last meeting were accepted and action points were either dealt with, to be discussed at this meeting or addressed with the exception that MC should ask if a NI representative would like to join the group. This action point will be picked up at the MEDIN partners meeting on June 30<sup>th</sup>.

2) Terms of Reference for Standards WG

The amended Terms of Reference for the group were displayed. Some changes were made to the definition of different types of standards. It was felt that one last opportunity to comment should be given and these terms should be sent around with the minutes of the meeting for agreement.

**AP. MC** to circulate the terms of reference for final opportunity to comment.

3) Discovery Metadata Standard and Tool.

- Update on guidelines and schema

An update on the development of the discovery standard guidance and schema was given and general comments received. It was appreciated that a lot of work had been put into the standard and that some minor deviations from the ISO 10139 standard were inevitable. The UK GEMINI 2 standard was to be revised in early May and the MEDIN standard may need some minor revision following its release. The guideline document and schema will be completed by 8<sup>th</sup> May and made available via the website. The schema will be reviewed by SeaZone who are also building the schematron which is a

tool to validate a metadata record. There are some other available schematrons in use such in SeaDataNet and these may prove useful to coordinate with.

**AP. BS/DL** to send final schema to MC by 8<sup>th</sup> May

**AP. OM/JR** to contact Karen Vickers at BODC who devised the schematron tool for SeaDataNet

It was not clear at which point the validation of metadata would happen in the flow of metadata from an organisation to a DAC to be harvested by the portal. This should be discussed by the portals development group and also a brief document outlining different use cases of; who and when people generate metadata (preferably as close to source as possible), where it is hosted, what the validation stage is and how it is harvested by the portal. There may be some information from the SeaDataNet project that could be used and the document should be alerted to the SeaDataNet technical task team when complete.

**AP. DC** to discuss with DC, RL,BS and draft a use case document to be sent to standards group and used in portal development workshop.

- Demonstration of tool

A demonstration of the prototype metadata generation tool was given and some comments received. The tool can generate individual records and batch ingest so it would be possible to build tools to allow current metadata holdings be imported to become MEDIN compliant. It was suggested that by filling out the forms the generated metadata should be compliant and the schematron tool may not be required for that use case. Trying some batch ingestion from other discovery records was thought to be a good test.

**AP. DL/BS** send around link for the tool to the MEDIN standards group by 8<sup>th</sup> May

**AP. ALL** comment on metadata generation tool by 1<sup>st</sup> June

**AP. RL** send an EDMED xml record to trail ingest into the tool.

- Discussion on keywords

The distributed document on keywords was introduced and discussed. In the metadata schema keywords would be entered using P021 (~380 terms) as it was used in the MDIP schema, is a managed standard and has a variety of mappings to other dictionaries. However a list was required that could be used to search for metadata records which had broader categories. It was decided that a new list at the level of the SeaDataNet P081 and MESH keywords should be devised that would be mapped to P031. P031 may need some modification to remove some badly defined terms. This new list should have a maximum of 15-20 keywords, be more 'marine' centric and should be used to categories throughout MEDIN. They should be distributed to the standards group and then commented on at the partners meeting on June 30<sup>th</sup>. It was also highlighted that when a data guideline was in the process of development it should not be signed off until there were parameter group terms in P021 that could be accommodated.

**AP. DC** to define geographic and thematic scope and then **MC** to derive a new high level search terms and send to standards WG for review.

**AP. MC** to send e-mail to UK rep on INSPIRE metadata drafting team to ask for clarification on using GEMET.

- Defining Coordinate Reference System Information

The paper on vertical coordinate reference systems was introduced and well received. It was thought that the EPSG for horizontal coordinates was very good and to develop a similar system to cover vertical CRS was a large task. There was debate on how useful vertical depth was for discovering data sets and it was felt that it was essential. A number of scenarios and issues were discussed including the conditional element for the vertical CRS and possibility of using an extra keyword vocabulary. It was decided that the Vertical extent information should be left in the MEDIN discovery standard however only filled in if the vertical coordinate reference system was registered in EPSG (it was recognised that there were only 2 vertical CRS currently registered but that these would fulfil most use cases). The SeaDataNet keyword list for vertical coverage should be made mandatory for the keywords in the schema until a full operational system could be employed. The possibility of MEDIN partners delivering a catalogue which could be served up should be investigated as a longer term solution.

**AP. BS** to amend MEDIN discovery standard to reflect above.

**AP. MC** to distribute the SeaDataNet vertical coverage terms

**AP. JR/OM** to discuss the possibility of providing a catalogue of vertical CRSs as a longer term solution which could be served by the NERC vocab server.

- Discussion on Unique Identifiers

The methodology for defining unique identifiers as provided by GeoData was discussed. GeoData had proposed the use of UUIDs which are generated using an algorithm for all data sets. It was felt that all DACs and most organisations already had unique identifiers for their data sets and using UUIDs as well would add a layer of complexity which is not helpful. It was felt that where data was submitted across more than 1 DAC then these UUIDs could be used however importantly this would have to be a managed process.

**AP. DACs** to use own unique identifiers with a name space for their organisation

**AP. DC** to organise a workshop to discuss how unique identifiers will be practically managed for data sets that are spread over more than one DAC using the GeoData guidance.

- Discussion on MD Scope Code

There is a desire to easily identify derived GIS data sets using the MEDIN discovery standard and it was not clear if this could be done using the element 'MD Scope Code'. It was discussed that this information should be included in the Element 'Data Format' and that there were a number of vocabularies that could be used so that searching could be done using this field in the portal.

**AP. BS** and **RH** to send **MC** their vocabularies that may be used for this element; **MC** to discuss with **RL** and serve using the NERC vocab server.

**AP. MC** to let a contract to draw up cross-platform guidance on data and information for GIS storage and transfer

- Discussion on Granularity

Thought that what was recorded in the minutes of the last meeting was a good start on defining levels of granularity. BODC will be considering what is a data set in due course and ERFF are also drawing up advice which MEDIN has provided information to.

**AP. DC** use existing guidance in minutes for CCO meeting and provide feedback to group

**AP. MC** to circulate ERFF guidance to group

**AP. RL** to feed back to group on BODC investigations

- Tools

There is a possibility of developing tools that may be used to assist people generating or transferring MEDIN discovery metadata. IT was felt that it would be a case by case basis and generic tools could probably not be built. It was felt that any tool to generate MEDIN metadata from Marine Recorder should not be funded by MEDIN.

#### 4) Data Guidelines

- Agreeing template

The data guideline for benthic sampling by grab or core was presented and explained that it should be used to provide data in a consistent format to DACs, so data within that theme can be exchanged easily and so it can be used as an appendix to contracts.

There were a number of comments including:

- Link to the SeaDataNet vocab lists rather than the NERC vocab server
- Remove section 1.4
- Add in a line for horizontal accuracy
- Remove the example line from the xls template and give a full spreadsheet example
- QC transcription error flags

The revised guideline should be sent to MMG and other contractors for comment before being signed off.

**AP. ALL** to send further comments to **MC** immediately

**AP. MC** to revise and then send to MMG for review

- Plans for development of other guidelines.

There was general discussion on how to develop the guidelines for other data types. For the geology and geophysics theme it was felt that the seabed mapping WG under UKMMAS should be made aware of the 5 guidelines available at present and be tasked

(via a contract) to devise a single one for use. Also, the latest version of the BGS guidelines should be put on the MEDIN website.

**AP. MC** to contact chair of UKMMAS seabed mapping WG and progress funding of work

Biology guidelines would largely be developed following the CCW/DASSH work and could start immediately. Some of these guidelines would not meet all needs such as the photographs and videos and this will have to be extended by MEDIN. Guidelines for fisheries should be discussed at the Fisheries DAC meeting

**AP. DC** to raise at next fisheries DAC meeting

It was felt that there were more suited guidelines available for the oceanography theme than the ICES ones presented and these should be explored further. There is a litter group under the UKMMAS which can be used to define a guideline for litter.

**AP. MC** to coordinate with UKMMAS Litter group and Marine Conservation Society in particular for development of a litter guideline and liaise with Lesley Rickards (BODC) to research oceanographic guidelines better.

- QA/QC data procedures

It was discussed if QA/QC information (eg number of transcriptions checked) should be stored in the data guideline. In the case of benthic macroinvertebrates it was felt that it should be and that this should be flagged at record level.

5) Plan of work for 2009-2010.

- Feedback from sponsors board

It was felt that there are resources available to carry out work this year and that if there are pieces of work that need to be done then these should be progressed quickly. The work plan for the MEDIN standard work was shown and amendments suggested.

**AP. MC** to send revised work plan to group for comment.

- Discussion on evaluation metadata and work for portal development

It was felt that evaluation metadata was needed but that it should not be formulated as an extension of the discovery standard. In the first instance the MOLES 3 format used in NERC Data Grid and the 'Observations and Measurements' work should be reviewed. For specific applications it may be possible to use evaluation metadata through WxS services and use cases particularly addressing the needs of the MCA should be written to give further indication of how the portal could meet various demands in the future. The link between discovery and evaluation metadata could be through the MEDIN discovery standard element 19 'Additional information source'. It could be explored how to do this using UKDMOS for the 30<sup>th</sup> Junes MEDIN partners meeting.

**AP. SW** to write a use case scenario of how the portal could meet various application demands at the evaluation metadata level such as those required by the MCA.

**AP. MC** to review current evaluation metadata usage and write document on how the implementation may be progressed.

## 6) World Register of Marine Species (WoRMS)

A contract has been let by the EA to UNICOMARINE to be able to export the UNICORN taxon list to be WoRMS compliant. Once this has been completed the Marine Recorder taxon list will be rerun against the WoRMS list and then MEDIN will work with JNCC to ensure all the Marine Recorder taxon are registered in WoRMS. NHM should also be updated at regular intervals.

It was agreed that the WoRMS taxon list should be promoted as a MEDIN standard.

**AP. MC** to move the WoRMS taxon list from the 'Other available standards' to the 'MEDIN approved standards' webpages.

## 7) MEDIN webpages on marine standards

The MEDIN webpages on standards were shown and some comments received. Further comments on the standards pages should be sent to the MEDIN team and a review of the whole website made as some felt that the banner was too large.

**AP. DC** to organize a review of the whole MEDIN website

**AP. ALL** Send comments on the MEDIN standards to MC

## 8) AOB

James Rappaport, may organize a meeting between UKSDI, MEDIN and UKGEMINI. The proposed post should be termed 'portal population'.

## 9) Date and location of next meeting

Next meeting in mid September but two telephone conferences with key personnel should be arranged before then to review progress.