WORLD METEOROLOGICAL ORGANIZATION

COMMISSION FOR CLIMATOLOGY

REPORT OF THE CCl EXPERT TEAM ON WCP REQUIREMENTS FOR METADATA

Toulouse, France, 11-13 March 2009

WCDMP-No. 73

CONTENTS

EXECUTIVE SUMMARY

The CCl Expert Team 1.1 on Climate Data Management including Metadata held its second meeting on 11-13 March 2009 in Toulouse, France. The primary goals were clearer definition of station and dataset metadata requirements for climatology and support of the WMO Information System (WIS), applicability of the WMO's Core Metadata Profile, and deciding how best to complete the team's metadata tasks. A secondary goal was refining the requirements and the approach for a new survey of Climate Data Management Systems (CDMS) currently installed in member countries.

Sessions were chaired by Team Lead Denis Stuber (France), along with OPAG 1 Chair Raino Heino (Finland), Jeff Arnfield (USA) and Omar Badour (WMO). Additional experts from the team, along with Meteo France and the WMO, provided background and updates on WMO Information System (WIS), ISO 19115 metadata and CDMS-related topics.

Discussions surrounding metadata, in particular station metadata, were complex and intense. Defining a station metadata format that uses existing standards, accommodates comprehensive station details and time series information, yet remains simple and accessible, is a difficult challenge. While a variety of groups, both within the WMO and in external organizations, have worked to address these issues, the work has often been independent and lacked coordination, and no clear solution yet exists.

AGENDA

11 March

1. OPENING OF THE MEETING

1.1 Opening and welcoming remarks
A.2 Adoption of the agenda
1.3 Organization of the meeting

2. KEY NOTE PRESENTATIONS

2.1 CCl Long term vision and future challenges (*Pierre Bessemoulin, President of CCl)*
2.2 Past and present CCl Activities regarding Metadata *(Raino Heino chair CCl-OPAG.I)*
2.3 WMO Climate programs *(Omar Baddour, WMO)*

3. CURRENT STATUS OF WIS IMPLEMENTATION AND WMO CORE METADATA PROFILE

3.1 WIS Project (David THOMAS, WMO)
3.2 THE WMO CORE METADATA PROFILE, purpose and status *(Atsushi SHIMAZAKI, WMO)*
3.3 General Discussions including on prospects for Climate activities

4. REVIEW OF WMO METADATA AND DATA REPRESENTATION

4.1 Review of ET1.1 Task 4.1 on Station metadata definition for climatological practices
4.2 Review of ET1.1 Task 4.2 in Dataset Metadata definition for climatological practices
4.3 Review of the ET1.1 Task 4.3 in Station Metadata exchange suitable to WIS
4.4 Review of the status of work of the Inter-Program Expert Team on Metadata Implementation, its recommendations and implications for CCl
4.5 Review the outcome of the Expert Team on Assessment of Data Representation Systems (ET-ADRS).

12 March

5. EXPERT GROUP FOCUS ON TECHNICAL DEVELOPMENT FOR ADVANCING ET1.1 TASKS ON METADATA

5.1 Preliminary presentation on Siting classification (*Michel Leroy)*
5.2 Implementation of WMO Core Metadata Profile: GTS Metadata generated in the framework of the VGISC project (*Jean Pierre Aubagnac)*
5.3 Group Work on Task 4.1
5.4 Group Work on Task 4.2
5.5 Group work on Task 4.3

13 March

6. REVIEW OF OTHER ET1.1 TASKS

6.1 Review Task 2.1 on List of CDMSs in use in NMHSs
6.2 Review Task 2.1 on Questionnaire about CDMSs in NMHSs
6.3 Review Task 3.2 on updating the requirement for CDMSs
6.4 Review Task 4.2 on CLIREP and its sustainability

7. CONCLUSION AND RECOMMENDATIONS FOR THE REPORT

SESSION DETAILS

1 OPENING OF THE MEETING

1.1 Opening and welcoming remarks

1.2 Adoption of the agenda

Denis Stuber provided some additional context for the agenda items, and indicated that the agenda could be modified during the meeting if necessary to accommodate vital discussions and additional presentations. The group readily approved the agenda.

1.3 Organization of the meeting

The group agreed to convene from 9:00 to 12:30 and from 14:30 to 17:30 in the afternoon, as proposed. Denis Stuber assured the team that the facility would remain available later, should the work at hand necessitate later hours.

2 KEY NOTE PRESENTATIONS

2.1 CCl Long term vision and future challenges (*Pierre Bessemoulin, President of the CCl)*

Pierre Bessemoulin, President of the CCl, welcomed the participants to Meteo France's facility and to the meeting. He proceeded to outline the CCl's history, structure, long-term vision and future challenges. CCl's vision and structure are being evaluated for revision, as are other WMO programs. Some information can be obtained from www.wmo.int/pages/prog/wcp/ccl/index_en.html. It was noted that changes are in the proposal stage, and must undergo further discussion before any approval might be given.

During the general discussion that ensued, several experts noted overlapping efforts in several WMO program areas and project, particularly in regard to metadata. Overlaps between WIGOS, WIS, CBS and CCl were cited. The group also noted the need for, and difficulty of, closer communication and coordination between different WMO teams and projects.

2.2 Past and present CCl Activities regarding Metadata (Raino Heino chair CCl-OPAG.I)

Raino Heino, OPAG 1 chair, also welcomed the attendees. Raino briefly presented some history of the CCl, particularly with regard to metadata, and discussed its current status and challenges.

2.3 WMO Climate programs (Omar Baddour, WMO)

Omar Badour welcomed the participants, and expressed hope that the sessions would provide good progress toward clearer specifications of and solutions to the challenges of metadata for climatology. He then provided an overview of various WMO climate programs, helping the group put the team into a broader context.

3 CURRENT STATUS OF WIS IMPLEMENTATION AND WMO CORE METADATA PROFILE

3.1 WIS Project Overview (David Thomas, WMO)

The WMO's David Thomas made a detailed presentation regarding the status of the WIS project. Initial live implementations of the WIS products and services catalogue are anticipated in late 2009. Products can be registered via an automated metadata harvest, or may be sent to the GISC.

Questions arose regarding VGISCs. In 2006 there were 2 parallel projects, one the France / Germany / UK VGISC, the other the German UNIDART project. What happened to UNIDART? DT said that UNIDART was a fundamentally different design in that it required a central registry. There is also a project involving Germany, Japan and China, in which metadata will be sent to the RA II VGISC and users will then get data via UNIDART.

OB asked if it is good to have multiple DCPCs within one country. DT said that these centres are typically virtual centres within other organisations. For example, in Germany there are several DCPCs that are not within DWD. It is also worth noting that a lot of effort goes into the process of assessing would-be DCPCs. DT also commented that most, but not all, of the NCs are hosted by the relevant NMHS.

3.2 WMO Core Metadata Profile, purpose and status (Atsushi Shimazaki, WMO)

An overview of the current WMO Core Metadata Profile revision was provided by the WMO's Atsushi Shimazaki.

3.3 General discussions including prospects for climate activities

4 REVIEW OF WMO METADATA AND DATA REPRESENTATION

4.1 Review of ET1.1 Task 4.1 on Station metadata definition for climatological practices

Jeff Arnfield provided a brief review of tasks and objectives for these sessions. Discussions often covered aspects of several topics.

The group briefly discussed The Guidelines on Climate Metadata and Homogenization (WCDMP-53), which provides general guidelines on metadata important in climatology, but concluded that a more formal specification is needed for effective exchange of metadata.

Lengthy discussions focused on the applicability of the WMO Core Profile to station metadata, and the implications for WIS. A variety of points were made:

- From the WIS perspective, generating a minimally compliant ISO 19115 record using WMO Core Profile permits an individual station to be discoverable using the WIS search.

- It was strongly expressed during debates that some feel the Core Profile is an inappropriate representation of station metadata. Even if applicable, the Core Profile may not be suitable for all station metadata elements, and may not readily represent the temporal aspects of a station's history (changes in location over time, for example)

Several possible approaches to creating records for station metadata using the WMO Core Profile were offered:

- Jean Pierre Aubagnac's VGISC example illustrated including very basic station metadata (ID, Name, Latitude and Longitude, with a link to a web service for additional information) as an inline feature catalog. This apparently permits the stations to be discovered in WIS. This also links stations to the datasets to which they contribute, addressing a separate requirement that was mentioned during the meeting.

- Create a summary record for each station, with link to NMHS source (web, manual, etc) for additional details

The group agreed that a first step would bet to develop a specific list of station metadata parameters, designating each as mandatory for climate, desirable, optional or unnecessary. This will provide immediate, tangible output and a basis for evaluation of various methods of representation and transfer.

4.2 Review of ET1.1 Task 4.2 in Dataset Metadata definition for climatological practices

Note: agenda items 4.1, 4.2 and 4.3 blended into a series of discussions amongst the experts and guests in attendance. The substance of these discussions is summarized within the other sections of this document.

4.3 Review of the ET1.1 Task 4.3 in Station Metadata exchange suitable to WIS

4.4 Review of the status of work of the Inter-Program Expert Team on Metadata Implementation, its recommendations and implications for CCl

4.5 Review the outcome of the Expert Team on Assessment of Data Representation Systems (ET-ADRS).

Jeff Arnfield provided a brief summary of the April 2008 ET-ADRS meeting held in Washington, DC, USA. The cross-commission group examined a variety of existing and emerging standards, along with some of the systems that had successfully implemented those standards. Discussions covered the gamut between overview level and highly detailed. Both syntactic and semantic interoperability were discussed, with a consensus that interoperability on the semantic, or shared meaning, level was more complex yet critical. Many affirmed the need to maintain operational capabilities, and expressed concerns over resource implications of mandating more complex standards. Key considerations included compatibility with existing systems and processes, and the importance of accommodating bandwidth and other infrastructure limitations.

4.6 WIS implementation implementation by Meteo France International (Patrick Benichou)

Patrick Benichou discussed WIS compliance for CDMSs, illustrated by a discussion of how MFI approaches climate data in a WIS-compliant way in their projects. MFI focuses on Information Systems aspects to help NMHSs modernize infrastructure and develop regional (intra-country) capabilities.

Key WIS concepts include the extension of “GTS-like” functionality to other WMO programs besides CBS, providing a global information system, and allowing data dissemination via “Pull”, “Delayed Push” and “Real-Time Push”. A key component of compliance will be the “WIS layer”, allowing NMHSs to act as GISCs or DCPCs. The “WIS layer” must do various adaptations to ensure WIS compliance.

For WIS purposes, a climate database would be seen as a component of the NMHSs overall information system. A CDMS (such as CliSys) should be able to automatically generate the metadata required by the WIS. Consequently, MeteoFrance International are “WIS-enabling” the systems they provide to the countries with whom they are working.

An in-depth demonstration of accessing the SIMDAT VGISC followed.

5 EXPERT GROUP FOCUS ON TECHNICAL DEVELOPMENT FOR ADVANCING ET1.1 TASKS ON METADATA

5.1 Preliminary presentation on siting classification (*Michel Leroy)*

Michel Leroy provided information on a station classification scheme developed for MeteoFrance in which for each basic parameter there is a rating on a scale of 1 to 5 for the siting and a rating on a scale of A to E for “maintained performance”. This scheme hides some details but is easy to use. With the siting classification, 1 means “follows all WMO recommendations”, 5 means “avoid absolutely for large- or meso-scale applications” (a rain gauge under a tree would merit a 5), with 2 – 4 being intermediate (with clearly-defined criteria for each level for each parameter). He outlined, as an example, what a wind sensor would need to do to meet each of the classifications. MeteoFrance are progressively classifying all their sites. The work is 65% - 95% complete (with completeness varying by region).

5.2 Implementation of WMO Core Metadata Profile: GTS Metadata generated in the framework of the VGISC project (*Jean Pierre Aubagnac)*

5.3 Group Work on Task 4.1

Discussions regarding station metadata content frequently included exchange formats. The variety of expert experiences and orientations, combined with a strong desire to thoroughly consider opinions and possibilities, resulted in spirited debate.

General guidelines for station metadata exist, but they are not specific enough to ensure a consistent implementation and an interoperable result

Much work has been done in various WMO and external teams regarding this topic, but there has been little coordination between those teams. This emphasizes the need for and benefits of a cross-commission, task oriented approach. The proposed CBS expert team on metadata and interoperability may be the best means of achieving this cooperation amongst WMO teams.

Different users desire vastly different levels of station information detail. Any definition must accommodate, but not require, very detailed station information, and must be readily extensible to accommodate information not previously planned for.

One suggestion was to exchange only very basic station metadata, relying upon records maintained separately by the station's host NMHS or other agency for further detail. In this scenario, the shared station metadata would serve a role similar to dataset metadata: provide enough detail for discovery, with a link to a separate repository for the actual data, in this case the detailed station records. Such an approach could simplify the station metadata exchange format, but might result in critical details being unavailable if they were not archived by or were otherwise inaccessible from the host NMHS. Some station operators might discard older station configuration and maintenance details after the operational need for them has passed.

The group was unable to reach consensus as to whether the WMO Core Metadata Profile is useful and appropriate to represent station metadata. Following lengthy discussion, opinions still were split between yes, maybe and definitely not. While it is likely that partial adoption of the WMO Core Profile for station metadata would still be possible in a variety of ways, there remain some station details that might not be adequately described using the profile, and thus other existing standards should be considered in addition to, or instead of, ISO 19115. Such standards include, but are not limited to, the ISO 19110 standard for feature catalogs and such OGC Sensor Web standards as SensorML. The possible use of feature catalogs to link stations and datasets was discussed.

Specification of mandatory elements must be tempered by the fact that detailed station metadata may not have been rigorously collected or maintained over time, and that even now collecting such metadata may strain the resources of developing countries.

In discussing possible exchange and distribution formats for station metadata, the group identified possible enhancements to the current Publication 9 Volume A as a possible way forward. Some limitations to the current publication were also noted.

- Flat files are inherently difficult to extend for additional content, however. Modifications to the Pub 9 Vol A would have an impact on existing applications that rely on that publication's structure. Adding additional codes to the current Observation Remarks column might be a simple, if inelegant, workaround.

- Some data programs, such as GCOS, are not included. One simple way to denote which programs a station contributes to might a series of flag columns; while an XML schema would permit additional entries of a given type, such an approach would require altering a flat file's structures.

- Stations that do not contribute to one of the usual programs may still deliver observations that are eventually archived, but there is no way to denote that such observations are archived and available.

- Observing frequencies other than synoptic, hourly and half-hourly are not currently represented. Continuous observations, such as strip charts, are not represented, and special or irregular observations not handled.

- Is it necessary, or even desirable, to maintain historical configuration information? Currently the only means of tracking a station's entire history would be to review successive versions of Volume A. While changes to a station over time are a fundamental aspect of station metadata, including multiple historical configurations would add complexity to the publication, and could increase its size significantly.

- Currently closed stations do not appear in the publication. The group believed they should, although there would need to be some indication that the station no longer was in operation. One means of addressing this might be inclusion of an overall begin and end date for each station

- If Publication 9 Volume A were expanded to include additional observing systems, such as those making RADAR observations, it is unknown whether all key metadata elements are present. Currently there is no mechanism to accommodate new types of information except as a remark.

- If a significant number of new stations were added, the current WMO Index Number scheme may not permit each to be assigned a unique identifier.

5.4 Group Work on Task 4.2

Apart from a brief overview during the session, there was inadequate time to cover this in detail

The ISO 19115 is a robust, flexible standard, and many organizations have expressed intent to use it as their standard for dataset metadata. The latest revision of the WMO Core Metadata Profile included many fewer extensions than the original version, affirming the ability of the standard to accommodate a variety of collection metadata requirements. From casual inspection, it seems that the primary focus should be evaluation of the WMO Core Profile code tables to ensure they support adequate categorization of metadata for climatological uses.

During the WIS presentation, it was pointed out that the metadata record's abstract, which is usually entered as free text, often played a central role in locating information of interest. Guidance on creating an adequate, useful abstract might be useful.

While many metadata elements may be of interest to climatologists, whether climatology has unique metadata requirements remains an open question.

5.5 Group work on Task 4.3

Ability and limitations of the WIS to use separate feature catalog to discover stations and provide linkage to detailed station metadata needs to be more thoroughly explored

There was general agreement that a well-defined, system-neutral format that can be extended, such as an XML schema, would be the most useful method of exchanging station information. Such a schema should be standards-based, although additional code tables might need to be defined as they have been for other WMO purposes. Any format for station metadata transfer must be readily extensible to accommodate information not previously planned for.

6 REVIEW OF OTHER ET1.1 TASKS

6.1 Review Task 2.1 on List of CDMSs in use in NMHSs

6.2 Review Task 2.1 on Questionnaire about CDMSs in NMHSs

6.3 Review Task 3.2 on updating the requirement for CDMSs

6.4 Review Task 4.2 on CLIREP and its sustainability

7 CONCLUSION AND RECOMMENDATIONS FOR THE REPORT

Tasks and Work Plan

- There are many user communities, and uses, for climate data. Metadata must permit each community to locate, evaluate and access information of interest. Documenting some basic use cases for both station level and dataset and product level metadata will help us refine and validate metadata requirements.

- It is unclear whether climatology has unique metadata requirements, distinct from those of other users. Defining the metadata parameters and vocabularies that are critical for climatology will permit us to objectively evaluate the adequacy of metadata structures and keywords.

- Develop a specific list of station metadata parameters, using Publication 9 Volume A and other resources as a starting point. Each attribute should be designated as either mandatory, desirable, optional or unnecessary for climatological use. This will provide immediate, tangible output and a basis for evaluation of various methods of representation and transfer.

- Because there was no consensus on the suitability and adequacy of the WMO Core Metadata Profile for station metadata, the group should determine what other standard(s) might be applicable as a complementary or alternative approach. The list of metadata parameters defined above should help to objectively validate these standards.

- Regardless of the standard used to represent station metadata, clearly defined keywords and code values will be necessary. Existing WMO code lists should be used where applicable. Where the WMO code lists are incomplete, additions should be proposed. Where no appropriate WMO code list exists, other standard vocabularies should be identified, evaluated and selected if suitable.

- Defining and implementing summary quantification and scoring techniques for station siting and operating procedures based on clearly defined criteria, such as used in Michel Leroy's quality factor technique, would add value to metadata while simplifying it and reduce its volume for end users; however, such a task could be complex, and might not be possible in many cases where adequate historical metadata were not available in digital form.

- The WMO Core Metadata Profile, as applied to datasets, will also use WMO code lists and keywords. The current code lists should be reviewed to ensure their adequacy for climate purposes

- Proposed team activities:

    - Propose list of specific "mandatory" and "recommended" station metadata elements, along with definitions, based on WCDMP-53 and taking into account work done by ET-AWS (Geneva 2006)

    - Review and provide feedback on metadata elements

    - Define code tables necessary to exchange basic station metadata

- Evaluate code tables to ensure their adequacy for climate purposes
- Identify any metadata elements or keyword requirements that are unique to climatology
- Provide comments, suggestions on Alexander's proposed revision to Pub 9 Vol A
- Recommend that the proposed revisions to Pub 9 Vol A be elevated in importance and considered for implementation
- Provide working team with proposed GML schema for Station Metadata Feature Catalog along with reference to supporting standards.
- Evaluate applicability of WMO Core Metadata Profile to station Metadata by developing a basic, compliant record using minimal station metadata as defined under task 4.1
- Develop guidance for generating effective metadata abstracts to better facilitate free text searches.

Recommendations

Active coordination between various WMO and external groups focusing on station and dataset metadata is necessary to minimize duplication of efforts and increase quality of output. The team recommends that a multi-commission, task-focused team be formed to address metadata issues, perhaps including representatives from related external groups where appropriate.

The WMO's Publication 9 Volume A is currently the WMO's most widely distributed station metadata format. While extensive work has been accomplished to define content and format enhancements to the WMO's Publication 9 Volume A, formal action has not been taken to refine and implement the proposed changes. The team recommends that the proposed changes be refined and clarified as needed, then submitted for review and adoption by the WMO.

PARTICIPANTS

Mr Jeffrey Arnfield

National Climatic Data Center/NOAA
151, Patton Avenue
Asheville, NC 28801-5001
USA
Tel: (1-828) 271 4444
Fax: (1-828) 271 4022
Email: jeff.arnfield@noaa.gov

Mr Alexander Besprozvannyh

Russian Research Institute for Hydrometeorological Information – World Data Center (RIHMI-WDC)
6, Korolyov Str.
Obninsk, Kaluga Region 249035
Russian Federation
Tel: (7-095) 2552467
Fax: (7-095) 2552225
Email: alexander@meteo.ru

Mr Pierre Bessemoulin

Météo France
Direction de la Climatologie
42, avenue Gustave-Coriolis
31057 TOULOUSE-Cedex
France
Tel (33-5) 61-07-83-00
Fax (33-5) 61-07-83-09
pierre.bessemoulin@meteo.fr

Dr Raino Heino

Finnish Meteorological Institute (FMI)
P.O. Box 503
00101 Helsinki
Finland
Tel: (358-9) 19294120
Fax: (358-9) 19293503
Email: raino.heino@fmi.fi

Mr John Shortridge

Bureau of Meteorology
Supervisor, Computing Support,
National Climate Centre
Australia
Tel: 61-3-96694282
Email: J.Shortridge@bom.gov.au

Mr Denis Stuber

Météo-France
42 avenue Gustave Coriolis
31057 Toulouse Cedex
France
Tel: (33-5) 61078337
Fax: (33-5) 61078309
Email: denis.stuber@meteo.fr

Invited Experts fom Météo-France

Mr Philippe Dandin

Météo-France
42 avenue Gustave Coriolis
31057 Toulouse Cedex
France
Tel: (33-5) 61078300
Fax: (33-5) 61078309
Email: philippe.dandin@meteo.fr

Mr Michel Leroy

Météo-France
7 rue Teisserenc De Bort, B.P. 202
78195 Trappes
France
Tel: (33-5) 130166405
Fax: (33-5) 130166010
Email: michel.leroy@meteo.fr

Mr Jean-Pierre Aubagnac

Météo-France
42 avenue Gustave Coriolis
31057 Toulouse Cedex
France
Tel: (33-5) 61078245
Fax: (33-5) 61078109
Email: jean-pierre.aubagnac@meteo.fr

WMO Secretariat

Mr Omar Baddour

Chief, Data Management Applications
WMO Information System Branch
Observing and Information Systems Department
Tel: 41 22 730 8268
Fax: 41 22 730 8042
Email: obaddour@wmo.int

Mr David Thomas

Project Manager, WMO Information System Branch
Observing and Information Systems Department
Tel: 41 22 730 8241
Fax: 41 22 730 8021
Mobile:+41 79 425 59 60
Email: dthomas@wmo.int

Mr Atsushi Shimazaki

Scientific Officer
Data Representation, Metadata and Monitoring (DRMM)
WMO Information System Branch
Observing and Information Systems Department
Tel: 41 22 730 8313
Fax: 41 22 730 8042
Email: ashimazaki@wmo.int

REFERENCES AND PRESENTATIONS

Agenda Item	Document
2.2	http://www.wmo.int/pages/prog/wcp/ccl/index_en.html
2.3	http://www.wmo.int/pages/prog/wcp/wcdmp/wcdmp_series/documents/WCDMP-53.pdf http://www.wmo.int/pages/prog/wcp/wcdmp/documents/WCDMPNo62.pdf
4.1	Technical Note No. 35: Site Classification, Michel Leroy, Meteo France, November 1999 http://www.wmo.int/pages/prog/wcp/wcdmp/wcdmp_series/documents/WCDMP-53.pdf http://www.wmo.int/pages/prog/www/OSY/Reports/ET-AWS4_Geneva2006.pdf http://www.wmo.int/pages/prog/www/ois/ois-home.htm http://www.wmo.int/pages/prog/www/WMOCodes.html
4.2	http://www.wmo.int/pages/prog/www/WDM/Metadata/documents.html
4.3	http://www.wmo.int/pages/prog/www/OSY/WMO-9A/proposal-rapporteur.pdf http://www.opengeospatial.org/standards/sensorml http://www.ogcnetwork.net/SensorML_Examples
4.4	http://www.wmo.int/pages/prog/www/CBS/CBS-Ext06_tecoWIS/1-3-2_UK-Foreman_WMO-CoreMetadata_20060909.doc http://www.wmo.int/pages/prog/www/wdm/ipet-mi-iii/ipet-mi-iii_report.doc
4.5	http://www.wmo.int/pages/prog/www/WDM/ET-ADRS-1/ET-ADRS-1_draft_report_v7.doc http://www.wmo.int/pages/prog/www/CBS/Meetings/MG_9/documents/Doc5-2_2_OPAG-ISS_ETS2008.doc
5.1	http://www.wmo.int/pages/prog/www/IMOP/reports/2008/ET-ST&MT-2_Final%20Report.pdf