World Meteorological Organization

IPET-MDRD TASK TEAM ON THE APPLICATION OF METADATA
TOKYO, JAPAN, 25-28 JUNE 2013

TT-ApMD-2/ FINAL REPORT

Date:	15 VII 2013
Original Language:	English
Status:	Final

TASK TEAM ON THE APPLICATION OF METADATA - SECOND MEETING - FINAL REPORT

DISCLAIMER

Regulation 43

Recommendations of working groups shall have no status within the Organization until they have been approved by the responsible constituent body. In the case of joint working groups the recommendations must be concurred with by the presidents of the constituent bodies concerned before being submitted to the designated constituent body.

Regulation 44

In the case of a recommendation made by a working group between sessions of the responsible constituent body, either in a session of a working group or by correspondence, the president of the body may, as an exceptional measure, approve the recommendation on behalf of the constituent body when the matter is, in his opinion, urgent, and does not appear to imply new obligations for Members. He may then submit this recommendation for adoption by the Executive Council or to the President of the Organization for action in accordance with Regulation 9(5).

_____

TASK TEAM ON THE APLICATION OF METADATA, SECOND MEETING, 25-28 JUNE 2013.
FINAL REPORT

1 ORGANIZATION OF THE MEETING

1.1 Opening of the meeting

1. Mr Hiroyoshi Nakayana, Director of the Information and Telecommunication Division, welcomed the members of TT-APMD to the headquarters of the Japan Meteorological Agency. He emphasized the importance of WIS Discovery Metadata to the operation of the WMO Information System.

2. Dr Foreman, Chief of Data Representation, Metadata and Monitoring of WMO, representing the Secretary General, Director of Observations, and the Director of WIS of WMO also welcomed participants to the meeting. He noted that the two objectives of the meeting were to provide guidance on writing WIS Discovery Metadata records that could be used to create supporting documentation by the middle of August, and to make sure that it was possible to provide from metadata records the information needed to fulfil the use cases for WMO No. 9 Volume C1.

1.2 Approval of the agenda

3. The agenda was approved. The list of participants is at Annex A, and the agenda at Annex B.

1.3 Working arrangements

4. The meeting agreed its working times.

2 VALIDATION OF WMO CORE PROFILE 1.3

2.1 Key points of Core Profile 1.3

5. Dr Toyoda outlined the key features of the WMO Core Profile version 1.3. The WMO Core Profile is a subset of the complete ISO 19115 standard that limits the items that have to be specified to those that are essential for discovering information in the WIS. The Core Profile has been designed to allow automated validation of the contents, and has used controlled vocabularies in many places to make it easier for the many authors of metadata in the WMO communities to create consistent metadata.

6. In defining the WMO Core Profile version 1.3, all mandatory components of the ISO 19115 were retained. Additional items were made mandatory for all records, and further items were made mandatory for information that is intended for global exchange.

Decision: 1. TT-ApMD recommended that all metadata records should indicate the data policy (WMOEssential, WMOAdditional, or WMOOther), not only those records for global exchange. The guidance needs to state that if there is no constraint listed, users are likely to assume that there are no limitations on the use of the data, and that authors of metadata records for datasets that are not exchanged on the GTS who feel unable to use the term "WMOEssential" and who place no constraints on the use of the data should include the text used by the INSPIRE rules to indicate that there are no restrictions on the use of the data ("No limitation").

7. The guidance documentation must give examples of metadata for a variety of methods of data distribution (for example: link to data, online instructions to access data, or contact (post, email, other) to request data).

2.2 Automated validation

8. Mr Kozimor outlined the validation tools developed by NOAA and JMA. These check for well-formed XML, for conformance with ISO 19139, and with the WMO Core Profile v1.3.

9.   The "rubric" report evaluates conformance with WMO Core Profile 1.3, and provides a report against each requirement and provides guidance on how to improve the metadata for those items that are not completed correctly.

10.   The schematron allows the metadata record to be checked for conformance with coding rules.

11.   Schematron and rubric tests are available at www.ngdc.noaa.gov/docucomp/validationServicesWmo. It would be possible to validate groups of metadata records in a way similar to www.ngdc.noaa.gov/docucomp/collectionServices.

12.   In its own internal metadata management procedures, NOAA identifies common sequences in collections of metadata records, such as contact details, and stores these as reusable components that are subsequently merged with the records as they are delivered to users. This makes it simpler to maintain changes to those components.

13.   Mr Kozimor also presented statistics over time showing progress in the conformance to another ISO profile for a different set of metadata. TT-ApMD considered that this kind of information could be used to populate reports showing improvements on the WIS infrastructure, which could be extremely valuable for raising interest of managers for the WIS, complementing other statistics, for example those on search.

14.   TT-ApMD recommended to add to the monitoring system of the WIS a service offering continous health checks (daily, weekly) on the metadata validity and compliancy. This could be performed by one GISC (GISC Washington).

Action: 1.   (Secretariat, September 2013) TT-ApMD will propose to the proposed workshop on WIS monitoring that a metadata health check service would be useful as one of the service of the WIS monitoring infrastructure.

Action: 2.   (Toyoda, September 2013) Dr Toyoda will develop additional validation rules to allow metadata to be checked for recommended content as defined in the guidance developed by the meeting.

15.   Although it will not be possible to validate the contents of the abstracts of metadata records, the team has to provide strong guidance on how to write useful abstracts in Standard English, including examples of good practice.

2.3   Publication of validation tool

Decision: 2.   NOAA is willing to host the validation tool to assist WMO Members validate their metadata. In theory it would be possible to harvest the WIS catalogue and produce statistics on the scores obtained by running the validation tools. A logical approach would be for the validation tool to access GISC Washington.

Decision: 3.   The NOAA validation page will be used to provide a metadata validation service for metadata authors.

Action: 3.   (Kozimor, December 2013) The following changes are needed to the validation tool: checks should be against version 1.3, the missing checks need to be added, additional "spirals" are needed (for example, recommended practices), validation for collections of WIS metadata records, hosting of validation other than at NOAA EMMA (so that WIS centres can include such monitoring in their internal processes), support to metadata authors including wiki documentation and extended recommendations on how to respond to findings by the rubric.

3   PUBLICATION OF GUIDANCE

3.1   Categories of guidance needed

16.   TT-ApMD agreed to work with the list of areas of guidance proposed, with the addition of Space Weather. Topics that were not discussed during the meeting would be completed after the meeting, with each topic being assigned to a team member to provide the guidance.

Action: 4. (Toyoda, September 2013) IPET-DRMM will be asked to define the set of physical parameters that should be associated with each TAC and BUFR guidance area and their sub-divisions (see Annex C). This will allow the team to complete the examples.

17.   SRU allows search by keyword. Separating out physical parameters into keywords will make them more accessible to search using SRU. Keywords could to be used to modify the ranking of the metadata records returned in search results. They could be used to in such a way that records with the search keywords would be returned prioritized and returned first.

Decision: 4. TT-ApMD decided that physical parameters should be recorded in keywords.

18.   Analysis of the search requests for data at EUMETSAT shows that most first time users discover the data through the browsing categories and then transfer to full text search. This is not possible using the current WIS Discovery metadata records as no recommended taxonomy/categorization has been defined in the WMO Core Profile. This reopens the issue discussed at IPET-DRM-1.

Action: 5. (Toyoda, October 2013) TT-ApMD will ask IPET-MDRD to consider introducing a mechanism that would allow facetted searches where the catalogue can be browse following different themes (e.g. by type of products, by earth science domain, by originating centres, etc).

3.2   Consistency

3.2.1 Synoptic information

19.   CMA had provided an analysis of metadata at TT-ApMD-1. This was based on version 1.2 of the Core Profile. The analysis had been repeated to compare current metadata from GISCs Offenbach, Beijing, Tokyo and the start-up metadata. The metadata from each centre are valid (as defined by the validator), but they differ in style and content. In particular, it is not always possible to find the information needed to manage the WIS.

20.   The first comparison was between the metadata and Volume C1. The start-up metadata and metadata from Beijing did not include the WMO Region.

21.   There are still differences between how the centres handle the update frequency (only one uses userMaintenanceUpdateFrequency in the PnYnMnDnHnMnS form).

22.   Specification of station location differs.

23.   Consolidated recommendation:

- Metadata contact – use pointOfContact
- Title – include a basic description of the dataset that is meaningful to non-specialists and following the TT-ApMD recommendation for title
- Abstract – include a description in standard english meaningful to non-specialists and following the TT-ApMD guidelines for abstracts. Include the data type, centre creating reports, region, station names
- Data contact – use originator
- resourceMaintenance – TT-ApMD decided not to provide guidance
- resourceFormat – use, for example, "FM 12 SYNOP" and "XII EXT" forms
- keywords for elements – place each element in a separate keyword
- Specification of the time of reports – use keywords T00Z etc to allow datasets that include data for specific times to be identified.
- Place keywords – use station identifiers, station name, station country (and use a thesaurus specified for this purpose to allow them to be identified)
- Do NOT include the lat/long in the keywords – maintenance is too difficult
- Distribution format: use FM12 SYNOP as format, XII EXT as version
- Distribution role – pointofContact
- Transfer options – current guidance to GISCs is that the GISC search detects data for global exchange and generate the URL for downloading from the cache – which means that this URL need not be in the metadata – thus we only need the URL of any non-GISC source of the data.

However, the link to the cache of the copy of the data on the Principal GISC is useful if the data are found using GEOSS (for example).

Decision: 5. When Station identifiers are included as keywords with VolA as the thesaurus for the station identifier and station name, TT-ApMD recommended that "common names" should also be provided for non-specialist users (for example "Moscow").

Decision: 6. TT-ApMD concluded that a URL that retrieves the data from the cache of the Principal GISC should be included in metadata records corresponding to GTS traffic (if technically possible), and that the metadata records may also contain URLs that point to additional options for accessing the data.

24. It is good practice to describe gmd:online/*gmd:function (codelist CI_OnLineFunctionCode) to differentiate a direct link to a resource (code "download") from other functions such as online request for information (code "order").

Decision: 7. TT-ApMD recommended that a link to download the data should be provided whenever possible, but noted that there is no limit to the number of options that can be defined with gmd:online.

Decision: 8. Users will need a point of contact if they experience problems, so distributorTransferOptions is the preferred place to specify the distributor options.

Action: 6. (Toyoda, July 2013). The changes in recommendations need to be passed to ET-WISC for information.

Decision: 9. Upper air, ship and buoy reports share many characteristics with surface observations and should be based on that template. However, where the sensor platform is mobile (e.g. AMDAR, ships or buoys) the bounding box will describe the maximum *reasonable* extent within which the platform may operate.

3.3 Satellite information

25. Mr Aubert presented a metadata record for a satellite product produced in EUMETSAT. He presented the different parts that need to be filled by the metadata creator in order to create a WMO Core Profile 1.3 compliant record describing satellite data

26. Mr Aubert stressed that metadata record editorial parts such as title and abstract are not only intended to be indexed by the GISCs search engines but are also used to describe the discovered products in the search results pages. It is therefore extremely important to have for each record, information that contains standard english and that can be understood by specialist and non-specialist users looking for products. He presented in his example, guidelines to create an abstract for satellite products. Guidelines should be produced on to write Title and Abstract in included in the outcome documentation of the TT-ApMD.

Decision: 10. An example of data policies for data not intended for Global exchange, therefore without the need to use one of the resolution 40 data policies (WMOAdditional and WMOEssential), was presented. In this example WMOOther was added in a <gmd:otherConstraints> element in addition to the <useLimitation>, <accessConstraints> and <useConstraints> element containing the DCPC data policy information to signify that this data policy is valid in the WIS context. A metadata record could contain multiple data policies (<gmd:resourceConstraints>) and DCPCs would like to have a way to indicate which data policies are valid for the WIS communities and therefore can be shown by the GISCs (the others can be ignored).

Decision: 11. TT-ApMD concluded that it would recommended WMOOther to be used for any data, not just those intended for global exchange.

Action: 7. (Toyoda, December 2013). The text describing the "WMOOther" data policy needs to be changed so that it may be applied to any data made available through the WIS, and not just those data for global exchange. This should be achieved through the "fast track" procedure by IPET-MDRD.

27.   Use of hierarchyLevel that can be used to create a facetted search was presented. Mr Aubert noted that there are no clear guidelines on which categories to use and therefore used the taxonomy from the EUMETSAT product catalogue. He noted that the facetted search of NASA GCMD catalogue could be used as a starting for the WIS needs. TT-ApMD noted that MD_hierarchyLevel is still needed in version 1.3, but in the next version of ISO 19115-1 it will be replaced by MD_Scope that will be more versatile.

28.   Mr Kozimor reported that a thesaurus, such as GCMD, contains vocabularies that correspond to many keyword types. In order to describe keyword types, the keywords have to be grouped into separated MD_Keyword blocks, which causes problems with validation against the reuirement 8.2.3 of the WMO Core Metatada Profile 1.3. Such thesauri are not necessarily subdivided into ISO-MD_KeywordType. In consequence, TT-ApMD recommended that the thesaurus title should supplemented with keyword type, such as "Global Master Change Directory (for theme)" or "Global Master Change Directory (for place)" to avoid the problem with validation.

29.   The geographical information required by the WMO Core Profile was presented and Mr Heene indicated that at the next version of ISO 19115-1 it will be possible to provide a named geographic extent as well as the bounding box. It will then be a good practice to add a text-based name for the geographic extent.

Decision: 12. TT-ApMD concluded that the use of thesauri was appropriate to categorize keywords.

Decision: 13. For the keyword components of the metadata record, only the keywords from WMO_CategoryCode had been used in the example. TT-ApMD remarked that additional keywords could be provided. The guidance material should make it clear that authors can use additional keywords that are not in the standard thesaurus.

30.   Mr Aubert asked if there was any process in place for updating the documentation created by TT-ApMD when WMO changes are made.

Decision: 14. TT-ApMD confirmed that the Guidance documentation needs to be updated when changes are made to the WMO Core Profile specifications. This needs to be part of the "project" to update the Core Profile.

Action: 8. (Toyoda, August 2013). TT-ApMD will ask IPET-DRMM to prepare a list of physical parameters that can be used as keywords within metadata records.

Decision: 15. TT-ApMD recommended to IPET-DRMM the list of parameters contained in OSCAR (as presented by Dr Toyoda and collected by DWD) as a tool to creating standard names for physical properties. These lists could be used as the starting point for the reference parameter list for metadata.

31.   For non-GTS data where individual datasets, rather than collections of datasets, are described, such as numerical weather prediction, satellite products, and climatology products, temporal extent provides the start date and end date of the data or product, and the date the dataset was produced should be recorded in the publication date. Normal WIS practice would be to produce a single metadata record that described the complete set of datasets that points to a retrieval service at the DCPC. In that case the start of the temporal extent is the date of the first dataset in the series and the end is the date of the last dataset in the series; the frequency of producing datasets within the series is recorded using gml:AbstractTimePrimitive using the description "Production Frequency" (see ISO 8601 for how to represent repetition).

Action: 9. (Toyoda, August 2013). Dr Toyoda will provide guidance on how to specify the different types of time in metadata records.

Action: 10. (Aubert, August 2013) Recording geological time may be achieved by using negative times. Dr Aubert will find out whether controlled vocabularies can be used to denote the limits of vertical or temporal extents, and provide appropriate advice in the guidance.

Decision: 16. Vertical extents should be measured either in metres (height above mean sea level, either geometric height or geopotential height), or in hPa where pressure is the vertical coordinate. The guidance needs to recommend text to use for the vertical Coordinate Reference System.

Decision: 17. Limitations on the times for which information is appropriate for use (for example the validity period of a TAF) should be coded in the useLimitation using the period notation (e.g. "PT12H"). This must be included in the guidance.

3.4   Climate information

Action: 11. (Wilson, August) Mr Wilson will create an example of climate metadata before the end of August.

Decision: 18. TT-ApMD will use the categories in the CDMS (Climate Data Management System) specifications for discovery metadata to structure its own documentation.

32.   Metadata records with human-readable title and abstract will contain words that are commonly ignored by search engines ("stop words"), and some searches produce different results with and without these words included in the search term. TT-ApMD recommended that titles and abstracts should be written in standard English, which means that developers of applications that search the WIS catalogue must design their search to work correctly if these stop words are present.

3.5   GAW

Decision: 19. When it is not possible to provide data to complete a component of a metadata record, the entry should be represented in the XML by gco:nilReason. However, for the mandatory elements of a WMO Core Profile metadata record contents should always be provided.

Decision: 20. GAW uses three letter codes as station identifiers. These should be handled in the same way as for surface observations – that is, two sets of keywords, one for the station identifier and the other for the station name. The latitude and longitude for each station are not included (unless the record is for a single station, in which case the location is given by the bounding box).

33.   GAW will propose a new Category Code keyword atmosphericComposition to be considered by the fast track procedure.

Action: 12. (Toyoda, August 2013). The list of physical parameters will need to include the chemical substances and aerosol types. Dr Toyoda will ask GAW to propose a list of these.

34.   GAW (and CHy) would like to be able to issue digital object identifiers (DOIs) for their data streams. Metadata records already provide a Uniform Resource Identifier (URI), the fileIdentifier. In addition to the URI, a DOI needs a published method of resolving the DOI to the underlying resource.

35.   Although there is no requirement for WIS to retain old versions of metadata records, the emerging requirement for Digital Object Identifiers means that (for metadata records other than those for GTS traffic) any change to the metadata records that impacts on the description of the data should result in a change to the fileIdentifier (examples of changes that would not require a change are the correction of spelling errors, introduction of multi-lingual abstracts, updating contact details or updating distribution information).

Action: 13. (Toyoda, September 2013) Dr Toyoda will modify an existing GAW metadata record to provide an example for others to use.

3.6   Hydrology

Action: 14. (Secretariat, August 2013). The Secretariat will ask the Commission for Hydrology to nominate an expert to work with a member of IPET-MDRD to prepare guidance specific to hydrological information.

3.7 Documents and Standards

36.　Because resolution 40 does not cover documents, the INSPIRE term for "No limitation" should be used to describe access constraints.

37.　Documents frequently have abstracts and keywords assigned by professional librarians. These should be used if available. Additions to the abstract may be appropriate to ensure that search terms used in the context of WIS will discover the document. Further, the ISO entries for Edition and ISBN are relevant to documents, and should be used to hold the relevant information.

38.　TT-ApMD recognised that harvesting existing library catalogues would be useful, but this would only be possible if it the relevant ISO fields could be extracted from Dublin Core metadata records.

3.8 Products (model analyses, forecasts, ensembles and other products)

39.　Mr Heene described a metadata record for a GRIB file produced by Offenbach. The abstract was generated automatically from the GTS heading for the file.

40.　Although, for products exchanged on the GTS, much of the information required for the abstract can be created from the GTS heading, this heading is not available for products not exchanged on the GTS. Also, for data generated by numerical models, users will need more information about the model used (and that the data being described are a product from a model), and about fields that cannot be described by the GTS headings (such as potential vorticity).

41.　Two examples are appropriate for GRIB datasets. One should be for a generic GRIB dataset, and the other for GRIB datasets exchanged on the GTS.

42.　The generic GRIB example should include:

- Model (or process) name

- Type of model (eg deterministic, ensemble, spectral, grid and type of grid, number of levels, resolution, etc)

- Parameter(s), level(s), area, reference time, forecast time(s), analysis time(s).

43.　For GRIB fields, keywords describing parameters could be derived from the GRIB tables and the GRIB code table used as the thesaurus- these would need to refer to the GRIB code entry.

44.　Mr Heene also showed an example from a regional climate model.

45.　Dr Toyoda described the metadata records for products managed by GISC Tokyo. These were generated from the Tokyo internal product catalogue that contains more information about the product than Volume C1. He noted that the majority of entries in Vol C1 correspond to GRIB files from models that are created by a small number of centres. Non-specialist users might be helped if the model itself were recorded in a metadata record that was a parent to the records from the specific metadata records for each file distributed on the GTS – this would be a use of parentIdentifier. However, the most appropriate linkage would be from the summary to the detailed records, but this capability is not available in the ISO 19115:2003 version of the standard. It could be considered for version 2.0 of the WMO Core Metadata Profile.

46.　Dr Toyoda suggested that vertical levels could be described by keywords of type "stratum" drawn from GRIB code table 4.5. Forecast times could be denoted by keywords with type "temporal" with a thesaurus called "forecast time".

47.　TT-AvXML was creating a web register to hold BUFR tables and other reference code lists that could also include GRIB. TT-AvXML will use this to resolve URIs for each of the definitions used within XML messages, but the resource will be available for metadata records to use.

48.	TT-ApMD recognised that there is no standard convention for the sign for ocean depths. TT-ApMD recommended that vertical extents should be described so that distance above mean sea level is positive and distance below mean sea level is negative.

Action: 15. (Besprozvannykh, August 2013) Confirm that the ISO standards permit negative vertical positions in an extent.

## 3.9	Extensions to code lists

49.	Mr Wilson identified issues associated with code lists:

- Need to update code lists (License information, CategoryCode, the descriptions in MD_RestrictionCode)

- Guidance should tell users to use the 1.3 versions of code list tables for all the tables listed in Part 2 of the Core Profile.

- What to do if code list does not describe the data

- What to do if your data fits into more than one category (list them all – for example coupled ocean-atmosphere model output might be listed as both meteorology and oceanography)

Action: 16. (Secretariat, August 2013) The documentation of the WMO_CategoryCode, capitalization of the description text differs between the code lists in the reference document and the xml file. This needs to be corrected and the word "draft" removed from the version identification.

50.	Metadata guidelines need to explain to users the procedure for updating the code lists.

Decision: 21. The WMO_CategoryCode list needs to be extended and rationalised so that the current meteorology, climate and water disciplines can be described. The granularity of the current entries differs widely; a uniform approach to granularity is needed. This work must be completed before version 2.0 of the Core Profile is defined.

Action: 17. (Toyoda, September 2013) Code lists are required for physical quantities that have to be described, for the processing of information (such as forecast, analysis), and for observing methods (such as radar, satellite). TT-ApMD will ask IPET-MDRD, IPET-MDRD and TT-WMD to collaborate to produce these lists.

Decision: 22. TT-ApMD recommended to IPET-MDRD that in future versions of the WMO Core Profile the data policy and GTS Priorities should be referenced as code lists rather than as free text in the XML.

Action: 18. (Toyoda, August 2013). TT-ApMD noted that the use of the scope of distribution keyword (e.g. GlobalExchange) is optional for GTS traffic. It recommended to IPET-MDRD that the use of the keyword be made mandatory.

Action: 19. (Secretariat, August 2013). TT-ApMD asked that the Secretariat publish the date of publication of code list versions so that this can be used in the metadata records

51.	When a code list element is specified in the xml, the text of the code list value should appear as both an attribute value and as the text associated with that attribute (this is consistent with INSPIRE recommended practices). This should be checked by the validation tools.

Decision: 23. TT-ApMD recommended that all dates (as well as times) have a time zone associated with them. This would normally be UTC (e.g. 2013-06-29Z).

52.	Recommendation 14 of IPET-MDI-1 (2010) Final Report recommends that identifiers should be added for every bounding extent: boundingGeographicBoundingBox, boundingTemporalExtent, boundingVerticalExtent. These are most useful for metadata records that contain multiple domains. The bounding box that encloses all the data contained in the dataset should use these identifiers. Other bounding boxes should have identifiers that precede the standard identifier text with a string that distinguishes between them.

Decision: 24. TT-ApMD guidance will recommend that gml identifiers (gmd:id) are used in metadata records that contain multiple bounding boxes of the same type to label the bounding box that best describes the entire content of the dataset. Metadata records that contain only one bounding box can omit the identifier.

Decision: 25. Within WIS Discovery Metadata, the only requirement for using a gmd:id attribute will be to identify the bounding box that summarizes the total extent of the dataset being described if the metadata record contains multiple bounding boxes. Users may include gmd:id in other contexts, but these have no significance.

Decision: 26. For GTS bulletins the date that the GTS bulletins started or when the content of the bulletin changed, should be recorded in the citation as publication date and revision date respectively.

4   TRAINING PLAN FOR WMO CORE METADATA PROFILE

4.1   Formal documents

53.   Dr Toyoda introduced the presentation prepared by Mr Colavecchia. This described the procedures used in Environment Canada to produce metadata records for both their internal and external datasets. These procedures are supported by a spreadsheet that allows the key aspects of the metadata to be captured.

54.   Several centres are using a similar approach based on spreadsheets.

55.   Experience at DWD is that users take about 30 minutes to create a metadata record using a spreadsheet or metadata editor. That means that it is necessary to automate production of metadata records – especially those corresponding to model outputs.

56.   Rather than concentrating on manual modification of metadata records describing GTS information, the guidance needs to be aimed at generating such records automatically. The principal aspects of such guidance are, therefore: procedures to notifiy the GISC of changes to the content of datasets; guidance to GISCs on the structure and content required of these metadata records ("templates"); and standard keywords to use. The GISCs should re-create the metadata records for standard data types, and ask users only to check the contact information and the data policy.

Decision: 27.   TT-ApMD will create standard templates for representative GTS reports and DCPC products. This will require templates for record types that are not in the planned guidance (such as warnings or tsunami reports).

Decision: 28. 4.1.6   Following the definition of the standard templates the GISCs and WIMMS will recreate the metadata records for GTS data.

Decision: 29. In parallel with recreating the GTS records, all Members will be asked by their Principal GISCs to check and correct the station list, bounding box, contact and data policy elements of their metadata.

57.   Detailed guidance should be aimed at those centres creating metadata records for non GTS datasets (that is DCPC and NC products), and because of the variety of such information the guidance needs to concentrate on the thought processes needed to create good metadata. Nevertheless templates describing well identified DCPC and NC products such as climate records, satellite products, forecast data, etc) will be created by TT-ApMD and added in the WIS metadata portfolio.

58.   Good examples of metadata records should be identified, initially by TT-ApMD, and the examples published so that those generating metadata for these or new types of information, can see what makes a good metadata record.

59.   In summary, TT-ApMD decided that the following guidance materials were required:

- Examples
  - Spreadsheets – for users (each element labelled so metadata experts can link to xml sections)
  - Completed XML for the spreadsheet example (with cross-references in the documentation to help metadata experts match the XML to the spreadsheet)
  - XML template with "fill in here" for elements that change between instances
- Validation tools
- Presentations – one for users, and one for metadata experts, with completed "speakers notes" so that consistent presentations can be given using the material.
- Documentation of guidance – wiki (or other web) that can be used to produce "paper" document

5 CREATING VOL C1 FROM METADATA CATALOGUE

5.1 Missing information

60. Mr Wilson described the test performed by GISC Melbourne to determine the obstacles to collecting the information required for Volume C1 from the WIS Metadata Catalogue. Country was derived from the Contact information. The date was retrieved from the metadata update date. The Code Form was taken from the format and version. There were problems with reconstructing the Content and Remark fields. The Region and RTH information was not available from the metadata records.

61. A survey of operational GISCs showed that most centres did not think that recreating Volume C1 in its present form was needed.

62. Under WIS, the meaning of "Region" becomes ambiguous. For telecommunications roles, the relevant meaning is the Principal GISC responsible for the centre. For data content matters, it is the Regional Association.

Decision: 30. TT-ApMD decided that for GTS data the Regional Association in which the data originate must be recorded in the metadata as the AdministrativeArea of the originator contact information.

63. Volume C1 does not identify the originator of the data, and the "centre" is that centre injecting the information onto the GTS. This is not available in the WIS metadata.

Decision: 31. TT-ApMD decided that both the originator (the centre creating the data) and the distributor (the centre placing the data onto the GTS) should be recorded through the contact information. The role should be used to achieve this.

64. To extract the Contents and Remarks fields, it is necessary to find the text following the "--- WMO No.9 – Volume C1 'Remarks' field:" text in the abstract.

Decision: 32. TT-ApMD decided that keywords from a thesaurus "VolA" should be used to to identify station identifier keywords. Similarly, reporting times should use a thesaurus "ReportingTimes" with the pattern TxxZ for reports issued at specific time.

5.2 Governance of updates to the metadata catalogue

65. Current regulations require changes, additions or deletions of GTS bulletins to be notified two months in advance. This notification period is needed to allow time for routing changes to be implemented. In the WIS, such a notification period is possible for bulletins that are being added, but changes to metadata corresponding to changes in bulletin contents or to deletions of bulletins cannot be implemented until the day of the change because of the requirement that a GTS bulletin can only have one metadata record.

Action: 20. (Toyoda, July 2013). ET-WISC will be asked to confirm that their intention is that advance notification (currently two months for the GTS) is only given for additions of GTS bulletins.

66.   If that is not the case, then it will be necessary to introduce versioning of metadata records to allow more than one record corresponding to a GTS bulletin to be retained in the WIS catalogue (using an identifier such as urn:x-wmo:int.wmo.wis:date_of_implementation:AHL to distinguish it from the operational version that does not have a date in the identifier – this would also allow an archive to be retained if required).

Decision: 33. TT-ApMD recommended that the formal procedure for Permanent Representatives to notify changes to the content of GTS traffic and the data policy to be applied to information should be through providing updated metadata records to the Principal GISC. This needs to be approved by CBS and EC.

67.   Dr Toyoda reported that, at GISC Tokyo, 13% of bulletin headings circulating on the GTS did not have corresponding WIS Discovery Metadata records. Most implementation so f the GIAC catalogue require metadata to manage the flow of data. There is a need to help data custodians to create and maintain metadata. Given the high proportion of GTS traffic without valid metadata records, TT-ApMD considered it appropriate for an automated process to notify data providers when data were received on the GTS that did not have an associated metadata record. JMA was willing to provide such a service.

6   ANY OTHER BUSINESS

68.   There was no other business.

7   Closure of meeting

69.   The meeting closed at 1813 on 28 June 2013

ACTION AND DECISION SUMMARY

ACTIONS

Action: 1. (Secretariat, September 2013) TT-ApMD will propose to the proposed workshop on WIS monitoring that a metadata health check service would be useful as one of the service of the WIS monitoring infrastructure.

Action: 2. (Toyoda, September 2013) Dr Toyoda will develop additional validation rules to allow metadata to be checked for recommended content as defined in the guidance developed by the meeting.

Action: 3. (Kozimor, December 2013) The following changes are needed to the validation tool: checks should be against version 1.3, the missing checks need to be added, additional "spirals" are needed (for example, recommended practices), validation for collections of WIS metadata records, hosting of validation other than at NOAA EMMA (so that WIS centres can include such monitoring in their internal processes), support to metadata authors including wiki documentation and extended recommendations on how to respond to findings by the rubric.

Action: 4. (Toyoda, September 2013) IPET-DRMM will be asked to define the set of physical parameters that should be associated with each TAC and BUFR guidance area and their sub-divisions (see Annex C). This will allow the team to complete the examples.

Action: 5. (Toyoda, October 2013) TT-ApMD will ask IPET-MDRD to consider introducing a mechanism that would allow facetted searches where the catalogue can be browse following different themes (e.g. by type of products, by earth science domain, by originating centres, etc).

Action: 6. (Toyoda, July 2013). The changes in recommendations need to be passed to ET-WISC for information.

Action: 7. (Toyoda, December 2013). The text describing the "WMOOther" data policy needs to be changed so that it may be applied to any data made available through the WIS, and not just those data for global exchange. This should be achieved through the "fast track" procedure by IPET-MDRD.

Action: 8. (Toyoda, August 2013). TT-ApMD will ask IPET-DRMM to prepare a list of physical parameters that can be used as keywords within metadata records.

Action: 9. (Toyoda, August 2013). Dr Toyoda will provide guidance on how to specify the different types of time in metadata records.

Action: 10. (Aubert, August 2013) Recording geological time may be achieved by using negative times. Dr Aubert will find out whether controlled vocabularies can be used to denote the limits of vertical or temporal extents, and provide appropriate advice in the guidance.

Action: 11. (Wilson, August) Mr Wilson will create an example of climate metadata before the end of August.

Action: 12. (Toyoda, August 2013). The list of physical parameters will need to include the chemical substances and aerosol types. Dr Toyoda will ask GAW to propose a list of these.

Action: 13. (Toyoda, September 2013) Dr Toyoda will modify an existing GAW metadata record to provide an example for others to use.

Action: 14. (Secretariat, August 2013). The Secretariat will ask the Commission for Hydrology to nominate an expert to work with a member of IPET-MDRD to prepare guidance specific to hydrological information.

Action: 15. (Besprozvannykh, August 2013) Confirm that the ISO standards permit negative vertical positions in an extent.

Action: 16. (Secretariat, August 2013) The documentation of the WMO_CategoryCode, capitalization of the description text differs between the code lists in the reference document and the xml file. This needs to be corrected and the word "draft" removed from the version identification.

Action: 17. (Toyoda, September 2013) Code lists are required for physical quantities that have to be described, for the processing of information (such as forecast, analysis), and for observing methods (such as radar, satellite). TT-ApMD will ask IPET-MDRD, IPET-MDRD and TT-WMD to collaborate to produce these lists.

Action: 18. (Toyoda, August 2013). TT-ApMD noted that the use of the scope of distribution keyword (e.g. GlobalExchange) is optional for GTS traffic. It recommended to IPET-MDRD that the use of the keyword be made mandatory.
Action: 19. (Secretariat, August 2013). TT-ApMD asked that the Secretariat publish the date of publication of code list versions so that this can be used in the metadata records
Action: 20. (Toyoda, July 2013). ET-WISC will be asked to confirm that their intention is that advance notification (currently two months for the GTS) is only given for additions of GTS bulletins.

DECISIONS

Decision: 1. TT-ApMD recommended that all metadata records should indicate the data policy (WMOEssential, WMOAdditional, or WMOOther), not only those records for global exchange. The guidance needs to state that if there is no constraint listed, users are likely to assume that there are no limitations on the use of the data, and that authors of metadata records for datasets that are not exchanged on the GTS who feel unable to use the term "WMOEssential" and who place no constraints on the use of the data should include the text used by the INSPIRE rules to indicate that there are no restrictions on the use of the data ("No limitation").
Decision: 2. NOAA is willing to host the validation tool to assist WMO Members validate their metadata. In theory it would be possible to harvest the WIS catalogue and produce statistics on the scores obtained by running the validation tools. A logical approach would be for the validation tool to access GISC Washington.
Decision: 3. The NOAA validation page will be used to provide a metadata validation service for metadata authors.
Decision: 4. TT-ApMD decided that physical parameters should be recorded in keywords.
Decision: 5. When Station identifiers are included as keywords with VolA as the thesaurus for the station identifier and station name, TT-ApMD recommended that "common names" should also be provided for non-specialist users (for example "Moscow").
Decision: 6. TT-ApMD concluded that a URL that retrieves the data from the cache of the Principal GISC should be included in metadata records corresponding to GTS traffic (if technically possible), and that the metadata records may also contain URLs that point to additional options for accessing the data.
Decision: 7. TT-ApMD recommended that a link to download the data should be provided whenever possible, but noted that there is no limit to the number of options that can be defined with gmd:online.
Decision: 8. Users will need a point of contact if they experience problems, so distributorTransferOptions is the preferred place to specify the distributor options.
Decision: 9. Upper air, ship and buoy reports share many characteristics with surface observations and should be based on that template. However, where the sensor platform is mobile (e.g. AMDAR, ships or buoys) the bounding box will describe the maximum *reasonable* extent within which the platform may operate.
Decision: 10. An example of data policies for data not intended for Global exchange, therefore without the need to use one of the resolution 40 data policies (WMOAdditional and WMOEssential), was presented. In this example WMOOther was added in a <gmd:otherConstraints> element in addition to the <useLimitation>, <accessConstraints> and <useConstraints> element containing the DCPC data policy information to signify that this data policy is valid in the WIS context. A metadata record could contain multiple data policies (<gmd:resourceConstraints>) and DCPCs would like to have a way to indicate which data policies are valid for the WIS communities and therefore can be shown by the GISCs (the others can be ignored).
Decision: 11. TT-ApMD concluded that it would recommended WMOOther to be used for any data, not just those intended for global exchange.
Decision: 12. TT-ApMD concluded that the use of thesauri was appropriate to categorize keywords.

Decision: 13. For the keyword components of the metadata record, only the keywords from WMO_CategoryCode had been used in the example. TT-ApMD remarked that additional keywords could be provided. The guidance material should make it clear that authors can use additional keywords that are not in the standard thesaurus.

Decision: 14. TT-ApMD confirmed that the Guidance documentation needs to be updated when changes are made to the WMO Core Profile specifications. This needs to be part of the "project" to update the Core Profile.

Decision: 15. TT-ApMD recommended to IPET-DRMM the list of parameters contained in OSCAR (as presented by Dr Toyoda and collected by DWD) as a tool to creating standard names for physical properties. These lists could be used as the starting point for the reference parameter list for metadata.

Decision: 16. Vertical extents should be measured either in metres (height above mean sea level, either geometric height or geopotential height), or in hPa where pressure is the vertical coordinate. The guidance needs to recommend text to use for the vertical Coordinate Reference System.

Decision: 17. Limitations on the times for which information is appropriate for use (for example the validity period of a TAF) should be coded in the useLimitation using the period notation (e.g. "PT12H"). This must be included in the guidance.

Decision: 18. TT-ApMD will use the categories in the CDMS (Climate Data Management System) specifications for discovery metadata to structure its own documentation.

Decision: 19. When it is not possible to provide data to complete a component of a metadata record, the entry should be represented in the XML by gco:nilReason. However, for the mandatory elements of a WMO Core Profile metadata record contents should always be provided.

Decision: 20. GAW uses three letter codes as station identifiers. These should be handled in the same way as for surface observations – that is, two sets of keywords, one for the station identifier and the other for the station name. The latitude and longitude for each station are not included (unless the record is for a single station, in which case the location is given by the bounding box).

Decision: 21. The WMO_CategoryCode list needs to be extended and rationalised so that the current meteorology, climate and water disciplines can be described. The granularity of the current entries differs widely; a uniform approach to granularity is needed. This work must be completed before version 2.0 of the Core Profile is defined.

Decision: 22. TT-ApMD recommended to IPET-MDRD that in future versions of the WMO Core Profile the data policy and GTS Priorities should be referenced as code lists rather than as free text in the XML.

Decision: 23. TT-ApMD recommended that all dates (as well as times) have a time zone associated with them. This would normally be UTC (e.g. 2013-06-29Z).

Decision: 24. TT-ApMD guidance will recommend that gml identifiers (gmd:id) are used in metadata records that contain multiple bounding boxes of the same type to label the bounding box that best describes the entire content of the dataset. Metadata records that contain only one bounding box can omit the identifier.

Decision: 25. Within WIS Discovery Metadata, the only requirement for using a gmd:id attribute will be to identify the bounding box that summarizes the total extent of the dataset being described if the metadata record contains multiple bounding boxes. Users may include gmd:id in other contexts, but these have no significance.

Decision: 26. For GTS bulletins the date that the GTS bulletins started or when the content of the bulletin changed, should be recorded in the citation as publication date and revision date respectively.

Decision: 27. TT-ApMD will create standard templates for representative GTS reports and DCPC products. This will require templates for record types that are not in the planned guidance (such as warnings or tsunami reports).

Decision: 28. 4.1.6 Following the definition of the standard templates the GISCs and WIMMS will recreate the metadata records for GTS data.

Decision: 29.   In parallel with recreating the GTS records, all Members will be asked by their Principal GISCs to check and correct the station list, bounding box, contact and data policy elements of their metadata.

Decision: 30.   TT-ApMD decided that for GTS data the Regional Association in which the data originate must be recorded in the metadata as the AdministrativeArea of the originator contact information.

Decision: 31.   TT-ApMD decided that both the originator (the centre creating the data) and the distributor (the centre placing the data onto the GTS) should be recorded through the contact information. The role should be used to achieve this.

Decision: 32.   TT-ApMD decided that keywords from a thesaurus "VolA" should be used to to identify station identifier keywords. Similarly, reporting times should use a thesaurus "ReportingTimes" with the pattern TxxZ for reports issued at specific time.

Decision: 33.   TT-ApMD recommended that the formal procedure for Permanent Representatives to notify changes to the content of GTS traffic and the data policy to be applied to information should be through providing updated metadata records to the Principal GISC. This needs to be approved by CBS and EC.

_____

ANNEX A: LIST OF PARTICIPANTS

Eiji TOYODA (Lead)

Japan Meteorological Agency
Otemachi 1-3-4, Chiyoda-ku
100-8122 TOKYO
Japan
Fax: +81-3 3284 8180
Tel: +81-3 3212 8341 exr 3311
Email: toyoda@npd.kishou.go.jp

Guillaume AUBERT (Member)

EUMETSAT
Postfach 100555
D-64205 DARMSTADT
Germany
Fax: +49 6151 807 426
Tel: +49 6151 807 7196
Email: guillaume.aubert@eumetsat.int

Alexander BESPROZVANNYKH (Member)

Russian Federal Service for Hydrometeorology
and Environmental Monitoring (RIHMI-WDC)
WORLD DATA CENTER
Korolov St. 6
249 035 OBNINSK
Russian Federation
Fax: +(7-095) 2552225
Tel: +(7-08439) 74959
Email: alexander@meteo.ru

Markus HEENE (Member)

Deutscher Wetterdienst
Frankfurter Strasse 135
D-63067 OFFENBACH AM MAIN
Germany
Fax: +49 69 8062 4162
Tel: +49 69 8062 2648
Email: markus.heene@dwd.de

John KOZIMOR (Member)

NOAA/NESDIS/NGOC
325 Broadway
80301 Boulder
Colorado
United States of America
Fax: n/a
Tel: +1 303 497 6729
Email: John.Kozimor@noaa.gov

Chi Kin PAN (Member)

Hong Kong Observatory
134A Nathan Road
KOWLOON
Hong Kong, China
Fax: +852 2311 9448
Tel: +852 2926 8357
Email: ckpan@hko.gov.hk

Kenji TSUNODA (Member)

Japan Meteorological Agency
Otemachi 1-3-4, Chiyoda-ku
100-8122 TOKYO
Japan
Fax: +81-3 3211 8404
Tel: +81-3 3212 8341
Email: tsunoda@met.kishou.go.jp

Peng WANG (Member)

China Meteorological Administration
46 Zhongguancun Nandajie
100081 BEIJING
China
Fax: +86 010 6217 6241
Tel: +86 010 6840 6564
Email: wangpeng@cma.gov.cn

Robert McLeod WILSON (Associate member)

Bureau of Meteorology
G.P.O. Box 1289
3001 MELBOURNE
VIC
Australia
Fax: +61 9669 4128
Tel: +61 9669 4349
Email: b.wilson@bom.gov.au

Licheng ZHAO (Associate member)

China Meteorological Administration
46 Zhongguancun Nandajie
100081 BEIJING
China
Fax: +86 10 6217 4797
Tel: +86 10 6840 7924
Email: lczhao@cma.gov.cn

Shigeharu NISHIKAWA (Interested expert)

Japan Meteorological Agency
Otemachi 1-3-4, Chiyoda-ku
100-8122 TOKYO
Japan
Fax: +81-3 3211 8404
Tel: +81-3 3212 8341
Email: s_nishikawa@met.kishou.go.jp;
gce00463@gmail.com

Steve FOREMAN (Secretariat)

World Meteorological Organization
7 bis, avenue de la Paix
Case postale No. 2300
CH-1211 2 GENEVA
Switzerland
Fax: +41 22 730 8021
Tel: +41 22 730 8171
Email: sforeman@wmo.int

Atsushi SHIMAZAKI (Secretariat)   World Meteorological Organization
7 bis, avenue de la Paix
Case postale No. 2300
CH-1211 2 GENEVA
Switzerland
Fax: +41 22 730 8021
Tel: +41 22 730 8313
Email: ashimazaki@wmo.int

ANNEX B: AGENDA

Item	Topic
1	Organization of the meeting
1.1	Opening of the meeting
1.2	Approval of the agenda
1.3	Working arrangements
2	Validation of WMO Core Profile 1.3
2.1	Key points of Core Profie 1.3
2.2	Automated validation
2.3	Publication of validation tool
3	Publication of guidance
3.1	Categories of guidance needed
3.2	Synoptic information
3.3	Satellite information
3.4	Climate information
3.5	GAW
3.6	Hydrology
3.7	Documents and Standards
3.8	Products (model analyses, forecasts, ensembles and other products)
3.9	Extensions to code lists
4	Training plan for WMO Core Metadata Profile
5	Creating Vol C1 from metadata catalogue
6	Any Other Business
7	Closure of meeting

ANNEX C: TOPICS FOR WHICH TT-APMD WILL PROVIDE GUIDANCE ON PREPARING METADATA.

Guidance area	Sub-division
World Weather Watch	Surface synoptic
World Weather Watch	Upper air
World Weather Watch	Satellite radiances
World Weather Watch	Scatterometer
World Weather Watch	Radar altimeter waves
World Weather Watch	SAR waves
World Weather Watch	Satellite water vapour
World Weather Watch	GPS water vapour
World Weather Watch	Weather radar
World Weather Watch	NWP weather forecast
World Weather Watch	Weather alert
World Weather Watch	Weather warning
Aviation	TAF
Aviation	METAR
Aviation	SigWx
Climatology	CLIMAT
Climatology	Monthly/seasonal/annual/decadal analyses
Climatology	IPCC reports
Climatology	Climate simulations
Oceanography	Sub-surface observations (physical)
Oceanography	Biological observations
Oceanography	Surface observations
Oceanography	chemistry observations
Cryosphere	Sea ice
Cryosphere	Glaciers
Crysophere	Snowfall
Atmospheric chemistry	Radiative measurements
Atmospheric chemistry	Chemical concentrations
Hydrology	Add items
Space Weather	Add items