|  |  |
| --- | --- |
| **World Meteorological Organization**  **COMMISSION FOR BASIC SYSTEMS**  **Inter-Programme Expert Team on Data representation Development** 18-20 February 2019, Zagreb (Croatia) | **IPET-DD 1** |
| 21.3.2019  Submitted by: Secretariat |

# Final report OF THE First Meeting of the Inter-Programme Expert team on Data representation development

# 18-20 February 2019



**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).

© World Meteorological Organization, 2018

The right of publication in print, electronic and any other form and in any language is reserved by WMO. Short extracts from WMO publications may be reproduced without authorization, provided that the complete source is clearly indicated. Editorial correspondence and requests to publish, reproduce or translate this publication in part or in whole should be addressed to:

Chairperson, Publications Board

World Meteorological Organization (WMO)

7 bis, avenue de la Paix Tel.: +41 (0)22 730 84 03

P.O. Box No. 2300 Fax: +41 (0)22 730 80 40

CH-1211 Geneva 2, Switzerland E-mail: [Publications@wmo.int](mailto:Publications@wmo.int)

NOTE:  
The designations employed in WMO publications and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of WMO concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries.

The mention of specific companies or products does not imply that they are endorsed or recommended by WMO in preference to others of a similar nature which are not mentioned or advertised.

This document (or report) is not an official publication of WMO and has not been subjected to its standard editorial procedures. The views expressed herein do not necessarily reflect those of WMO or its Members.

# 1 OPENING OF THE MEETING

1. Thorsten Büsselberg, Chair of IPET-DD, opened the session expressing gratitude to DHMZ for hosting the meeting and recalling the importance of the first IPET-DD meeting in consideration of the plans for WIS 2.0 and in view of widening the user base and improving the services provided by WIS Centres.
2. Dr. Kreso Pandzic, Advisor of DHMZ Director, welcomed the participants on behalf of the Director Dr. Branka Ivančan-Picek, highlighted the importance of the IPET-DD team and wished a fruitful meeting.
3. Djiana Klaric, co-Chair of IPET-DD, welcomed the participants.
4. After a round of introduction of the participants the Chair recalled the principal objectives of the meeting:

* Improve search and provide metadata guidance in view of the implementation plan of WIS 2.0
* Discuss the new aviation codes to become operational in 2020
* WIGOS metadata representation and other XML based formats
* netCDF-CF data requirements with particular attention to new radar data format

1. Thorsten Büsselberg presented a video on WIS 2.0 developed by DWD. After the video he highlighted the importance of improving WIS portals functionalities starting from the users’ needs, moving the focus towards services as well as data, and making sure that metadata are fit for purpose and provide what is expected by the user.
2. Enrico Fucile presented the progress on WIS 2.0 implementation plan by introducing the 11 principles that are at the core of the renewed plan, and have to be the focus of the team in all the future work regarding metadata. The principles are reported in Annex 3 and were discussed actively by the team. The comments were in general supportive of the new WIS 2.0 principles, and the presentation was able to set the ground for the future work of the Team in a view of the future changes. There were only two comments that need to be considered in future review of the implementation plan:
   * Daniel Michelson noted that the term “reduction”, in the principle 5, is not clear and can be interpreted as a suggestion to provide a degraded version of the original data. It will be necessary to be clearer in the implementation phase to be sure that the reduction will result in an operation that is required by users and fit for the purpose of providing useful information for a big dataset. Direct interaction of the user with the big data to run a specific process closer to the dataset can be provided by some Centres, but this is a tailored service that is not expected to be cost free and will be provided only by a limited number of Centres.
   * Kate Roberts noted that in principle 7 the cache provided by the GISCs is still 24 hours, which arguably is insufficient for some data such as climatological reports that are going to be present in the cache for a time that is not sufficient for the user to make use of them. It is suggested that a longer cache persistence is provided for some data.
3. The Team decided to keep the 11 principles [see Annex 3] as a reference for future work on metadata to make sure that changes are aligned with future WIS 2.0 plans.

# 2 TT-MDS

1. Guillaume Aubert, Chair of TT-MDS, made a presentation on the status of the work of TT-MDS. TT-MDS main objective is to explore and improve data discoverability and challenges experienced by WIS users seeking information on data and services. The team is requested to conduct a user-centric analysis of the effectiveness of GISC catalogue search interfaces in the context of core WMO audiences and their needs. Working in close consultation with related WIS/GISC teams, and based on the above analysis and in the context of known other technologies, the team shall discuss and recommend on proposed solutions, providing

- Guidance for GISC/DCPC providers on search interface design

- Guidance for metadata authors including assessment of ‘metadata granularity’

- Guidance for metadata tool developers

- Guidance on improved discovery and searching, including making WIS metadata records discoverable via industry search engines

- Guidance on Quality assessment monitoring, and tools / other solutions to expedite the transition to improved metadata content (in conjunction with TT-MDG).

1. TT-MDS has conducted an analysis on the different GISC catalogue discovery services based on the experience of the different members and the following challenges for users have been highlighted:
   * The WIS Portals are often focusing on the GTS community using the GTS technical jargon as well as specific discovery services for the GTS (TTAAii) hence omitting to serve a larger community that is only looking to find meteorological products in its broader terms.
   * The different portals provide very different search experiences that do not always follow industry standards. This is confusing for users that are looking for consistency between the different portals, or who expect a behaviour for such discovery services derived from their day to day interactivity with the web industry search engines, or from their previous use of well-designed catalogue interfaces (incorporating such functionality as 'faceted searching', simple/advanced search options, fielded searching, popup hints, saved searches, etc).
   * The different portals are not fully focusing on helping users to discover products and instead tend to be very technical, therefore creating a steep learning curve for neophyte users. The product information presentation is also very often following the ISO Metadata internal structure. ISO 19115 is an interesting format for capturing the product information; however if the metadata presented in a catalogue follows the 19115 structure, sections need to be clearly indented and labelled (or otherwise arranged), to make it user-friendly and to avoid ambiguity.
   * A very large part of the Metadata content has been automatically generated from VOL C, with a very limited set of information for qualifying accurately the products, and differentiating them sufficiently to allow a search engine to perform efficiently, or to allow a catalogue user to easily identify a relevant product.
   * Due to the nature of the GTS bulletins (and the absence of metadata content to signify that the product is a subset of a "collection"), the WIS portal catalogues treat metadata for collections of products, and metadata for instances of products (GTS bulleting metadata), in the same way. Metadata for product collections describe families of products such as for instance 3 hourly forecasts without drilling down to the last parameters of the products for selecting a specific data file with a meteorological parameter and a given forecast time. They provide general discovery information on the products regrouped under the collection. Product instances are individual instances of a product family with all the necessary parameters to access the product data file itself (ie temperature of the 3 hourly forecast from the 20190110 3 am). This mix of discovery metadata and “access” metadata has the side effect of flooding the discovery results with lots of product instances results and therefore creating a degraded search experience for the user.
2. The full paper, with the guidance on how to address the issues highlighted under 8, is in Annex 4.
3. The IPET-DD discussed the presentation given by Guillaume Aubert, the issue of resourcing and tasks already shared across TT-MDS and TT-MDG, and expressed the following comments, decisions and actions.
4. IPET-DD considered the necessity for TT-MDG and TT-MDS to work in closer interaction given the existing overlap of some tasks of the two teams (such as improving metadata quality, addressing metadata granularity), the dependence of the action of TT-MDS on the work plan of TT-MDG, and the fact that almost all the members of one team are in the other team.
5. IPET-DD decided that TT-MDG and TT-MDS will work together with a single work plan, and a focus on delivering on common objectives
6. IPET-DD noted that a good discovery service is very dependent on the metadata content it uses, regardless of whether the discovery mechanism is a search engine or catalogue GUI. Poor quality content, the absence of good user-help, and the absence of a good GUI design which guides users towards more efficient search actions, can result in too many matches in a search result, without a clear way to refine the search. As well as metadata content, GUI design features such as faceted searching, the option to "mark/select" records and put them in a shopping-cart, to save searches, etc are all options which can enhance a user's search effectiveness.   
   TT-MDS and TT-MDG were both previously tasked with identifying ways to improve metadata quality, though TT-MDS's focus has been on the quality which affects search effectiveness.
7. In order to improve the quality content, TT-MDS and TT-MDG shall define a set of Key Performance indicators to evaluate and monitor the WIS metadata content. Based on the list of created key performance , a first evaluation shall be conducted by the TT-MDS and TT-MDG.  
   Results shall be published and provided to ET-WIS. TT-MDS and TT-MDG shall liaise with ET-WIS about having the quality content monitoring integrated in the GISC Watch.  
   In addition, the defined KPIs shall be regularly reviewed and, as needed, modified or new ones developed.
8. TT-MDS and TT-MDG to develop Metadata quality KPIs, and recommend a process for improving the quality of the metadata content
9. IPET-DD noted that a set of discovery components and navigation elements based on the type of users shall be described and recommended by TT-MDS and TT-MDG to improve the discovery services. A set of mockups shall be provided to detail the type of navigation and discovery components that should be provided by the WIS catalogues.
10. TT-MDS and TT-MDG to define and recommend best practices for implementing web discovery services (via catalogue GUIs and search engines).
11. IPET-DD recommends that TT-MDS and TT-MDG produce guidance material which explains the issues created by the different level of metadata granularity in the WIS Catalogues, and which recommends an approach where product collection metadata records and product instance records shall be separated, or made distinct, so that discovery services can / do support the user to discover meteorological products based on collection information, and then sub-select and access the appropriate instances using the adequate parameters (right time step, right geographical, right elevation). TT-MDS and TT-MDG shall also demonstrate that this two-step approach can be used to link data discovery and data access services, as well as to facilitate "data access services" discovery.
12. TT-MDS and TT-MDG to recommend a separation between product collection and product instances in the catalogues.
13. TT-MDS and TT-MDG shall explore and recommend on preferred machine to machine interfaces, to be implemented by the different WIS catalogues for increasing their interoperability. A set of functionalities and goals for the APIs shall be defined, and a machine to machine standard or protocol recommended to be implemented depending on particular needs.
14. TT-MDS and TT-MDG shall develop a mapping between the information contained in the WCMP and the dataset fragment of [schema.org](http://schema.org/). This shall allow WIS records and WIS catalogues to be better indexed by industry search engines.

# 3 TT-MDG

1. Kate Roberts, Chair of TT-MDG, presented the activity of the team, with reference to the 'MDG team\_overview\_DiscussionPoints' paper (Annex 9). The paper highlighted the various drivers which have influenced the changing direction and focus of TT-MDG over time. The paper provides the sections of IPET-DD's TOR, and IPET-DD's Workplan which are relevant to TT-MDG, and includes the relevant sections of "High priority items for IPET-DD", as well as the TT-MDG TOR.
2. The paper recommends improved communications processes for advertising new guidance, new support, new examples; as well as for getting feedback from metadata creators, and monitoring the effectiveness of these.
3. The paper also highlighted some findings related to improved metadata quality (for which improved guidance and support is needed). Some of the findings overlap those of TT-MDS, some are related to recent Guidance material that has not been implemented, and some are new recommendations. Most guidance / recommendations aim to improve a user's capacity to determine the suitability of that product. Findings include :

* the need for better and more consistent use of keywords,
* use of thesauri,
* the need for clear, consistently applied and correct details of the temporal and spatial coverage of the product;
* the need for a clear declaration about the retention policy (availability of the data) for GTS products;
* the need to include well-maintained urls to supporting material (such as to product /format specifications, data dictionaries);
* improved declarations about how to access data, and about data licencing;
* fuller descriptions of the data (and, where relevant, the data lineage and data's quality);
* clearer descriptions of the data format and how to decode it.

Some possible additional solutions include metadata re-use, such as use of xlinks, or cross-references to an umbrella/collection-level record (which more fully describes common details of format, encoding, data parameters, release schedules, and so on).

1. The paper also provided a brief overview of previous activities to address metadata quality monitoring, and recommends improved interaction with GISC managers and their monitoring activities.
2. The paper also noted that a workplan was developed by TT-MDG in June 2017, and after commencing its review of previous guidance and support, TT-MDG, in September 2017, was encouraged to move its focus towards the "IPET-DD issues list" posted on the wiki. After identifying the subset of issues that were relevant to TT-MDG, the team struggled a little with resolving the issues, due to not knowing, nor knowing how to find out, the initiating background / context , however, as a number of TT-MDG activities were downstream of recommendations from TT-MDS or WIS2 discussions, or were meant to be collaborative efforts with TT-MDS, a higher prioritisation was given to TT-MDS activities, and to WIS2 discussions.
3. Kate Roberts reminded the team that there is a list of tasks that needs to be reviewed. IPET-DD decided to review the task list during the parallel session in a way that the TTs will come back with a refined list of tasks. (Tasks were reviewed and prioritised during the breakout session, on Day 3)
4. Two additional TT-MDG papers were also tabled and discussed by the meeting participants: a document reviewing the content of a metadata record for CSAU01; and a document providing an overview of evolving metadata standards (including ISO19115-3 (2015)), and related issues and considerations needed.

A6 TT-MDS and TT-MDG shall highlight the benefits for migrating WMCP toward ISO 19115-3 and recommend a migration path from ISO 19115 to ISO 19115-3. Examples of changes and additional information to be added shall be provided as necessary.

1. IPET-DD decided to use GitHub to manage the future tasks of TT-MDG and TT-MDS, and recommended that the two teams should work, as a matter of high priority area, on moving the tasks to the new GitHub area that the Secretariat will make available to them.

# 4 TT-AvXML

1. Mark Hedley, Chair of TT-AvXML, introduced IWXXM and explained the change management and the use of GitHub in the maintenance of the project. He explained that the Team is adopting semantic versioning with a Major.Minor.Patch version model. Where
   * Major version is increased when the change is not backward compatible, which means that the change is affecting user software and forcing the user to make some changes.
   * Minor version is increased when a new functionality is added in a backward compatible manner. This means that the new functionality is not forcing a change in the user software except that for the use of the new functionality itself.
   * Patch version is increased when a bug fix is made in a backward compatible manner.
2. The codes registry codes.wmo.int can be considered as functional to the XML schema of IWXXM in the sense that provides permanent identifiers and URLs for the code lists used in IWXXM. This is a good design choice as changes in the codes registry can be used to implement new functionality without any release of the schema and without affecting users’ code. The problem of how long the identifiers are going to persist in codes.wmo.int has been discussed and the Secretariat is committed to maintain the code lists in the code registry for decades to come being the list of terms an important operational part of WMO codes and an important resource for the users to access authoritative terms.
3. Mark Hedley to prepare some text on the codes.wmo.int main page to explain the function of the registry as provider of permanent identifier and the importance of maintaining the identifiers to persist for very long time. Deprecation and requested changes should be also documented to give the users a view of what they can expect in terms of changes.
4. Mark Hedley continued presenting the official public and institutional repository which can be accessed at <http://schemas.wmo.int> . This is the official repository where everyone can access the schemas, as compared to the repository in GitHub available at github.com/wmo-im/iwxxm . schema.wmo.int provides access to the latest official version, to the release candidate versions (denoted with RCx where x is version number) and past official versions. The RC versions are going to be deleted immediately after the official version is released, while older official versions can persist for long time to avoid operational problems to users who are still using them.
5. Mark Hedley to link schemas.wmo.int to the GitHub repository.
6. The role of GitHub and schemas.wmo.int were discussed and the team agreed on the following decisions.
7. GitHub is the development repository for IWXXM and other XML based data representations. Access to the GitHub repository is granted to all the members of a team and comments/issues can be made by anyone who is interested in the development of the schema. Examples and documentation should be part of the repository, and releases made available on it for as long as possible, allowing comparison back in time of old changes and issues resolution. GitHub should be the memory of the development process and the live tool used for developing new features and versions.
8. schemas.wmo.int is the public, institutional, operational repository. The service on schemas.wmo.int is not maintained at critical level in the sense that operational services should not rely on the resilience of the service. schemas.wmo.int can be down for some time and this should not affect operations around the world. On schemas.wmo.int only release candidate and final approved versions will be published and the release candidates are published only for the purpose of testing by users or in the formal approval process.
9. Mark Hedley explained the new structure of the GitHub repository that now allows a proper continuous integration model to be implemented. In the past, different directories for each version were introduced into the repository; and the mechanism of tagging the releases was not used at all. This has been changed and at the moment a repository with a single version of the schemas, examples and other required files is provided. From now on the releases will be tagged and a travis integration system has been implemented. The system is currently performing simple regression testing. An increase of the tests in future will provide more confidence that a new version is not breaking users’ software.
10. IPET-DD agrees that the continuous integration model , and the use of automatic tests with travis is an important tool to ensure the quality of the releases, and safeguard the user from unexpected behaviours in the operational systems.
11. TT-MDR to provide examples in GitHub and to implement tests for continuous integration.
12. During the presentation it has been noted that the examples are using http://icao.int/iwxxm/xx that is not maintained by ICAO and it is a redirect to schema.wmo.int. The team agreed that this should be a domain maintained by ICAO
13. TT-AvXLM to present a paper to MIE/5 to ask ICAO to maintain the domain icao.int for the purpose of IWXXM.
14. B.L. Choy presented the process of updating IWXXM schema. The changes are made in the logical model that is developed using Enterprise Architect (EA). In the early days of the development several years ago another software FullMoon was used to convert the model from XML Metadata Interchange (XMI) to a schema that the users can use for validation of their xml data. Recently Enterprise Architect provided the capacity to produce schemas from xmi and therefore FullMoon is not used any more. However, some scripts and a special setting in EA are used to produce IWXXM schemas due to a particular technical decision that doesn’t appear to be relevant anymore and is making the process of producing xml schemas unnecessarily complicated.
15. TT-AvXML to avoid using special settings for the production of xml schemas from the logical model (xmi files).
16. BL Choy presented IWXXM 3.0 and proposed the modification to the Manual on Codes Vol. I.3 aimed to the introduction of the new version of IWXXM in the technical regulations. During the discussion it was noted that the proposed changes are at a higher level of detail compared with the other XML schemas documented in the Manual on Codes and that a shorter version of the IWXXM 3.0 manual would be more suitable.
17. TT-AvXML to propose, by end of March, the text for the Manual on Codes for the introduction of IWXXM 3.0.
18. IWXXM 3.0 RC3 will be made publicly available on <http://schemas.wmo.int> and submitted for approval between sessions. Changes proposed during the approval process will be included in the 3.0 release, to be made public after formal endorsement by President of WMO.
19. Mark Hedley presented a paper on the sustainability of WMO xml standards. The complexity of the schemas, and the very frequent change requirements from ICAO, make it challenging for a small team to deliver on time, and with sufficient testing to ensure the required quality. The process is based on modelling (with EA) the new features or changes, and producing new schemas that are going to be tested and reviewed by the users. Although the process is made complex by the use of EA, the complexity of the schemas does require a design tool to be able to model the data representation, before implementing the changes into an xml framework. A better streamline of the conversion from the model to the schemas is needed, but the use of EA was considered of value by IPET-DD. Another important element is that there are only two active members who are able to perform data modelling and this is not a workforce that can produce continuous changes to the schemas. The same problem of lack of human resources is present in all the team dealings with xml formats which change with some regularity (including TT-WMD). This is considered as a critical element for the future of the xml formats.
20. IPET-DD agrees that there are challenges in the sustainability of the xml formats and considers three main important factors to be critical: streamlining the process of conversion from model to schema, implementing sufficient testing to deliver quality representations, and having sufficient workforce to perform the required maintenance and development tasks
21. TT-AvXML to provide a streamlined process from logical model to schema.
22. All teams maintaining xml formats to provide tests, and a continuous development setup.
23. Secretariat to resolve the lack of experts for the purpose of xml data modelling.

# 5 WMDR and other XML formats

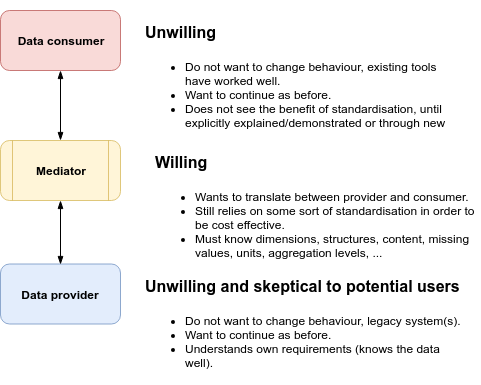
1. Enrico Fucile made a presentation on the status of WIGOS metadata representation 1.0 on behalf of Joerg Klausen (Chair of TT-WMD) who apologised for not being able to participate in the meeting.
2. As reported by Enrico Fucile, the WIGOS metadata representation version 1.0 (wmdr 1.0) is currently under CBS consultation on WMO wiswiki (<https://wiswiki.wmo.int/tiki-index.php?page=CBS-consult-2018> ). The consultation process will be closed by end of February and the proposed changes to the Manual on Codes Vol. I.3 will be proposed to Congress 18 for final approval. The proposed changes to the Manual on Codes include a URL to the following document [WMDR\_ModelAndSchemaSpecification.pdf](http://schemas.wmo.int/wmdr/1.0/WMDR_ModelAndSchemaSpecification.pdf) describing wmdr 1.0 schema and data model. This document is of great value for the users seeking to understand the metadata representation, and IPET-DD is requested to consider its publication as an appendix to the Manual on Codes Vol. I.3.
3. IPET-DD decided to support the request of publication of [WMDR\_ModelAndSchemaSpecification.pdf](http://schemas.wmo.int/wmdr/1.0/WMDR_ModelAndSchemaSpecification.pdf) as appendix to the Manual on Codes Vol. I.3.
4. IPET-DD with the support of Secretariat to review [WMDR\_ModelAndSchemaSpecification.pdf](http://schemas.wmo.int/wmdr/1.0/WMDR_ModelAndSchemaSpecification.pdf) and make the required editorial changes to publish it in the Manual on Codes Vol. I.3.
5. The WIGOS metadata representation is supported by several code tables already published in the final report of EC-69 (https://library.wmo.int/doc\_num.php?explnum\_id=3645 ) .   
   The code tables also support the wmdr 1.0 schema, and they need to be made available in the codes registry ( <http://codes.wmo.int> ) to be accessed as URLs and provide the supporting information to the xml data.   
   The tables are maintained for development purposes in the following GitHub repository <https://github.com/wmo-im/wmds> and a consolidated set as recommended by TT-WMD is ready to be proposed for fast track.
6. IPET-DD agreed on the submission of wmdr tables for fast track.
7. TT-WMD to provide examples for wmdr 1.0 in the GitHub repository.
8. Silvano Pecora, Vice-President of the Commission for Hydrology, made a presentation GroundWaterML2. He explained that OGC adopted
   * WaterML2 Part1 - Timeseries in February 2014,
   * WaterML2 Part2 - Ratings, Gaugings and Sections in February 2016,
   * the conceptual model WaterML2 Part3 - Surface Hydrology Features in January 2018,
   * WaterML2 Part 4 - GroundWaterML2 in March 2017
9. and WMO Executive Council 69th session meeting adopted WaterML2 Part 1 & 2 in May 2017. He stressed the importance for WMO activities to adopt WaterML2 Part 4 - GroundWaterML2 and explained function and use of the standard. GroundWaterML2 has been tested in the WMO Hydrological Observing System (WHOS) for sharing groundwater information and he reminded that GroundWaterML2 will complete the standard WaterML2 already adopted in the Manual on Codes Vol. I.3. Silvano Pecora requested IPET-DD to review and submit for approval in the Manual on Codes the proposed change including GroundWaterML2.
10. IPET-DD agreed to support the publication of GroundWaterML2 on the Manual on Codes Vol. I.3
11. IPET-DD with the support of Secretariat to review the proposed change to the Manual on Codes Vol. I.3 regarding GroundWaterML2 reported in Annex 5.

# 6 WMO Codes registry and GitHub

1. Nina Horat from MeteoSwiss made a presentation via webex on the use of GitHub by the TT-WMDS team for the development of the tables of the WIGOS metadata representation (wmdr). She explained that the tables approved at EC-69 have been loaded on a GitHub repository and that work is in progress to make the required changes in several feature branches connected with issues opened by the Task Team regarding modifications and additions to the existing tables. The git workflow adopted by the team requires opening an issue every time there is a request for a change and the team members to exchange opinions on the change and define the characteristics of the change in the issue itself. When the issue is mature enough a branch with the name or number of the issue will be created and the changes on the tables implemented in the branch. When the issue is closed the branch can be merged to a development branch. Many issues can be processed in parallel and when it is the time for a fast track or another official approval procedure, the changes that are already present in the development branch and have been merged from closed feature branches are going to be merged in the master branch which is then used to provide the documents for approval by Members. This process is simple and allows to incorporate in the fast track only the changes that have already been approved by the members of the team, while changes that are not ready yet can be deferred to another fast track.
2. Chair of IPET-DD welcomed the use of GitHub and highlighted the importance of such a tool to enable interaction between teams and between members of a team with a tool that allows to work around the world and around the clock as requested in the experts and task teams work.
3. IPET-DD decided to use GitHub for all the task teams, and suggested the same practice of other IPETs and TTs.
4. Secretariat to open a space for TT-MDS and TT-MDG with a repository of the WIS Metadata Core Profile (WMCP) for future work and discussion.
5. TT-MDS and TT-MDG to maintain a list of tasks on GitHub for the future workplan.
6. Enrico Fucile explained that the GitHub repository for WIGOS metadata representation contains also python scripts for the conversion of the csv file of the wmdr tables to Terse RDF triple language (Turtle) with the purpose to publish the tables on the WMO codes registry codes.wmo.int. The scripts are able at the moment to provide files used to upload full table onto the codes registry.
7. Mark Hedley made a presentation regarding the Codes registry. He explained that codes.wmo.int is hosted by the Met Office in the UK on behalf of the WMO. The codes registry service provides machine and human access to individual elements published within WMO publications – entities , and the content is published using the W3C standard Resource Description Framework (RDF). This enables rich machine to machine interface capabilities and supports multi-lingual representation. The codes registry provides:
   * Direct access and identification of every published entity,
   * Flexible information about entities,
   * Entity history and status.
8. The status of the registry is that the WMO owns and maintains the content and is responsible for the validity and suitability of the content. The publication is partial, compared to WMO official publications and it is the authoritative source of information only for some of the data representations, mainly the once using xml.
9. IPET-DD discussed the need to work in coordination with IPET-CM and agreed that GitHub and codes.wmo.int are suitable tools to foster collaborative work.
10. IPET-DD recommends a consistent use of GitHub and codes.wmo.int by IPET-CM and its TTs and proposes to reinforce the collaboration between the two expert teams.

# 7 netCDF-CF data formats

1. The aim of the session on netCDF and CF convention based formats is to establish the requirements for CF based data representation in WMO Programmes and to decide if there is a need to establish some work in WMO to support those requirements.
2. Daniel Michelson, Chair of IPET-OWR, presented the new format for radar and lidar data in radial coordinates (CfRadial2) based on netCDF-CF standards. The format is based on CF conventions and CfRadial2 has been proposed to the CF conventions governance to be accepted as part of the CF conventions. However the process of acceptance is not progressing at the moment and it is not clear if CfRadial2 will be accepted by the CF governance committee. IPET-OWR has written a Best Practices Guide with a part on Data Representation and Exchange which is reported in Annex 6.
3. Mark Hedley expressed the position of UK Met Office as being not sure that the new CfRadial2 format has been sufficiently discussed within the Eumetnet community and accepted as the official format to be use in global radar data exchange on the basis that there is already a HDF based format ODIM that is commonly used for radar data exchange between Eumetnet members and it is a standard used in UK national data exchange.
4. IPET-DD requests the Chair of IPET-OWR to make sure that the new CfRadial2 format is accepted in Eumetnet context and by UK in particular before the final drafting of the CfRadial2 part for the Manual on Codes.
5. IPET-DD recognises the importance of having a new data format for operational radar and agrees that CfRadial2 has to be included in the Manual on Codes as a standard for the exchange of operational weather radar data.
6. IPET-DD recognises also the complexity of introducing a data representation based on netCDF and CF convention, being this the first format with this characteristics, and proposed IPET-OWR to use the format for experimental exchange with the aim to consolidate the data representation and exchange practices.
7. IPET-DD to work in collaboration with IPET-OWR on a draft of CfRadial2 for the Manual on Codes to be proposed at the next Technical Commission Session.
8. Øystein Godøy from Norwegian Meteorological Institute joined the meeting via webex to represent the needs of the WMO Global Cryosphere Watch Programme regarding netCDF and the CF convention. The entire paper explaining the need of CF conventions and netCDF format for GCW is included in Annex 7. The main reason for the use of netCDF is in a very inhomogeneous community based on researchers and with a big difficulty in changing the tools that are commonly used for the data production. The proposed solution is a data portal that is going to act as a mediator between the provider and the consumer as shown in the following picture.



The data portal is able to convert data to BUFR to exchange data on GTS, but a standardisation of the metadata both for discovery and use purposes is needed to enable interoperability. For this purpose CF convention is used, but it is considered too wide for the use in operational systems and there is a requirement to produce narrower CF profiles to be maintained by WMO for the benefit of GCW Programme and similar needs can be envisaged for other WMO Programmes. However the definition and maintenance of CF profiles by WMO can easily deviate from the CF conventions if they are developed independently and without coordination with CF governing committee. Therefore a collaboration with CF community is seen as beneficial to avoid making conflicting standards that are difficult to reconcile.

1. IPET-DD recognises the needs for CF profiles to be maintained by WMO, but also recognises the need to open a dialogue and collaboration with the CF community to avoid to build conflicting data representations.
2. Guillaume Aubert made a presentation, on behalf of Daniel Lee from Eumetsat, on the position of EUMETSAT regarding CF conventions (full document in Annex 8). EUMETSAT recognises the wide use of netCDF and has plans to use it as native format for many of the future products and has recognised that CF conventions are the most suitable conventions to be used for satellite data. As early as 2008 the WMO Expert Team on Assessment of Data Representation Systems (ET-ADRS) noted that the CF Conventions are likely the most suitable standard for use in WMO systems when encoding data products in netCDF 4 [1]. The Expert Team also noted that the Conventions have certain shortcomings. The CF Conventions are already in their second decade of existence and were created when netCDF-3 was the most recent version of the netCDF format. Experience with using the Conventions have shown that:

* Many attractive features of netCDF-4 are not covered by the Conventions that would be useful for encoding data products, and
* The Conventions have evolved largely around forecast and in-situ observation data, so that improvements could be made to fit more closely with the needs of the satellite community.

In contrast to the WMO standards, the CF community does not have dedicated resources for evolving the standard and coordinating usage. This has led EUMETSAT to increase its involvement with the CF governance process. This paper describes the current state of this involvement and proposes ways to improve the use of the CF Conventions so that they better match the needs of the WMO community.

The WMO community would profit from:

* Making use of the existing CF Conventions when encoding data products in netCDF;
* Contributing to the development of the CF Conventions using existing processes;
* Specifying a subset of the CF Conventions that would be considered appropriate for operational use in WMO;
* Considering the use of one or more additional metadata attributes that would be similar but orthogonal to the CF standard\_name attribute that would allow the unambiguous description of variables, as well as mappings between WMO and CF metadata standards, while removing any dependency between WMO and CF governance processes.

It is proposed to explore these options in a workshop involving representatives of both WMO and CF. The workshop should pursue the goal of producing clear terms of reference for future work in contributing to the evolution of the CF Conventions and governing the use of these conventions within WMO.

1. Sebastien Villaume explained the work that ECMWF has done to establish a CF based ECMWF profile for various types of NWP data having a GRIB representation. He explained that the base format for ECMWF is still GRIB and they don’t have plans to move to netCDF, but there is a requirement by the users to have netCDF for the wide use of that format in the research communities. He explained in details why the CF convention is not sufficient for operational purposes and it is too weak in the representation of time and space variables. ECMWF has setup a set of rules, that can be collected in a specific NWP profile, aimed to reduce the variability in the possible representations of a field and to define uniquely concepts that are used in the context of GRIB and for NWP purposes. The need of a NWP profile and a better way to map netCDF to GRIB is clear for the presentation. ECMWF is not providing any specific request or way forward to improve the situation.
2. IPET-DD agrees that :

* there is a need for netCDF in the WMO communities and Programmes because the format is widely used in the research community, and to reduce the research to operation time, it is a good strategy to use in operation and in research formats that are as close as possible.
* netCDF is not a good standard for interoperability purposes and that CF conventions are partially covering the lack of specific structure of netCDF.
* CF conventions are too wide to enforce full interoperability and to facilitate the operational activities, therefore in many Programmes there is the need for specific profiles that could be maintained by WMO
* defining WMO netCDF profiles can produce data representations that are in conflict with CF conventions and this is not in the interest of the communities and WMO, therefore it is important to coordinate WMO activities of development of specific netCDF profiles based on CF conventions that WMO works in close collaboration with CF community to avoid conflicting data representaions.

1. IPET-DD proposes to WMO Secretariat to organise a workshop on “CF conventions in WMO Programmes” to bring together experts from WMO data representation and programmes activities and the CF communities with the aim to establish possible collaboration for the development of CF profiles for the use in WMO Programmes.
2. WMO Secretariat to organise a workshop on “CF conventions in WMO Programmes”

# 8 Task Teams parallel sessions and workplan

1. The Task Teams worked in parallel sessions. TT-MDS and TT-MDG worked as a single team to produce a common work plan. TT-AvXML discussed the use of Enterprise Architect and refined its work plan toward IWXMM 3.0. No members of the TT-WMD were present.
2. Task Teams to update their work plan on their GitHub wiki or repositories.

# ACTIONS AND DECISION SUMMARY

## Decisions

**D1** The Team decided to keep the 11 principles [see Annex 3] as a reference for future work on metadata to make sure that changes are aligned with future WIS 2.0 plans.

**D2** IPET-DD decided that TT-MDG and TT-MDS will work together with a single work plan, and a focus on delivering on common objectives

**D3** In order to improve the quality content, TT-MDS and TT-MDG shall define a set of Key Performance indicators to evaluate and monitor the WIS metadata content. Based on the list of created key performance , a first evaluation shall be conducted by the TT-MDS and TT-MDG.  
Results shall be published and provided to ET-WIS. TT-MDS and TT-MDG shall liaise with ET-WIS about having the quality content monitoring integrated in the GISC Watch.  
In addition, the defined KPIs shall be regularly reviewed and, as needed, modified or new ones developed.

**D4** IPET-DD noted that a set of discovery components and navigation elements based on the type of users shall be described and recommended by TT-MDS and TT-MDG to improve the discovery services. A set of mockups shall be provided to detail the type of navigation and discovery components that should be provided by the WIS catalogues.

**D5** IPET-DD recommends that TT-MDS and TT-MDG produce guidance material which explains the issues created by the different level of metadata granularity in the WIS Catalogues, and which recommends an approach where product collection metadata records and product instance records shall be separated, or made distinct, so that discovery services can / do support the user to discover meteorological products based on collection information, and then sub-select and access the appropriate instances using the adequate parameters (right time step, right geographical, right elevation). TT-MDS and TT-MDG shall also demonstrate that this two-step approach can be used to link data discovery and data access services, as well as to facilitate "data access services" discovery.

**D6** IPET-DD decided to use GitHub to manage the future tasks of TT-MDG and TT-MDS, and recommended that the two teams should work, as a matter of high priority area, on moving the tasks to the new GitHub area that the Secretariat will make available to them.

**D7** GitHub is the development repository for IWXXM and other XML based data representations. Access to the GitHub repository is granted to all the members of a team and comments/issues can be made by anyone who is interested in the development of the schema. Examples and documentation should be part of the repository, and releases made available on it for as long as possible, allowing comparison back in time of old changes and issues resolution. GitHub should be the memory of the development process and the live tool used for developing new features and versions.

**D8** schemas.wmo.int is the public, institutional, operational repository. The service on schemas.wmo.int is not maintained at critical level in the sense that operational services should not rely on the resilience of the service. schemas.wmo.int can be down for some time and this should not affect operations around the world. On schemas.wmo.int only release candidate and final approved versions will be published and the release candidates are published only for the purpose of testing by users or in the formal approval process.

**D9** IPET-DD agrees that the continuous integration model , and the use of automatic tests with travis is an important tool to ensure the quality of the releases, and safeguard the user from unexpected behaviours in the operational systems.

**D10** IWXXM 3.0 RC3 will be made publicly available on http://schemas.wmo.int and submitted for approval between sessions. Changes proposed during the approval process will be included in the 3.0 release, to be made public after formal endorsement by President of WMO.

**D11** IPET-DD agrees that there are challenges in the sustainability of the xml formats and considers three main important factors to be critical: streamlining the process of conversion from model to schema, implementing sufficient testing to deliver quality representations, and having sufficient workforce to perform the required maintenance and development tasks

**D12** IPET-DD decided to support the request of publication of WMDR\_ModelAndSchemaSpecification.pdf as appendix to the Manual on Codes Vol. I.3.

**D13** IPET-DD agreed on the submission of wmdr tables for fast track.

**D14** IPET-DD agreed to support the publication of GroundWaterML2 on the Manual on Codes Vol. I.3

**D15** IPET-DD decided to use GitHub for all the task teams, and suggested the same practice of other IPETs and TTs.

**D16** IPET-DD recommends a consistent use of GitHub and codes.wmo.int by IPET-CM and its TTs and proposes to reinforce the collaboration between the two expert teams.

**D17** IPET-DD recognises the importance of having a new data format for operational radar and agrees that CfRadial2 has to be included in the Manual on Codes as a standard for the exchange of operational weather radar data.

**D18** IPET-DD recognises also the complexity of introducing a data representation based on netCDF and CF convention, being this the first format with this characteristics, and proposed IPET-OWR to use the format for experimental exchange with the aim to consolidate the data representation and exchange practices.

**D19** IPET-DD recognises the needs for CF profiles to be maintained by WMO, but also recognises the need to open a dialogue and collaboration with the CF community to avoid to build conflicting data representations.

**D20** IPET-DD agrees that :

* there is a need for netCDF in the WMO communities and Programmes because the format is widely used in the research community and to reduce the research to operation time it is a good strategy to use in operation and in research formats that are as close as possible.
* netCDF is not a good standard for interoperability purposes and that CF conventions are partially covering the lack of specific structure of netCDF.
* CF conventions are too wide to enforce full interoperability and to facilitate the operational activities, therefore in many Programmes there is the need for specific profiles that could be maintained by WMO
* defining WMO netCDF profiles can produce data representations that are in conflict with CF conventions and this is not in the interest of the communities and WMO, therefore it is important to coordinate WMO activities of development of specific netCDF profiles based on CF conventions that WMO works in close collaboration with CF community to avoid conflicting data representaions.

**D21** IPET-DD proposes to WMO Secretariat to organise a workshop on “CF conventions in WMO Programmes” to bring together experts from WMO data representation and programmes activities and the CF communities with the aim to establish possible collaboration for the development of CF profiles for the use in WMO Programmes.

## Actions

**A1** TT-MDS and TT-MDG to develop Metadata quality KPIs, and recommend a process for improving the quality of the metadata content

**A2** TT-MDS and TT-MDG to define and recommend best practices for implementing web discovery services (via catalogue GUIs and search engines).

**A3** TT-MDS and TT-MDG to recommend a separation between product collection and product instances in the catalogues.

**A4** TT-MDS and TT-MDG shall explore and recommend on preferred machine to machine interfaces, to be implemented by the different WIS catalogues for increasing their interoperability. A set of functionalities and goals for the APIs shall be defined, and a machine to machine standard or protocol recommended to be implemented depending on particular needs.

**A5** TT-MDS and TT-MDG shall develop a mapping between the information contained in the WCMP and the dataset fragment of schema.org. This shall allow WIS records and WIS catalogues to be better indexed by industry search engines.

**A6** Mark Hedley to prepare some text on the codes.wmo.int main page to explain the function of the registry as provider of permanent identifier and the importance of maintaining the identifiers to persist for very long time. Deprecation and requested changes should be also documented to give the users a view of what they can expect in terms of changes.

**A7** Mark Hedley to link schemas.wmo.int to the GitHub repository.

**A8** TT-MDR to provide examples in GitHub and to implement tests for continuous integration.

**A9** TT-AvXLM to present a paper to MIE/5 to ask ICAO to maintain the domain icao.int for the purpose of IWXXM.

**A10** TT-AvXML to avoid using special settings for the production of xml schemas from the logical model (xmi files).

**A11** TT-AvXML to propose, by end of March, the text for the Manual on Codes for the introduction of IWXXM 3.0.

**A12** TT-AvXML to provide a streamlined process from logical model to schema.

**A13** All teams maintaining xml formats to provide tests, and a continuous development setup.

**A14** Secretariat to resolve the lack of experts for the purpose of xml data modelling.

**A15** IPET-DD with the support of Secretariat to review WMDR\_ModelAndSchemaSpecification.pdf and make the required editorial changes to publish it in the Manual on Codes Vol. I.3.

**A16** TT-WMD to provide examples for wmdr 1.0 in the GitHub repository.

**A17** IPET-DD with the support of Secretariat to review the proposed change to the Manual on Codes Vol. I.3 regarding GroundWaterML2 reported in Annex 5.

**A18** Secretariat to open a space for TT-MDS and TT-MDG with a repository of the WIS Metadata Core Profile (WMCP) for future work and discussion.

**A19** TT-MDS and TT-MDG to maintain a list of tasks on GitHub for the future work plan.

**A20** IPET-DD requests the Chair of IPET-OWR to make sure that the new CfRadial2 format is accepted in Eumetnet context and by UK in particular before the final drafting of the CfRadial2 part for the Manual on Codes.

**A21** IPET-DD to work in collaboration with IPET-OWR on a draft of CfRadial2 for the Manual on Codes to be proposed at the next Technical Commission Session.

**A22** WMO Secretariat to organise a workshop on “CF conventions in WMO Programmes”

**A23** Task Teams to update their work plan on their GitHub wiki or repositories.

# ANNEX 1. AGENDA AND TIMETABLE

|  |  |  |  |
| --- | --- | --- | --- |
|  | **MONDAY, 18 February 2019** | |  |
|  | Agenda item |  |  |
|  | 1 | Opening |  |
| 9:00 |  | Welcome | Chair  Enrico Fucile  DHMZ representative |
|  |  | Personal introduction | Chair + all |
|  |  | Meeting arrangements | Chair |
|  |  | Agenda approval | Chair |
|  |  | Terms of reference and objectives of the meeting | Chair |
|  |  | WIS 2.0 Principles | Enrico Fucile |
| 10:30 |  | Break |  |
|  | 2 | TT-MDS |  |
| 10:45 |  | TT-MDS Objectives | Guillaume Aubert |
|  |  | TT-MDS Current Tasks | Guillaume Aubert |
|  |  | Discussion | All |
|  |  | RoadMap for completing the on-going work and future tasks | All |
| 12:30 |  | Lunch break |  |
|  | 3 | TT-MDG |  |
| 13:30 |  | TBD | Kate Roberts |
|  |  | Discussion | All |
| 15:00 |  | Break |  |
| 15:15 |  | Further discussion on items 2 and 3 and conclusions | All |
| 17:00 |  | Close of day 1 |  |

|  |  |  |  |
| --- | --- | --- | --- |
|  | **TUESDAY, 19 February 2019** | |  |
|  | Agenda item |  |  |
|  | 4 | TT-AvXML |  |
| 9:00 |  | Introduction of IWXXM | Mark Hedley |
| 9:15 |  | IWXXM 3.0 | BL Choy |
| 10:00 |  | Future development of IWXXM:  - Use of GML in aviation  - Introduction of generic weather objects in METCE | Mark Hedley  BL Choy |
| 10:30 |  | Break |  |
| 10:45 |  | Development and publication of IWXXM   * Use of Sparx Enterprise Architect in modelling and generation of XML schemas * Use of SVN and Git for development, testing and project management + GitHub best practices * Change management and control procedures | BL Choy  Mark Hedley |
| 12:00 |  | Discussion session for model development and project management | All |
| 12:30 |  | Lunch break |  |
| 13:30 |  | Challenges:   * Sustainability in particular on knowledge sharing and upkeeping of team experties * Resourcing | Mark Hedley |
|  | 5 | WMDR and other XML formats |  |
| 14:00 |  | WMDR 1.0 | Joerg Klausen |
| 14:30 |  | GroundWaterML 2 | Silvano Pecora (moved to day 1 14:00) |
| 15:00 |  | Break |  |
|  | 6 | WMO Codes Registry |  |
| 15:15 |  | WMDR tables in gitHub and codes registry | Nina Horat (moved to 9:00) |
| 16:00 |  | Discussion on using GitHub for development and maintenance of the content of the code registry |  |
| 17:00 |  | Close of day 2 |  |

|  |  |  |  |
| --- | --- | --- | --- |
|  | **WEDNESDAY, 20 February 2019** | |  |
|  | Agenda item |  |  |
|  | 7 | netCDF-CF data formats |  |
| 9 :00 |  | Adoption of CfRadial 2.0 as a WMO standard representation of radial weather radar data | Daniel Michelson |
|  |  | GCW | Øystein Godøy |
| 10:30 |  | Break |  |
| 10:45 |  | Satellite data | Daniel Lee |
| 11:45 |  | ECMWF CF conventions | Sebastien Villaume |
| 12:30 |  | Lunch break |  |
| 13:30 |  | Discussion on CF conventions and possible collaboration models |  |
| 14:30 | 8 | Parallel sessions (TTs workplan) |  |
| 16:30 | 9 | Plenary |  |
| 17:00 |  | Close of day 3 |  |

# ANNEX 2. IPET-DD-1 Participants

i

Participants present at the venue

* Thorsten BÜSSELBERG (Chair IPET-DD), Germany
* Dijana KLARIC (co-Chair IPET-DD), Croatia
* Kreso Pandzic, Croatia
* Kate ROBERTS (Chair TT-MDG), Australia
* Daniel MICHELSON (Chair IPET-OWR), Canada
* B. L. CHOY, Hong Kong, China
* Mark Hedley (Chair TT-AvXML), UK
* Guillaume Aubert (Chair TT-MDS), Eumetsat
* Hanane KAMIL, Morocco
* Hermann ASENSIO (co-Chair TT-IM), Germany
* Tero KOIVUNEN, Finland
* Enrico FUCILE, WMO Secretariat

Participants connected via teleconference

* Silvano PECORA (vice-President CHy), Italy
* Manuel FUENTES, ECMWF
* Sebastien VILLAUME, ECMWF
* Steve OLSON, USA
* Mark OBERFIELS, USA
* Øystein Godøy, Norway
* Jitsuko HASEGAWA (Chair IPET-CM), Japan

# ANNEX 3. WIS 2.0 principles

|  |  |  |
| --- | --- | --- |
|  | WIS 2.0 Principles |  |
| 1 | **WIS 2.0: adopts Web technologies and leverages industry best practices and open standards** .  **BENEFIT**: Use of widely adopted practices and open standards will enable a large population of users to conveniently interact with WIS 2.0 to discover, access, and use authoritative weather, water and climate data. |  |
| 2 | **WIS 2.0: uses Uniform Resource Locators (URL) to identify resources (i.e. Web pages, data, metadata, APIs) .**  **BENEFIT**: URLs uniquely identify a resource and describe the primary mechanism for retrieving or interacting with it (i.e. the network 'location' and the communications protocol to be used). |  |
| 3 | **WIS 2.0: prioritizes use of public telecommunications networks (i.e. Internet) when publishing digital resources .**  **BENEFIT**: Publishing digital resources on the Internet enables the meteorological community to retrieve or interact with those resources - it is unlikely that most of the community would be permitted to join managed networks such as Area Meteorological Data Communications Networks (AMDCN) employed by NMHS for data exchange with guaranteed service levels.  **BENEFIT**: Internet connections are significantly cheaper than the same bandwidth delivered through a managed networks. |  |
| 4 | **WIS 2.0: requires provision of Web service(s) to access or interact with digital resources (e.g. data, information, products) published using WIS.**  **BENEFIT**: Web services support 'machine-actionability' (i.e. the capacity of software systems to access, interoperate, and reuse data with little or no human intervention) because humans increasingly rely on computational support to deal with data as a result of increase in volume, complexity and velocity (i.e. creation speed) of data.  **BENEFIT**: NMHSs develop their capacity to build and operate Web services, allowing them to extract more value from their data holdings through delivery of higher value services to their users. |  |
| 5 | **WIS 2.0: encourages NCs and DCPCs to provide 'data reduction' services via WIS that process 'big data' to create results or products that are small enough to be conveniently downloaded and used by those with minimal technical infrastructure.**  **BENEFIT**: Using ‘data reduction’ Web services to process high volume, complex data remotely, Members’ agencies and institutions can deliver high-value, high-quality services to their governments and citizens helping them more effectively meet their national mandates without the need to invest in and operate their own data management infrastructure . |  |
| 6 | **WIS 2.0: will add open standard messaging protocols that use the publish-subscribe message pattern to the list of data exchange mechanisms approved for use within WIS and GTS.**  **BENEFIT**: Low effort for data providers to distribute data in real-time to large numbers of consumers. |  |
| 7 | **WIS 2.0: will require all services that provide real-time distribution of messages (containing data or notifications about data availability) to cache/store the messages for a minimum of 24-hours, and allow users to request cached messages for download.**  **BENEFIT**: Software systems that consume real-time data or notifications can recover from failure by requesting delivery of messages that were missed while the system was offline. |  |
| 8 | **WIS 2.0: will adopt direct data-exchange between provider and consumer.**  **BENEFIT**: Faster transmission of real-time data by avoiding latency introduced by message switches at intermediate GTS nodes. |  |
| 9 | **WIS 2.0: will phase out use of routeing tables and bulletin headers.**  **BENEFIT**: Simplified message switching operations for all Members because routeing table maintenance no longer required.  **BENEFIT**: Faster setup of new data-sharing arrangements as there is no need to wait for intermediate nodes to update their routeing table configuration. |  |
| 10 | **WIS 2.0: will provide a Catalogue containing metadata that describes both data and the service(s) provided to access that data.**  **BENEFIT**: Users will be able to easily find the data in WIS that interests them, locate the most convenient Web service with which to access that data, and determine how to best use that Web service to meet their needs. |  |
| 11 | **WIS 2.0: encourages data providers to publish metadata describing their data and Web services in a way that can be indexed by commercial search engines.**  **BENEFIT**: Indexing by commercial search engines will help users discover data and associated services using their preferred search engine rather than having to find and use a WIS portal. |  |

# ANNEX 4. TT-MDS Guidance

The following document summarizes the activities undergone by the TT MDS. It additionally outlines the work still be achieve and provides a list of future tasks to be completed.

## TT-MDS Term of Reference

Below is the TT-MDS TOR as defined initially by the Task Team to explore and improve “data discoverability” challenges experienced by WIS users seeking information on data and services:

The team shall conduct a user-centric analysis of the effectiveness of GISC catalog search interfaces, in the context of core WMO audiences and their needs.

Working in close consultation with related WIS/GISC teams, and based on the above analysis and in the context of known other technologies, the team shall discuss and recommend on proposed solutions, providing

- Guidance for GISC/DCPC providers on search interface design

- Guidance for metadata authors including assessment of ‘metadata granularity’

- Guidance for metadata tool developers

- Guidance on making WIS metadata records discoverable via industry search engines

- Guidance on Quality assessment monitoring, and tools / other solutions to expedite the transition to improved metadata content.

## TT-MDS initial WIS portal discovery services analysis

The TT-MDS TOR contains a set of actions regarding the monitoring and evolution of the WIS discovery services with the aim of improving the discovery services, the usability of such services. In addition, the team has to analyze and define recommendation for improving the quality of the metadata content to address some of the discovery and usability limitations.

The team has conducted such analysis on the different GISC catalogue discovery services based on the experience of the different members and the following challenges for users have been highlighted:

1. The WIS Portals are often focusing on the GTS community using the GTS technical jargon as well as specific discovery services for the GTS (TTAAii) hence omitting to serve a larger community that is only looking to find meteorological products in its broader terms.
2. The different portals provide very different search experiences that do not always follow industry standards. This is confusing for users that are looking for consistency between the different portals and expect a behaviour for such discovery services derived from their day to day interactivity with the web industry search engines.
3. The different portals are not fully focusing on helping users to discover products and instead tend to be very technical, therefore creating a steep learning curve for neophyte users. The product information presentation is also very often following the ISO Metadata internal structure. ISO 19115 is an interesting format for capturing the product information but its internal structure is not to be followed to display the product information itself as it can be misleading for the user.
4. A very large part of the Metadata content has been automatically generated from VOL C with a very limited set of information for qualifying accurately the products and differentiating them sufficiently for allowing a search engine to perform efficiently.
5. Due to the nature of the GTS bulletins, the WIS portal catalogues are treating the same way collections of products and instances of products. Product collections describe families of products such as for instance 3 hourly forecasts without drilling down to the last parameters of the products for selecting a specific data file with a meteorological parameter and a given forecast time. They provide general discovery information on the products regrouped under the collection. Product instances are individual instances of a product family with all the necessary parameters to access the product data file itself (ie temperature of the 3 hourly forecast from the 20190110 3 am). This mix of discovery metadata and “access” metadata has the side effect of flooding the discovery results with lots of product instances results and therefore creating a degraded search experience for the user.

From this assessment, TT-MDS choose to address the different limitations as follow:

*Points 1. 2. and 3.* are related to the WIS user definition, the discovery services presentation, look and feel and more importantly usability and how to incorporate discovery services web industry standards in the WIS services. For this reason the team decided to create a guidance documentation on the definition of the Web discovery services by going through a Web design process: capture the different user profiles to highlight the diversity of today’s WIS users and their motivations drawing on the diverse experiences of the Task team members, then derive a minimum set of discovery services as well as usability best practices to be recommended to the WIS Development team with mock-ups to highlight the design.

In addition, industry standards or best practices should be implemented to broaden the WIS discovery services. The Task Team should also study and recommend the most relevant web discovery industry best practices for the WIS.

*Point 4.* is touching on the Metadata content quality which needs to be tackled by the WIS in order to have efficient and accurate discovery services. TT-MDS can recommend implementing a set of KPIs for monitoring the quality of metadata content in order to improve the discovery services with quick wins on a very short period.

*Point 5.* can also be addressed by recommending a different way to provide access to the GTS datasets while not penalizing the complete set of discovery services.

## TT-MDS guidance documentation

Based on the assessment summarized in *2 TT-MDS initial WIS portal discovery services a* the overall following guidance documentation structure has been defined by TT-MDS:

1. WIS discovery challenges and potential improvements (based on the summary above)
2. Definition of the potential users of the WIS Catalogue in 2018.
3. Recommendation of a set of common functionalities and usability patterns across GISCs portals.
4. Rational for improving the WIS content and recommendation for addressing it.
5. Implementation of discovery industry standard and provision of a modern machine to machine interface.
6. Differentiate between general product information (collection) and instance product information.

Below is the status of each individual points covered and a summary of the conclusions to be provided in the report.

### WIS Discovery limitation and potential improvements

Most of the limitations and potential improvements have already been addressed in TT-MDS initial analysis. The final guidance documentation should develop each different points and define a set of recommendation for addressing the limitations.

### Definition of the potential users of the WIS Catalogue in 2018

The TT-MDS has performed a capture of the different user communities and user personas that have to be covered by the WIS discovery. The details analysis is available here: <http://bit.ly/WIS-userprofiles> and can be summarized in the following users’ type:

* ***Neophyte users*** with limited or null knowledge in meteorological datasets. This regroups a set of users that would like to discover the different type of datasets that are available in the meteorological/climatology domain or users that want to access meteorological data but do not know how they are described in term of organisation, format. This type of users needs a way to browse through a list of available products as they do not know which keywords to use to match any datasets. Based on that first discovery by browsing, they can create more targeted search requests using filtering and keywords.
* **Meteorological users**. This type of users know already the domain and type of information they are looking for and wants a fast way to target specific datasets using keywords with the ability to discover related (sibling) datasets or compare datasets and this can be provided by the use of filtering search results. Such users also require a fast clear path to data and they would like to quickly access and download the data products to start their work.
* **GTS specialized users**. The access to the GTS using its own jargon and codification has to be maintained without having to make it too prevalent to avoid impacting the discovery of the rest of the available datasets. The TT-MDS can recommend a Web interface for the GTS that is functional but stays reserved for specialists.

More details are available in <http://bit.ly/WIS-userprofiles> and the guidance documentation on discovery created for the TT-MDS has to describe the different categories of users and provide set recommendations for building discovery services for each of the different categories of users.

* 1. **Recommendation of a set of common functionalities and usability patterns across GISCs portals**

Based on the user analysis above, a set of discovery services and broad navigation for the different WIS portals can be recommended. Three types of discovery are recommended: Browsing for new users, full text search with filtering (faceting) for narrowing down the search for more advanced users and a targeted independent search for GTS specialists using the GTS conventions (TTAAii classification).

In addition simple navigation organised similarly to the main Web discovery services should be recommended to follow what is expected by all internet users in 2018 when accessing discovery services.

A fast path for accessing data should be as well recommended and identified when necessary.

Technicalities should be hidden to focus solely on the product: product information, product access need to be the focus for the services. Users do not want to see the ISO metadata structure or artefacts from the GTS technical organisation but instead want to have enough product information for understanding the purpose of the data product and then have a fast and clear path for accessing it.

This can be provided by the TT-MDS as a set of recommendations with a few illustrative Web Design mock-ups to allow the implementers better understanding of what should be developed.

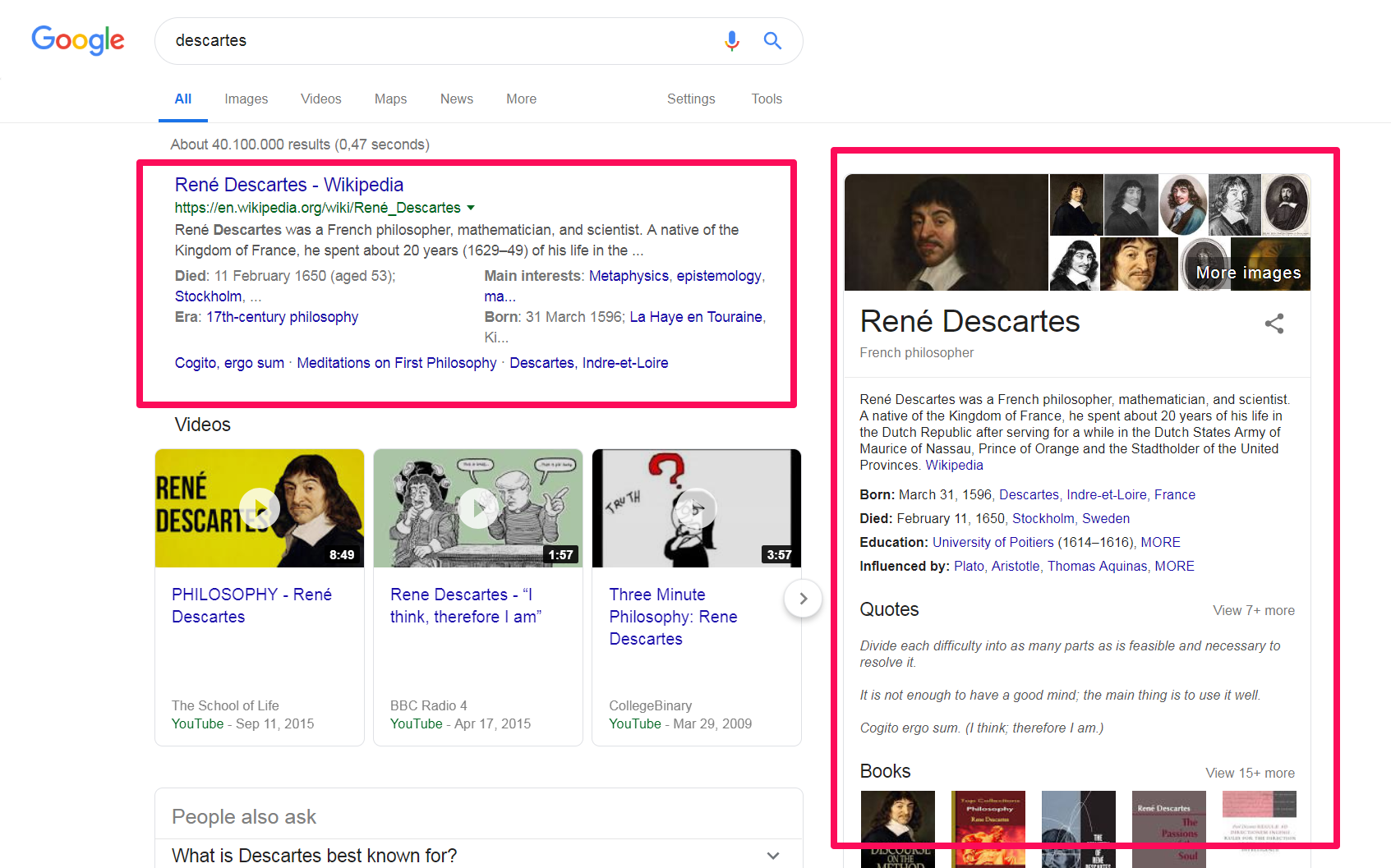
If possible a certain coherency and a set of commonalities in the look and feel in addition to the navigation commonalities should be implemented by all WIS discovery portals.

* 1. **Rational for improving the WIS content and recommendation to address it**

Any discovery service is as good as the content it is created from. Each different indexed products needs to contain enough information to differentiate them and have keywords accurately matching results. A big portion to the WIS metadata records has been generated using a standard template from VOL C to describe the GTS content and just changing the title and TTAAii information for each individual bulletin. It is in that case very difficult for any search engine to not return almost all those metadata records for each keyword matching the initial template. In general for avoiding such issues, the metadata quality should be monitored to understand how to improve it and allow modern search engines returning accurate results. The TT-MDS should develop and recommend the implementation of KPIs to monitor the quality of the metadata content. Additional work can be performed to define those KPIs and a Metadata quality service prototype could be implemented if time permits to demonstrate how the useful and profitable it would be for the WIS. This work has not yet started but it could be initiated in the next phase.

* 1. **Implementation of discovery industry standar**d **and provision of a modern machine to machine interface**

In today's internet, search engines like Google.com are the main initial source of information for the large majority of users. The WIS Portals should have the ability to export the content of their catalogue in the discovery industry services and have their content easily indexed by them otherwise they stay isolated from the rest of the internet and are cutting themselves from a very large portion of users who are searching for meteorological products. Modern industry search engines are now organized as a relationship graph for linking pages and information content and make use of machine learning technics based on that initial knowledge graph to direct the user towards similar or related search results. Using structured information in the pages, the search engines can also extract additional useful information to immediately provide it in the search results. The structured information can also be used to better categorize and index websites content.

Example of the utilisation of annotated structured information in industrial search engines

A standardization effort maintains and document in http://schema.org website has been created a few years ago within the search engine industry where they are documenting the structured information that needs to be provided for different types of information (movies, video games, howTo, …). A few years ago an initiative has been started to provide structured information for datasets and have them better indexed. The information can be found here: <https://schema.org/Dataset>. Google is also actively working on integrating and indexing very efficiently datasets that are documented using the datasets annotations (<https://developers.google.com/search/docs/data-types/dataset>). The dataset schema is very generic to be applicable to any datasets and it needs to be adapted and mapped to meteorological datasets.

The TT-MDS should review the dataset schema provided by Schema.org and the Google implementation to recommend a set of mandatory information elements that should be provided by the different WIS portals to improve the accessibility to WIS portals and the indexing and presentation of product pages.

In addition, modern machine to machine API would improve the interoperability of the different WIS portals and therefore increase their presence in the web. Modern API standard such as Restful or OpenSearch should be analyzed and if judged relevant recommended for implementation. OpenSearch, for instance, is a discovery engine industry standard that can be used for discovering datasets. Best Practices for serving scientific datasets have also been developed by some of the communities that are covered by the WIS (see <http://wiki.esipfed.org/images/9/97/Combined_Open_Search_Best_Practices_v0.4.pdf>). Such APIs are also understood by the search engine industry and it is helping them accurately indexing information. Such standards are also understood by many dataset Catalogues and used to make the WIS services interoperable and therefore increasing the presence of the WIS services and products on the Internet.

* 1. **Differentiate between general product information (collection) and instance product information.**

As described in 2 TT-MDS initial WIS portal discovery services analysis, the mix and indexing at the same level of the product collection metadata and product instance metadata are penalizing the discovery services by returning as search results product family (collections) and individual instances. The user is then confused by that selection of almost identical results and believes the search is not able to provide sufficient information. The TT-MDS should express the issue and recommend a set of potential paths for implementing a solution. For instance, a standard approach is to implement a first search based on product family and once the user has chosen the product family he wants to access offer him an integrated way to select the instances interesting for him.

## TT-MDS future work

Based on the previous paragraphs, the following proposal for the future tasks for the TT-MDS is proposed. This list of tasks has to be discussed, adjusted, prioritized and approved in session.

The guidance documentation for building WIS services can be created:

1. WIS discovery challenges and potential improvements.

The different challenges for users can be developed and added.

1. Definition of the potential users of the WIS Catalogue in 2018.

The different user profiles already defined and developed can be added. A set of high-level group or categories of users can be defined to focus on the implementation of a few categories.

1. Recommendation of a set of common functionalities and usability patterns across GISCs portals.

Simple navigation can be recommended and a set of minimum common discovery components should be defined and proposed for implementation and discussion with the WIS development team.

1. The rationale for improving the WIS content and recommendation for addressing it.

The challenges regarding that topic have to be described and TT-MDS should recommend defining a list of actions to define KPIs on metadata quality, create an implementation of such service to report on the current status. Future actions can then be defined to see how to improve the metadata content.

1. Implementation of discovery industry standard and provision of a modern machine to machine interface.

The rationale describing why it is important to follow industry standards has to be outlined in the document then future actions can be defined to create a mapping between the WIS metadata and the schema.org dataset fragments. This mapping could be maintained by IPET-DD and proposed for implementation by the different WIS/GISC/DCPC teams.

1. Differentiate between general product information (collection) and instance product information.

## This issue can be described in the guidance documentation to highlight the effect it has on the discovery services and how it is potentially confusing users that want to use the services. Potential implementation for solving the issue can be proposed for discussion to the WIS/GISC/DCPC technical teams.

--------------------

# Annex 5. GroundWaterML2

**FM 233: WMLGW**

**FM 233-16 WMLGW –XML WATERML2 GROUNDWATER**

232-16.1 **Scope**

WMLGW-XML shall be used for the exchange of hydrogeological information, in XML, conforming to the “WaterML 2: Part 4 – GroundWaterML 2 (GWML2)” schemas. WMLGW-XML may be used directly to encode information about key hydrogeological entities, such as aquifers, water wells, and enclosed water bodies, as well as related measurements and groundwater flows.

Notes:

1. WaterML2: Part 4 – GroundWaterML2 was developed jointly by WMO and the Open Geospatial Consortium.
2. The GWML2 conceptual, logical and physical schemas, and related XML encoding guidelines are all described in the document 16-032r2 OGC WaterML2.0: Part 4 – GroundWaterML2. The document is available at https://www.opengeospatial.org/standards/gwml2 and the reference version of the associated XML (physical) schema is available at <http://schemas.opengis.net/gwml/>. (WMO retains a copy of the schema at t <http://schemas.wmo.int/waterml/part4/1.0>).
3. Further information on handling application schema and data modelling can be found in the *Guidelines on Data Modelling for WMO Codes* (available in English only from <http://wis.wmo.int/metce-uml>).
4. Further articles on GWML2 can be found at:

<https://link.springer.com/article/10.1007/s10040-018-1747-9>

<https://link.springer.com/article/10.1186%2Fs40965-018-0058-3>

# Annex 6 Operational Weather Radar. Best Practices Guide. Data Representation and Exchange

IPET-OWR

Operational Weather Radar Best Practices Guide

Part G: Data Representation and Exchange

Contents

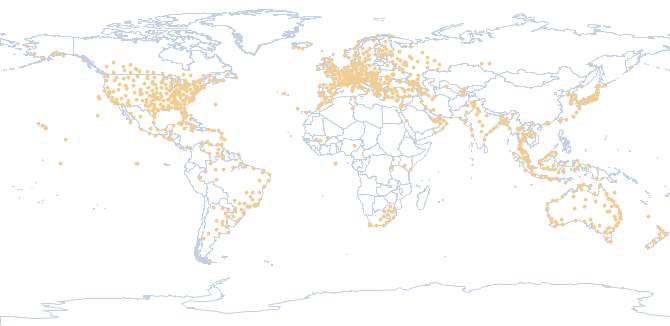
1. [Publishing Radar Site Metadata 3](#_bookmark0)
   1. [WMO Radar Database 3](#_bookmark1)
   2. [Site metadata versus scan metadata 3](#_bookmark3)
   3. [Recommendations 4](#_bookmark4)
   4. [References 4](#_bookmark5)
2. [Selection of Radar Products for Data Exchange 5](#_bookmark6)
   1. [Weather radar data levels 5](#_bookmark7)
   2. [Granularity of representation 6](#_bookmark9)
   3. [Recommendations 6](#_bookmark10)
   4. [References 6](#_bookmark11)
3. [Representation of Weather Radar Data 7](#_bookmark12)
   1. [CfRadial 2 file format 7](#_bookmark13)
   2. [Mandatory scan metadata 7](#_bookmark14)
   3. [Identification of Weather Radar Quantities 9](#_bookmark16)
   4. [Encoding of Radar Quantity Values 11](#_bookmark18)
   5. [File Naming Convention 11](#_bookmark19)
   6. [Recommendations 11](#_bookmark21)
   7. [References 12](#_bookmark22)
4. [Licensing of Exchanged Data 13](#_bookmark23)
   1. [Recommendations 13](#_bookmark24)
   2. [References 13](#_bookmark25)
5. [Methods of Data Exchange 14](#_bookmark26)
   1. [Considerations for the design of exchange mechanisms 14](#_bookmark27)
      1. [Security 14](#_bookmark28)
      2. [Resilience 14](#_bookmark29)
      3. [Performance 14](#_bookmark30)
      4. [Connection model 15](#_bookmark31)
   2. [Recommendations 15](#_bookmark32)
   3. [References 16](#_bookmark33)

## 1 Publishing Radar Site Metadata

The metadata describing a weather radar site is an important technical resource that is maintained by the operating member. When this metadata is also made available externally, it can form part of a valuable global resource.

### 1.1 WMO Radar Database

The WMO supports a database of weather radar site metadata called the “WMO Radar Database”. This database provides a centralized repository of information about the status and design of most operational weather radar networks globally. [Figure 1](#_bookmark2) shows the locations of radar sites which have entries in the database as of November 2018. This represents over 900 sites worldwide.



*Figure 1 Location of weather radars according to existing entries in the WMO Radar Database, November 2018*

The stated reasons for establishing the WMO Radar Database are:

* + - Presenting a comprehensive web-based database for network planning information and resource allocation for all members
    - Assisting in wide spread international exchange of radar data
    - Gathering radar information to protect radio-frequency spectrum allocation
    - Presenting common issues/problems and potential solutions

The submission of radar site metadata to the WMO Radar Database is encouraged as a best practice. This allows members to discover international radar sites of interest and may aid in the establishing of data exchanges.

### 1.2 Site metadata versus scan metadata

The metadata associated with a radar site is a subset of the metadata which is associated with individual radar scans. Site metadata is used to describe aspects of the radar system that do not generally change scan to scan. This includes metadata related to the infrastructure, instrument, signal processing, and

available products. These metadata tend to be long lived and typically change infrequently. Examples include the manufacturer, tower height, and beam width.

In contrast to site metadata, the metadata associated with an individual scan may change every scan, or even during a scan. Examples include the scan start time, pulse repetition frequency and range resolution. A subset of the site metadata will also typically be included with each scan to avoid the need for separate lookup of fundamental properties such as the antenna location and beam width.

Due to its infrequently changing nature, one of the major challenges faced by the WMO Radar Database is keeping the site metadata up to date. Metadata is often published once when the radar is setup and then allowed to go stale. It is considered best practice to ensure that the WMO Weather Radar Database is updated in a timely manner whenever radar site metadata changes.

### 1.3 Recommendations

1. Ensure that an appropriate National Focal Point for Weather Radar Metadata is nominated and listed within the WMO Country Profile Database
2. Submit weather radar site metadata to the WMO Weather Radar Database
3. Review and update entries in the WMO Weather Radar Database regularly

### 1.4 References

Turkish Meteorological Service (TMS) on behalf of World Meteorological Organization (WMO), 2018: WMO Radar Database, <http://wrd.mgm.gov.tr/>

World Meteorological Organization (WMO), 2018: Country Profile Database (National Focal Point for Weather Radar Metadata), <https://www.wmo.int/cpdb/workgroups/view/crm_FP%20WR>

## 2 Selection of Radar Products for Data Exchange

Operational weather radar systems can output many different products. These range from very low- level data such as a time series of samples measured by the receiver through to very high-level data such as multi-radar composite rainfall grids.

It is important that the correct level of weather radar data be targeted for exchange at the international level. If data exchanged is too low-level it will be very difficult to work with and utilize effectively. If the data exchanged is too high-level, then valuable information may be omitted which reduces its usefulness for applications.

### 2.1 Weather radar data levels

The WMO Information Model for Radial Radar and Lidar Data defines the basic levels of weather radar data shown in [Table 1](#_bookmark8).

*Table 1 Weather radar data levels*

|  |  |
| --- | --- |
| **Level Description** | |
| **0** | Data at full resolution as received at the sampling rate of the receiver. Generally only |
|  | available internal to the system. Special equipment may be required to measure and record |
|  | such data. |
| **1** | Data in sensor units also known as "time series" or "I/Q" (in-phase and quadrature) data. |
|  | Produced and processed by the instrument's signal processor. Generally not recorded except |
|  | for limited durations on operational radars. Commonly recorded on research radars. |
| **2** | Derived radar variables or moments (reflectivity, radial velocity, differential reflectivity, etc.) |
|  | at full resolution after aggregation and filtering. Organised in polar coordinates by rays, range |
|  | bins and quantities. Also, known as "sweep" and “volume scan" data. |
| **3** | Radar products which are derived primarily from level 2 data. May be in the level 2 polar |
|  | coordinates (particle ID, quality metrics, etc.), or in other coordinates systems such as vertical |
|  | profiles or Cartesian grids (CAPPI, rain rate estimates, etc.). |
| **4** | Higher order products which may include data from multiple measurements. This includes |
|  | products which composite multiple radars (mosaics) as well as those that blend data from |
|  | other sources (satellites, rain gauges, NWP etc.). |

In most operational radar networks, the Level 2 polar data is the lowest level of data which is widely available. This data corresponds to radar moments, such as reflectivity, which are organized in polar coordinates relative to the instrument (elevation, azimuth and slant range). Advanced and emerging applications such as direct assimilation for numerical weather prediction (NWP) require access to data at this level since it is the closest to 'raw' observations. For this reason, Level 2 data is considered the most suitable for international exchange purposes.

In addition to Level 2 data it may be beneficial to offer Level 3 and 4 radar products for exchange. These products may provide ready-to-use information such as Cartesian reflectivity composites or rainfall estimates. Access to such products is considered particularly helpful for users that are unable or do not wish to generate their own radar products from raw Level 2 data.

### 2.2 Granularity of representation

Representations of Level 2 weather radar data generally provide a high degree of flexibility with regard to granularity of storage. Scan data may be stored on a per-data type, per-sweep, per-volume, or even multi-volume basis. For example, a single volume scan might be stored as a set of single sweep files, or an entire day of volumes might be archived within a single file.

The choice of how granular to make a data exchange requires careful consideration. If the data is stored in large batches (e.g. multiple volumes per file) a large amount of latency will be introduced, limiting its usefulness in operational contexts. Similarly if the data is stored at a high level of granularity (e.g. one file per-data type, per-sweep) then exchanges become complex and a large burden is placed on downstream users to reassemble the data into whole volumes.

For the purpose of international data exchange, it is recommended to exchange data on a per-volume basis. This is a natural level of granularity for many radar applications while also facilitating reasonably low latency transfers for operational use. Where very low latency exchange is required the data could also be made available on a per-sweep or per-data type basis.

### 2.3 Recommendations

1. Prefer to exchange data as Level 2 (polar) data rather than higher order products
2. Prefer to exchange data on a per-volume basis
3. Offer Level 3 and 4 products for exchange as an auxiliary service when appropriate

### 2.4 References

Michelson D, Curtis M., Dixon M., Haase G., Horvat C., Joe P., Umehara A., 2018: WMO Information Model for Radial Radar and Lidar Data. Version 1.3. WMO IPET-OWR. 20pp

## 3 Representation of Weather Radar Data

The adoption of a common representation for weather radar data is important for the successful implementation of international data exchange at a large scale. While Level 2 radar data is generally output from a radar in a manufacturer specific format, exchanging data in these formats is problematic. The formats may not have openly available descriptions, require proprietary tools to work with, or not include enough metadata for downstream applications. The use of a common exchange format reduces the complexity of exchange mechanisms and eases the use of received data.

The WMO Information Model for Radial Radar and Lidar Data defines a model for Level 2 weather radar data which is suitable for use as a common exchange format. It describes the key objects, relationships and metadata that must be represented in a conceptual form, independent from the concrete implementation details.

The WMO Data Model for Radial Radar and Lidar Data elaborates on the Information Model by introducing logical, structural and representational constraints. These constraints1 form a bridge between the conceptual Information Model and the concrete description of an implementing file format or protocol.

### 3.1 CfRadial 2 file format

CfRadial 2 is an extension to the NetCDF Climate and Forecast (CF) Metadata Conventions which addresses the issue of representation of polar radar and lidar data. Together with the underlying NetCDF file format, CfRadial 2 establishes a concrete file-based implementation of the WMO Information Model for Radial Radar and Lidar Data, and the associated Data Model.

Files in the CfRadial 2 format are self-describing2 which means that they can be understood without the need for external information. In a context where files are to be exchanged internationally, this is an advantage over table driven codes such as BUFR and GRIB. This also eases the use of radar data in climate related applications where long-term archives are required.

CfRadial 2 is an extensible format which allows users to store non-standard data and metadata in the file while maintaining full compatibility with applications. This facilitates maximal retention of information during the conversion from the native Level 2 data format output by the instrument.

It is recommended as best practice that CfRadial 2 be used as the common international exchange format for Level 2 weather radar data.

### 3.2 Mandatory scan metadata

The Information Model, Data Model and CfRadial File Format specify a minimum set of metadata that must be included with each data file for the purpose of basic data integrity. [Table 2](#_bookmark15) lists these metadata and several others which are to be considered mandatory for international exchange purposes. Within this table:

1 Examples include encoding of strings as UTF-8, and permissible data packing and compression schemes.

2 Self-describing in this context means that information needed to understand the contained datasets is provided within the file itself. This includes information such as quantity names, units, valid ranges etc.

* + - **IMID** provides the identifier of the metadata within the WMO Information Model for Radial Radar and Lidar Data.
    - **Description** provides the description of the metadata as listed in the Information Model.
    - **CfRadial** provides the name of the equivalent dimension, variable or attribute which implements the metadata within the CfRadial 2 format.

The inclusion of comprehensive metadata beyond those identified by [Table 2](#_bookmark15) is strongly encouraged.

*Table 2 Mandatory weather radar metadata*

|  |  |  |
| --- | --- | --- |
| **IMID Description** | | **CfRadial** |
| **Volume metadata** | | |
| **1.0** | Instrument type, distinguishing between “radar” and | instrument\_type |
|  | “lidar” |  |
| **1.1** | Site identifier, WIGOS identifier (see below) | instrument\_name |
| **1.2** | Volume start time | time\_coverage\_start |
| **1.3** | Volume end time | time\_coverage\_end |
| **2.0** | Site longitude | longitude |
| **2.1** | Site latitude | latitude |
| **2.2** | Site altitude above geodetic datum. For a scanning | altitude |
|  | instrument this is the center of rotation of the |  |
|  | antenna. |  |
| **2.3** | Geodetic datum name | |
| **3.2** | Antenna beam width H | radar\_beam\_width\_h |
| **3.3** | Antenna beam width V | radar\_beam\_width\_v |
| **3.5** | Frequency | frequency |
| **Sweep metadata** | | |
| **5.1** | Target fixed angle | fixed\_angle |
| **5.4** | PRT mode | prt\_mode |
| **5.5** | Distance to centre of first range bin | meters\_to\_center\_of\_first\_gate |
| **Ray metadata** | | |
| **8.0** | Elevation angle | elevation |
| **8.1** | Azimuth angle | azimuth |
| **8.2** | Time of acquisition (relative to volume start time) | time |
| **8.8** | Pulse repetition time(s) | prt |
| **8.9** | Nyquist velocity | nyquist\_velocity |
| **Range bin metadata** | | |
| **11.0** | Length of range bin | meters\_between\_gates |
| **Dataset metadata** | | |
| **12.0** | Dataset identifier (user specified) | variable name |
| **12.1** | Quantity name | standard\_name |
| **12.2** | Quantity units | units |
| **12.3** | Quantity value used to indicate missing data | \_FillValue |
| **12.4** | Quantity value used to indicate no signal | \_Undetect |
| **13.0** | Identifiers of datasets which are qualified by this | qualified\_variables |
|  | dataset |  |

The site shall be identified (IMID 1.1) by its WIGOS identifier, the structure of which consists of four parts3. The part of the structure called “Local identifier” is the only part consisting of characters.

Following the ODIM NOD identifier convention (Michelson et al., 2014)4, it is suggested as a best practice that the local identifier be harmonized to a five-character string, where the first two characters are the member country’s ISO 3166-1 alpha 2 ccTLD5 code (lower case), and the latter three characters are freely-selectable (also lower case).

### 3.3 Identification of Weather Radar Quantities

The CfRadial 2 file format imposes no direct restriction on the naming of the variables used to store individual radar quantities. Rather, it specifies only a standard name attribute which identifies the basic physical quantity stored. For example, CfRadial makes no distinction between corrected and uncorrected6 reflectivity even though the difference is significant within an operational context.

To ensure that exchange of commonly used operational quantities, a naming convention is established for the dataset variables within a CfRadial file. [Table 3](#_bookmark17) identifies common operational quantities and provides the names that must be given to variables containing them when CfRadial 2 is used for international exchange. Within this table:

* + - **IMID** provides the identifier of the quantity within the WMO Information Model for Radial Radar and Lidar Data. This number is provided for reference only.
    - **Identifier** provides the name which must be used as the field data variable name (short\_name) within a CfRadial 2 file. This value corresponds to the IMID 12.0 "Dataset identifier" metadatum.
    - **Quantity** provides the value which must be used for the standard\_name attribute within a CfRadial 2 file. This value corresponds to the IMID 12.1 "Quantity name" metadatum.
    - **Description** provides the value which should be used for the long\_name attribute within a CfRadial 2 file. This value has no corresponding metadatum within the Information Model.

*Table 3 Standard Radar Quantity Identifiers*

|  |  |  |  |
| --- | --- | --- | --- |
| **IMID Identifier Quantity** | | | **Description** |
| **16.0** | DBZH | radar\_equivalent\_reflectivity\_factor\_h | Equivalent reflectivity factor H |
| **16.0** | DBZV | radar\_equivalent\_reflectivity\_factor\_v | Equivalent reflectivity factor V |
| **16.1** | ZH | radar\_linear\_equivalent\_reflectivity\_factor\_h | Linear equivalent reflectivity |
|  |  |  | factor H |
| **16.1** | ZV | radar\_linear\_equivalent\_reflectivity\_factor\_v | Linear equivalent reflectivity |
|  |  |  | factor V |
| **16.0** | DBTH | radar\_equivalent\_reflectivity\_factor\_h | Total power H (uncorrected |
|  |  |  | reflectivity) |
| **16.0** | DBTV | radar\_equivalent\_reflectivity\_factor\_v | Total power V (uncorrected |
|  |  |  | reflectivity) |

3 [http://wis.wmo.int/page=WIGOS-Identifiers](http://wis.wmo.int/page%3DWIGOS-Identifiers)

4 Michelson D.B., Lewandowski R., Szewczykowski M., Beekhuis H., and Haase G., 2014: EUMETNET OPERA weather radar information model for implementation with the HDF5 file format. Version 2.2. EUMETNET OPERA Output O4. 38 pp.

5 <http://www.iso.org/iso/country_codes>

6 Total power

|  |  |  |  |
| --- | --- | --- | --- |
| **16.1** | TH | radar\_linear\_equivalent\_reflectivity\_factor\_h | Linear total power H (uncorrected |
|  |  |  | reflectivity) |
| **16.1** | TV | radar\_linear\_equivalent\_reflectivity\_factor\_v | Linear total power V (uncorrected |
|  |  |  | reflectivity) |
| **16.2** | VRADH | radial\_velocity\_of\_scatterers\_away | Radial velocity of scatterers away |
|  |  | \_from\_instrument\_h | from instrument H |
| **16.2** | VRADV | radial\_velocity\_of\_scatterers\_away | Radial velocity of scatterers away |
|  |  | \_from\_instrument\_v | from instrument V |
| **16.3** | WRADH | radar\_doppler\_spectrum\_width\_h | Doppler spectrum width H |
| **16.3** | WRADV | radar\_doppler\_spectrum\_width\_v | Doppler spectrum width V |
| **16.4** | ZDR | radar\_differential\_reflectivity\_hv | Log differential reflectivity H/V |
| **16.5** | LDR | radar\_linear\_depolarization\_ratio | Log-linear depolarization ratio HV |
| **16.6** | LDRH | radar\_linear\_depolarization\_ratio\_h | Log-linear depolarization ratio H |
| **16.7** | LDRV | radar\_linear\_depolarization\_ratio\_v | Log-linear depolarization ratio V |
| **16.8** | PHIDP | radar\_differential\_phase\_hv | Differential phase HV |
| **16.9** | KDP | radar\_specific\_differential\_phase\_hv | Specific differential phase HV |
| **16.10** | PHIHX | radar\_differential\_phase\_copolar\_h | Cross-polar differential phase |
|  |  | \_crosspolar\_v |  |
| **16.11** | RHOHV | radar\_correlation\_coefficient\_hv | Correlation coefficient HV |
| **16.12** | RHOHX | radar\_correlation\_coefficient\_copolar\_h | Co-to-cross polar correlation |
|  |  | \_crosspolar\_v | coefficient H |
| **16.13** | RHOXV | radar\_correlation\_coefficient\_copolar\_v | Co-to-cross polar correlation |
|  |  | \_crosspolar\_h | coefficient V |
| **16.14** | DBM | radar\_received\_signal\_power | Log power |
| **16.15** | DBMHC | radar\_received\_signal\_power\_copolar\_h | Log power co-polar H |
| **16.16** | DBMHX | radar\_received\_signal\_power\_crosspolar\_h | Log power cross-polar H |
| **16.17** | DBMVC | radar\_received\_signal\_power\_copolar\_v | Log power co-polar V |
| **16.18** | DBMVX | radar\_received\_signal\_power\_crosspolar\_v | Log power cross-polar V |
| **16.19** | SNR | radar\_signal\_to\_noise\_ratio | Signal-to-noise ratio |
| **16.20** | SNRHC | radar\_signal\_to\_noise\_ratio\_copolar\_h | Signal-to-noise ratio co-polar H |
| **16.21** | SNRHX | radar\_signal\_to\_noise\_ratio\_crosspolar\_h | Signal-to-noise ratio cross-polar H |
| **16.22** | SNRVC | radar\_signal\_to\_noise\_ratio\_copolar\_v | Signal-to-noise ratio co-polar V |
| **16.23** | SNRVX | radar\_signal\_to\_noise\_ratio\_crosspolar\_v | Signal to noise ratio cross polar V |
| **16.24** | NCP | radar\_normalized\_coherent\_power | Normalized coherent power |
| **16.25** | RR | radar\_estimated\_precipitation\_rate | Rain rate |
| **16.26** | REC | radar\_scatterer\_classification | Radar echo classification |

### 3.4 Encoding of Radar Quantity Values

CfRadial 2 can store radar quantities at a range of bit depths using a linear offset and gain encoding scheme. This type of encoding is common within the native formats used by weather radars, with quantities typically being stored as either 8-bit or 16-bit integers.

When converting data from a native format produced by the instrument into CfRadial 2 it is recommended to preserve the existing encoding scheme if possible. This ensures that an exact representation of the original values is retained.

The radar quantities PHIDP, PHIHX, KDP and RR should always be represented with a depth of at least 16 bits. This is due to the increased need for precision in these types.

### 3.5 File Naming Convention

To simplify handling of data from multiple sources, a simple naming convention for exchanged files is recommended. This convention is based on the WIGOS identifier of the radar instrument, a timestamp, and a user defined product identifier. The addition of a user defined product identifier allows for the output of multiple files from a single radar at the same time. If a single product is output by the radar the identifier of "vol" is recommended. In this convention files are named using the format "**WIGOS**\_**timestamp**\_**product**.nc" where:

* + - **WIGOS** is the WIGOS station identifier for the radar
    - **timestamp** is the UTC product time in ISO 8601 basic format (YYYYMMDDThhmmssZ)
    - **product** is a user defined product name

An example of how the file name is constructed is shown in [Figure 2](#_bookmark20).



0.200100.12345.aumel\_20181110T233000Z\_vol.nc



Identifier Issuer of Issue Local Series Identifier Number Identifier



Timestamp



Product



WIGOS Station Identifier

*Figure 2 Example of radar data file naming convention*

### 3.6 Recommendations

1. Represent Level 2 weather radar data for exchange using the CfRadial 2 file format
2. Store all data associated with a radar volume in a single file
3. Include the metadata identified by [Table 2](#_bookmark15) in every file
4. Include all available metadata which is identified by the CfRadial 2 specification in every file
5. Included the following radar quantities in every file, where available based on the capabilities of the radar: DBZH, DBTH, VRADH, WRADH, ZDR, RHOHV, PHIDP
6. Store radar quantities in the file with the same encoding (bit depth) as output by the instrument when possible
7. Store the radar quantities PHIDP, PHIHX, KDP and RR with a depth of at least 16 bits
8. Name files for exchange according to the convention established in [3.5](#_bookmark19)

### 3.7 References

Michelson D, Curtis M., Dixon M., Haase G., Horvat C., Joe P., Umehara A., 2018: WMO Information Model for Radial Radar and Lidar Data. Version 1.3. WMO IPET-OWR. 20pp

Curtis M., Dixon M., Michelson D., 2018: WMO Data Model for Radial Radar and Lidar Data. Version 1.0. WMO IPET-OWR. 20pp

Dixon M., Curtis M., Michelson D., Hardin J., Kehoe K., Haimov S., 2018: CfRadial2 Data File Format. Version 2.0. NCAR. 59pp

World Meteorological Organization, 2015: Manual on the WMO Integrated Global Observing System, 2015 edition updated in 2017, WMO. 97pp

## 4 Licensing of Exchanged Data

An important consideration when exchanging data of any form is the license under which access to the data is granted. The data license that is used governs the conditions on how the data may be used and redistributed, as well as issues such as attribution and legal disclaimers.

It is recommended that weather radar data be exchanged as ‘open data’ with as few restrictions on the use and redistribution of the data as possible.

The data produced by operational weather radar networks is often seen as a significant national resource with exploitable commercial value. Given this, it is expected that licensing of data as ‘open data’ will not always be possible due to political and commercial considerations. In such cases, extra care must be taken to ensure that restrictions and conditions imposed on the use of the data are clearly stated.

Where open data exchange is permitted, it is recommended that the Creative Commons7 license be used. There are several variants of the Creative Commons license that are available. The most open variant is the public domain or ‘CC0’ license. This license waives all copyright claims over the data, allowing its use and redistribution for any purpose without requiring attribution to the original source. Another common variant is the attribution, or ‘CC BY’ license. This version also allows use and redistribution for any purpose provided that the original source of the data is attributed.

Where it is possible to exchange data under an open data license, it is recommended that either the CC0 or CC BY licenses be used since these permit the least encumbered use of the data.

### 4.1 Recommendations

1. Make weather radar data available under an open data license where possible.
2. Prefer to license open data with the Creative Commons license either as public domain (CC0) or with attribution condition (CC BY).
3. Ensure that data license conditions are clearly stated and known to all parties when open data license is not possible.
4. Identify the data license within the metadata of each exchanged file.

### 4.2 References

Creative Commons, 2018: About The Licenses, <https://creativecommons.org/licenses/>

7 <https://creativecommons.org/>

## 5 Methods of Data Exchange

When using CfRadial 2 to represent Level 2 radar data as recommended, the data payload takes the form of plain files. This allows operational data exchanges to be established using generic IT facilities which are not directly related to weather radar systems. The exact choice and design of an exchange mechanism must consider a range of technical, financial and legal requirements and may often be performed by IT or communications experts rather than radar experts.

### 5.1 Considerations for the design of exchange mechanisms

This section highlights some of the key factors that should be considered when designing an operational weather radar data exchange.

### Security

Security of the mechanism is a major concern for any data exchange. Some of the key characteristics of a secure exchange mechanism include:

* + - **Authentication** – Connections may be authenticated to ensure that data is exchanged with the correct end point.
    - **Encryption** – The data may be encrypted during transport. This is particularly important to ensure that data is not accessed outside of the terms of its license.
    - **Signing** – Digital signatures may be provided to allow integrity checking of the received data.
    - **Auditing** – Activity of the exchange may be logged to provide later analysis of security events.

### Resilience

An exchange mechanism suitable for use within an operational environment must be highly resilient. Some of the characteristics to consider include:

* + - **Reliability** – The mechanism should not be prone to failure and may be designed to meet a specific availability requirement imposed by the operational environment.
    - **Redundancy** – A redundant data flow may be established to ensure that data remains available for use operationally if the primary exchange fails for any reason.
    - **Recovery** – In the event of an outage or failure, consider how the mechanism should recover. This includes return to service procedures and issues such as the resending of missed data.
    - **Monitoring** – The mechanism may provide facilities for monitoring to ensure correct function, acceptable performance, and to assist with diagnosis of problems.

### Performance

The operational environment that is being supported by the exchange mechanism will dictate the required performance characteristics of the exchange. The primary considerations are:

* + - **Latency** – For most operational environments radar data must be delivered promptly if it is to be useful. The time between the data acquisition and delivery to the remote end of an exchange should therefore be a key consideration. This consideration extends beyond the latency of the exchange mechanism itself. All sources of latency between the radar instrument and the exchange (e.g. internal radar network, data processing chain) may contribute.
    - **Throughput** – The mechanism must be designed to ensure that it is capable of transferring the required quantity of files over time (while maintaining an acceptable latency). The throughput

requirement will vary greatly between exchanges depending on the number of radars, frequency of transmission, size of data files, number of clients etc.

* + - **Quality of Service** – The size of Level 2 radar data can change significantly based on factors such as weather conditions, scanning strategy and data processing. This means that the peak size of data files may be significantly larger than the average case. Where the throughput of a mechanism cannot be guaranteed under absolute worst case conditions (all files near peak size), a Quality of Service (QoS) feature may be provided to ensure that the most important data is prioritized.

### Connection model

The style of connection used by the exchange mechanism is an important consideration as it impacts the design of receiving systems:

* + - **Active or Passive** – In an active connection model, the receiving end of a data exchange is explicitly notified when new data is available, while in a passive model the receiver must periodically check (or poll) for the presence of new data.
    - **Push or Pull** – In a push model, the transfer of an individual data payload is initiated by the sender and 'pushes' data to the receiving system. In a pull model, the transfer of a payload is initiated by the receiver who 'pulls' the data from the sending system.

The choices between active/passive and push/pull are independent and can be used to create a variety of exchange mechanisms. An example of each combination is described below:

* + - A publish/subscribe mechanism, where a receiver initially subscribes to a data flow and then automatically receives new data over a TCP connection as soon as it becomes available, is an example of an **Active Push** mechanism.
    - A notify/pull mechanism, where the receiver subscribes to a data flow and then automatically receives notifications of new data availability over a TCP connection, but must then collect the payload itself (e.g. from a FTP server), is an example of an **Active Pull** mechanism.
    - A sender which uses FTP to automatically transfer files to a known remote host when they become available is an example of a **Passive Push** mechanism.
    - A sender which publishes files to an FTP server without otherwise notifying receivers is an example of a **Passive Pull** mechanism.

### 5.2 Recommendations

* + 1. FTP8 should be avoided due to its security vulnerabilities. This is particularly sensitive, and potentially impactful, to Members who experience challenges with capacity related to IT expertise. While it may not be reasonable to expect current message switching system solutions using FTP to be replaced in the short term, future WMO Information System (WIS) mechanisms for real-time data exchange should not be built on top of FTP.
    2. Members should consider the desired resilience and performance characteristics of the exchange when choosing between passive and active data exchange. Active mechanisms provide useful control and monitoring, at the expense of system complexity and overhead.

8 Note that the subject of this recommendation is the File Transfer Protocol (FTP), which is distinct from other similarly named technologies such as Secure File Transfer Protocol (SFTP).

* + 1. WMO data exchange mechanisms should be designed to accommodate weather-radar data payloads using the data representations proposed by IPET-OWR.
    2. Members wishing to exchange weather-radar data should use WMO data exchange mechanisms that are designed to accommodate weather-radar data payloads.
    3. Private (secure) networks can be used as an alternative to the Internet for additional safety and stability. Such networks, e.g. the Regional Meteorological Data Communication Network (RMDCN) in Europe, are private TCP/IP networks with regulated bandwidth designed for peer- to-peer data transfer. The Global Telecommunications System (GTS) is used by many European Members over the RMDCN.
    4. Alternative/redundant routing of data should be considered in exchange methods in order to increase the likelihood that data are successfully exchanged. For example, in a network with three nodes A, B, and C, where all nodes exchange data with the others, if the connection between nodes B and C is broken, data from these two nodes are routed through node A in order to reach each other. So-called multicasting or gridded methods may be a means of achieving such network robustness.

### 5.3 References

Michelson D., Curtis M., 2017: Weather radar data exchange methods. WMO IPET-OWR. 4pp

**Reasons for using NetCDF and importance of conventions**

# Annex 7 GCW requirements on netcdf

## Introduction

NetCDF provides the possibility to generate self-describing, machine independent data encoding. In order to achieve fully self-describing and machine interpretable data, conventions on how to encode data as well as the semantics required are also needed. This document complements a presentation providing an overview of considerations made in the context to [WMO GCW](https://www.globalcryospherewatch.org/) to ensure interoperability across the heterogeneous community contributing to GCW.

## WMO GCW data management

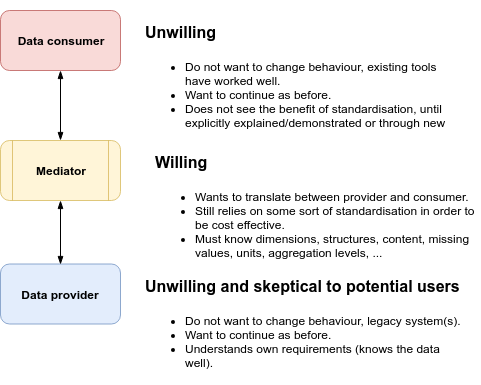
The intention of the GCW Portal is to be the entry point to datasets describing the cryosphere and form the information basis for the assessment activities of the Global Cryosphere Watch. It offers a web interface that contains information about datasets through discovery metadata provided by the data providers (or host data centre). These discovery metadata are harvested on a regular basis from data centres actually managing the data on behalf of the owners/providers of the data.

The GCW Portal utilises interoperability interfaces to metadata and data in order to provide a unified view on the datasets that are relevant for GCW activities. The GCW Portal will facilitate real time access to data through Internet and WMO GTS as requested by the user community. This require a certain level of interoperability at the data level in addition to at the metadata level. On GTS, WMO formats (BUFR and GRIB) are required and the GCW Portal can transform data into these formats in the dissemination process, provided contributing data centres are following the required standards for documentation and interfaces to data.

The GCW data management user community (data providers and consumers) is a heteogeneous community ranging from national meteorological and hydrological services through universities to independent research institutions. In the WMO context, the GCW user community is very similar to the WMO GAW user community. GCW needs to bridge between the traditional WMO communities and the

science community. The situation for data providers in the context of GCW is that many have existing systems, but that these are not interoperable. Some doesn’t have any structured data management.

The data management component is an enabling service in the sense that it identifies relevant datasets and their locations and provides an interface that can be used in the evaluation of GCW data and products. The portal will support simple visualization (generation of maps or diagrams like time series) and transformations such as reformatti g and re-projection of data, *if the data are served through the appropriate interfaces and forms*. The GCW Data Portal takes the role as a mediator between the data provider and the data consumer hiding some costs for the user community ([Figure 1](#_bookmark0)).

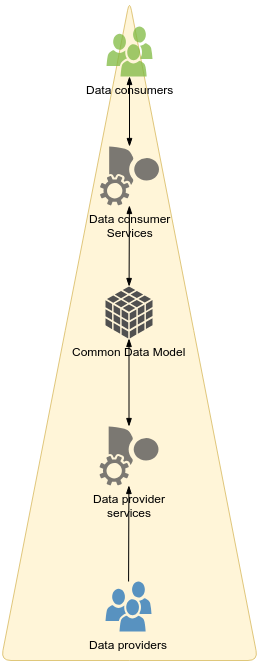


*Figure 1: The GCW Data Portal is a mediator between user communities.*

GCW data management shall integrate datasets and provides access to data and information on past, present, and future cryospheric conditions. To achieve these results, the data portal must be attached to real-time and near-real-time data management systems and to data archives. While interfacing with existing data management systems, GCW respects partnership and ownership. GCW itself will rely on distributed data management technologies and partners (e.g. CryoNet stations) to establish the GCW catalogue. This process will create a unified interface to datasets in an otherwise fragmented terrain. No information on data (discovery metadata) will be kept in the GCW catalogue without an agreement with the data producer/data owner.

In the GCW context, at least two types of metadata are relevant. One is “discovery” or index metadata identifying general characteristics of a dataset, including what was measured where and when, potential restrictions on data use, data custodians, and the available interfaces to the actual dataset. This is the type of metadata that will be exchanged within GCW. Another type, “use” metadata, is required when a

user has accessed a dataset and begins to use it. Such metadata typically include a specification of variables, units used, how missing values are encoded, and other details on the contents of the dataset. The ingested discovery metadata will be harvested from project specific, national, and international catalogues. Successful exchange of metadata will involve some degree of adaptation of systems on either side. However, in order to establish a sustainable system, the number of standards the GCW portal has to support cannot be too many. Furthermore, the actual data also has to be standardised to support integration of data among data providers in support of data consumers ([Figure 2](#_bookmark1)).



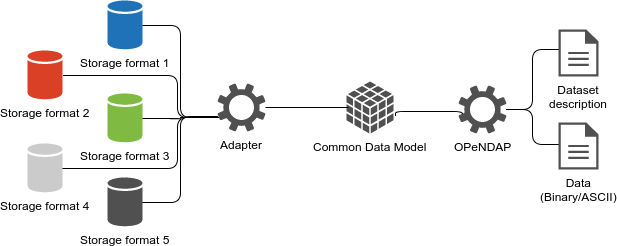
*Figure 2: Standardisation across existing data repositories in a cost efficient manner.*

## NetCDF/CF in the context of WMO GCW

NetCDF in itself doesn’t solve any interoperability issues between data providers and data consumers in the GCW context. However, when the [Climate and Forecast Conventions](http://cfconventions.org/) (CF) and the [Attribute](http://wiki.esipfed.org/index.php/Attribute_Convention_for_Data_Discovery_1-3)  [Convention for Dataset Discovery](http://wiki.esipfed.org/index.php/Attribute_Convention_for_Data_Discovery_1-3) (ACDD) are added, NetCDF files are standardised in manner which

allows building services on top of the data. Such services can be visualisation, transformation (e.g. subsetti g, reformatti g, combination of data etc).

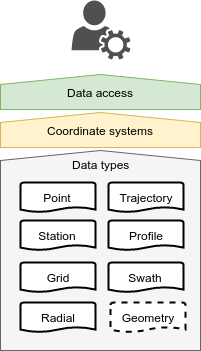
The primary issue in WMO GCW is not that data is made available as NetCDF, but that data are encoded according to CF and ACDD and made available using [OPeNDAP](https://opendap.org/). This way existing data can be integrated in GCW data management through adapters, removing the need for a full reimplementation of data centres, that map non standardised terminology into standardised terminology and structures ([Figure](#_bookmark2) [3](#_bookmark2)). The original data can be NetCDF, or it can be spreadsheets or relational databases. The mapping into a standard (CF) and application of the data streaming protocol OPeNDAP allows services on top of the data.



*Figure 3: Mapping existing non-standarised data to standardised data using adapters.*

## Benefits of using NetCDF/CF and ACDD

WMO table driven formats are primarily used within the WMO community. The process of reading, notspeaking of creating these table driven formats is difficult and in most scientific communities considered too difficult and time consuming. On the other hand, NetCDF is widely used in science communities, especially within climate modelling. The CF convention has support for gridded data, timeseries, profiles etc ([Figure 4](#_bookmark3)) which simplifies the process of combining data from multiple sources.



*Figure 4: Several types of data are supported or in the process of being supported through CF conventions.*

## Lessons learned

The CF conventions are too wide, narrower profiles are necessary to be able to build services on top of the data. The scientific communities not accepting NetCDF is often not suing any standards at all, but if supported with tools they are usually positive. Although CF is not focusing on the semantic web approach, the approach using standard names etc is useful in that context as well and part of several initiatives exploiting this approach for better discovery and use experiences across communities. The major insufficiencies of CF currently are lack of a standardised way to handle shapes (e.g. similar to KML, Shape etc) and hierarchies in data (grouping). When it comes to standardisation efforts, too few people and communities are engaged. It is not beneficial for the geoscientific communities to compete, but rather to join efforts and focus on extending the scope for standards that have sufficient momentum.

--------------------

# Annex 8. EUMETSAT’s plans for future use of netCDF-CF

## Introduction

The netCDF format enjoys widespread adoption and excellent software support for the use in encoding scientific data in the geophysical and remote sensing communities and beyond. An increasing number of robust software packages can readily consume and perform scientific processing on data stored in netCDF files. Therefore, the use of netCDF offers advantages to data producers and consumers, particularly in the case of large volumes of environmental data that will be consumed by users from outside the numerical weather prediction community. Examples of this abound in the meteorological community due to the large amount of forecast and ancillary data, as well as the large amount of observational data collected on space-borne platforms.

EUMETSAT, as a large data producer, acknowledges these facts. Current planning foresees the use of netCDF as the “native” format for data products generated using future generations of satellites such as Meteosat Third Generation (MTG) and EUMETSAT Polar Service Second Generation (EPS-SG). This means that netCDF will replace currently used, bespoke binary formats. In keeping with current practice, EUMETSAT will still convert data products to the WMO formats BUFR and GRIB on request by its member states.

The use of netCDF as a final product delivery format presents new challenges from a data governance perspective. WMO formats have a robust and well-defined governance policy, driven by the needs of Member States and extensive allocations of effort by member states for maintenance and evolutions. This ensures that needs of member states can be reflected quickly in new editions of the formats and the tables they use in order to store the controlled vocabulary that the formats use. Additionally, the controlled vocabulary of the WMO formats is an intrinsic part of the data formats’ semantics, meaning that it is difficult to encode data in a way that can be misinterpreted. This is not the case for netCDF.

NetCDF files can be encoded using arbitrary variable, dimension and attribute names. Therefore, there is no guarantee provided by the format itself that a data producer is encoding data in such a way that users can interpret it correctly. A standard must be declared within the netCDF file that clearly describes the meaning behind any names used, as well as the relationships between named objects in the file.

The Climate and Forecast (CF) Conventions are the most widely adopted standard in the area of environmental data and are widely supported by most major software libraries. They govern the use of metadata in data products, as well as the data’s presentation and the vocabulary used therein. This familiar standard is understood by many users and by their software, to the point that most users are not even aware that products they use conform to any specific standard; instead, they believe that netCDF “just works” because it is interpreted by their software transparently and correctly. It is for this reason that the CF Conventions are considered by many to be the best standard for use with netCDF data.

As early as 2008 the WMO Expert Team on Assessment of Data Representation Systems (ET-ADRS) noted that the CF Conventions are likely the most suitable standard for use in WMO systems when encoding data products in netCDF‑4 [1]. The Expert Team also noted that the Conventions have certain shortcomings. The CF Conventions are already in their second decade of existence and were created when netCDF-3 was the most recent version of the netCDF format. Experience with using the Conventions have shown that:

* Many attractive features of netCDF-4 are not covered by the Conventions that would be useful for encoding data products, and
* The Conventions have evolved largely around forecast and in-situ observation data, so that improvements could be made to fit more closely with the needs of the satellite community.

In contrast to the WMO standards, the CF community does not have dedicated resources for evolving the standard and coordinating usage. This has led EUMETSAT to increase its involvement with the CF governance process. This paper describes the current state of this involvement and proposes ways to improve the use of the CF Conventions so that they better match the needs of the WMO community.

## Compliance with CF Conventions

EUMETSAT plans maximal compliance with the CF Conventions for all netCDF products it produces. However, certain features will not be covered by the Conventions until the standard has evolved.

### Standard names

The CF Conventions define a controlled vocabulary. This vocabulary, consisting of names of variables that can be unambiguously interpreted, is encoded in the standard\_name attribute of a variable. Variables with this attribute can be looked up in a table of standard names. This table is updated at regular intervals by the CF community.

The standard\_name attribute is not required in a CF‑compliant netCDF file. However, it is advantageous to standardize the description of commonly used variables in order to avoid ambiguities. If the standard\_name attribute is used, its value must exist within the standard name table governed by the CF community in order to be compliant.

EUMETSAT uses standard names primarily in current-generation products. For the most part, these names can be found in the table of standard names. Next generation programmes make considerably less use of standard names, as the adoption process has proven quite slow due to limited resources within the CF community.

### Out-of-scope netCDF features

The CF Conventions were drafted with the intent of standardizing the use of netCDF‑3. Since its first adoption in 2003, the netCDF standard has evolved. The release of netCDF‑4 in 2008 brought several new features, among others:

* Use of HDF5 as a storage layer
* Multiple unlimited dimensions
* Groups, allowing for the hierarchical organization of data
* Several new primitive and user-defined data types

None of these new features are governed by the CF Conventions, and it is a matter of debate as to whether they can be used in products that are CF‑compliant. In any case, the fact that the CF Conventions provide no guidelines for these features of netCDF‑4 mean that both data producers and consumers may interpret data encoded using these features differently. For example, new data types may not be supported by software. Additionally, and more crucially, different users may have different interpretations of the semantics of groups.

An overview of netCDF‑4 features used in future core missions at EUMETSAT is provided in Table 1.

Table 1: netCDF‑4 features used in future core missions at EUMETSAT. “x” indicates the feature is used in products of the respective programme.

|  |  |  |
| --- | --- | --- |
| Feature / Programme | EPS Second Generation (EPS-SG) | Meteosat Third Generation (MTG) |
| New primitive data types | x | x |
| New user-defined data types | x | x |
| Multiple unlimited dimensions |  | x |
| Groups | x | x |

EUMETSAT is pursuing the adoption of guidance on the use of all of the features listed in Table 1 in the CF Conventions. This is a long-term endeavor. Additionally, participants in CGMS WGI have identified the need for guidance in the encoding of data in the viewing geometry of space-borne platforms. Although this is outside the scope of the CF Conventions, it is possible to do this in ways that are compliant with the Conventions. Further material on this matter is forthcoming.

## Collaboration with the CF community

At CGMS-46 a WGI liaison between CGMS and the CF community was nominated [2]. This has strengthened collaboration with the CF community. In the course of this collaboration, a proposal to extend the CF Conventions to govern the use of netCDF‑4 groups was submitted on 3 Aug 2018. This draft extension [3] is still in discussion now, but it is expected that it will be adopted by the release of the next version of the Conventions (v1.8).

Further extensions to the Conventions will be proposed following conclusion of discussion surrounding the current proposal [3].

Some aspects of collaboration with the CF community that have hindered the quick resolution of issues remain. For example, it has been noted by multiple data producers that the adoption of new standard names takes a very long time (in some cases years) due to the resource constraints on the side of the community; more cannot be expected from a group of experts who donate their time with only limited backing from their organisations. The same issue applies to wider-reaching change proposals, such as the one currently in discussion [3]. While this has been acknowledged by members of the Conventions Committee, it is unlikely to change. Nonetheless, contributions from new users of the CF Conventions have been welcomed.

## References

[1] [http://www.wmo.int/pages/prog/www/WDM/ET-ADRS-1/Documents.html (13](http://www.wmo.int/pages/prog/www/WDM/ET-ADRS-1/Documents.html%20(13) Feb 2019)

[2] <https://www.cgms-info.org/documents/CGMS-46_FINAL.pdf> (13 Feb 2019)

[3] <https://github.com/cf-convention/cf-conventions/pull/145> (13 Feb 2019)

## Recommended Text

The use of netCDF in WMO systems has been explored by many groups, notably by the ET-ADRS, which compared the merits of netCDF directly against those of other popular data formats. The primary shortcoming of netCDF has been described as the lack of a standard that fulfils all operational needs of WMO. As the WMO community increasingly embraces the use of netCDF, it is important to ensure that data products continue to be unambiguously interpretable by all users. The CF Conventions are an excellent starting point in reaching this aim; they are widely adopted, well tested and well understood.

Nonetheless, the CF Conventions have been created primarily by research users and due to a very slow governance process have not been updated to regulate many useful aspects of the new netCDF‑4 format. As a result, the Conventions do not completely satisfy the needs of the operational numerical weather prediction community. The primary reason for the slow evolution of the CF Conventions and lack of representation of the concerns of operational NWP (including concerns related to encoding observational data) is a lack of resources in CF governance process.

The WMO community would profit from:

* Making use of the existing CF Conventions when encoding data products in netCDF;
* Contributing to the development of the CF Conventions using existing processes;
* Specifying a subset of the CF Conventions that would be considered appropriate for operational use in WMO;
* Considering the use of one or more additional metadata attributes that would be similar but orthogonal to the CF standard\_name attribute that would allow the unambiguous description of variables, as well as mappings between WMO and CF metadata standards, while removing any dependency between WMO and CF governance processes.

It is proposed to explore these options in a workshop involving representatives of both WMO and CF. The workshop should pursue the goal of producing clear terms of reference for future work in contributing to the evolution of the CF Conventions and governing the use of these conventions within WMO.

# Annex 9. TT-MDG status report

## Drivers for current TT-MDG activities focus – and status

Various activity directions for TT-MDG have emerged from OPAG-ISS, GISCS' feedback, and previous METADATA activities. Another key driver is the need to align with WIS2.0 drivers and directions: such as providing guidance, and identifying metadata standard changes needed, in anticipation of the expanded content and scope of WIS and WIS metadata (such as for data services, data quality, and so on; and metadata for much more non-Bulletin data).

### 1a. IPET-DD TOR relevant to TT-MDG

*The words highlighted in yellow, in the IPET-DD TOR below are, in part, relevant to TT-MDG responsibilities. Those in green are solely the responsibility of TT-MDG*

(a) Review and further develop the WMO Core Metadata Profile, model driven code forms and WMO standards for metadata and data exchange to meet the needs of Members, WIS, WMO Programmes and cooperating organizations such as ICAO;

(b) Contribute to the review of and recommend updates to the Part D of the Manual on Codes (WMO No. 306) and the Manual on the WMO Information System (WMO-No. 1060) and associated reference and guidance material as required, publishing these in suitable electronic forms for human and automated use including codes.wmo.int;

(c) Facilitate proposals to standards development organizations, such as ISO TC 211 and OGC, changes to their respective standards that are required to meet the needs of WIS;

(d) Advise Members, technical commissions and ICG-WIGOS on WIS Discovery Metadata, the model driven code forms and associated application schema and data interoperability issues;

(e) Monitor metadata quality within the WIS DAR Catalogue and take follow-up action to address systematic issues identified;

(f) Identify implementation issues requiring the urgent consideration of the OPAG on ISS;

(g) Monitor and explore evolving data exchange and discovery technologies with the aim of improving the efficient operation and usability of WIS;

(h) Raise awareness of Members of the opportunities and risks associated with new data exchange and discovery technologies;

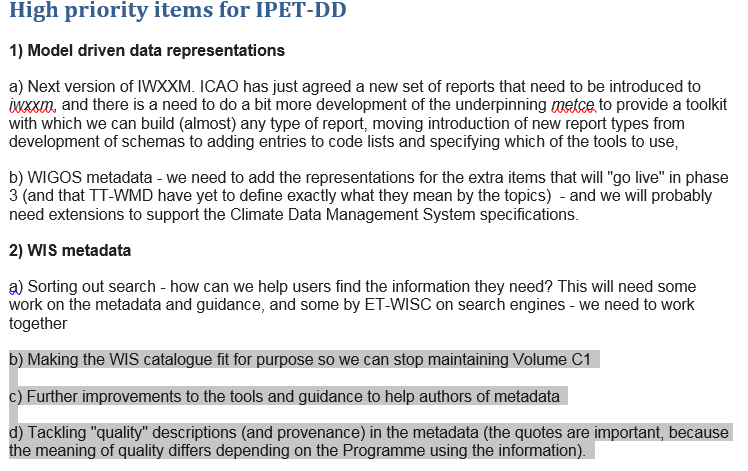
### 1b. IPET-DD Workplan, drawn from the OPAG ISS Work Plan

The table below is the IPET-DD Workplan (2017-2020), and aspects which are relevant to the TT-MDG are indicated in yellow

| **No.** | **Task** | **Deliverable/Activity** | **Due** | **Responsible** |
| --- | --- | --- | --- | --- |
| *IPET-DD01* | Improve relevance of WIS Discovery metadata | WIS users are able to find information on data and services:   * User-centred analysis of effectiveness of GISC catalogue search interface * Guidance for GISC/DCPC providers on search interface design * Guidance for metadata authors including assessment of ‘metadata granularity’ * Guidance for metadata tool developers * Guidance on making WIS metadata records discoverable via industry search engines   Periodical monitoring of metadata quality in WIS DAR catalogue   * Procedures for metadata quality monitoring- including tracking ‘missing’ metadata * Metadata catalogue monitoring tool(s) | Q2 2018  monthly | IPET-DD |
| *IPET-DD02* | WMO Core Metadata Profile (WCMP) compatible with community best practice | Releases of WCMP amended to accommodate needs of Members and Programmes (fast-track)  Assessment of value to be realised from migration of WCMP to ISO 19115-1:2014 & ISO/TS 19115-3:2016 | Biannual | IPET-DD |
| *IPET-DD03* | **Extend WMO data model(s) to address additional WMO information exchange requirements** | Updates to WIGOS METADATA Application Schema in support of OSCAR implementation  Updates to IWXXM in support of ICAO Annex 3 Amendments 78 and 79 and ICAO SWIM  Application Schema (and associated code form) to encode climate station metadata and climate data streams  Updated Guidelines on data modelling (reflecting changes to METCE etc.)  Guidance regarding use of model-driven code forms available to Members |  | IPET-DD |
| *IPET-DD04* | Publish standard vocabularies in support of WMO information exchange requirements | Consistent terminology  Definitions and language translations added to WMO Codes Registry  WMO data models (e.g. METCE, OPM, IWXXM, WIGOS METADATA) published as ontology resources within WMO Codes Registry |  | IPET-DD |

### 1c. High Priority items for IPET-DD, as advised at Meeting 1, 2017-05

The full list (from <https://wiswiki.wmo.int/tiki-download_file.php?fileId=3487> ) is below. Those aspect solely relevant to TT-MDG are 2b., 2c. and 2d. 2a is a collaborative effort with TT-MDS.



### 1d. TOR for TT-MDG as elaborated by the team

|  |  |
| --- | --- |
| (a) | Monitor metadata quality within the WIS DAR catalogue and take follow-up actions to address systematic issues identifier |
| (b) | Monitor effectiveness of metadata (the adequacy of the content of metadata records and its efficiency for supporting discovery), including developing an approach to assessing effectiveness |
| (c) | Assess the effectiveness of the examples for authors and update and maintain the examples on the wiki |
| (d) | Seek feedback from authors of metadata on the guidance available |
| (e) | Review and improve the guidance provided to authors, users and software developer |
| (f) | Develop training materials to support existing and new guidance for authors |
| (g) | Recommend tools and other products to assist with the creation of metadata records |
| (h) | Maintain consistency and currency across the guidance, support and training materials. |

### 1e. TT-MDG workplan evolution / prioritisation

A Workplan was developed by TT-MDG in June 2017, and the group commenced a review of previous guidance and support (a broad summary of previous support is below).   
In September 2017 the team was encouraged to move its focus towards the "IPET-DD issues list" posted on the wiki.   
After identifying the subset of issues in that list that were relevant to TT-MDG, the team struggled a little with resolving the issues, due to not knowing, nor knowing how to find out, the initiating background / context (and therefore the actual problem / need) related to a number of the logged issues. Some progress was made, with some issues still outstanding.

As a number of TT-MDG activities were downstream of anticipated recommendations still being developed by TT-MDS and WIS discussions, or were intended as collaborative actions with TT-MDS, from January 2018, a higher prioritisation was given to TT-MDS activities, and to WIS2 discussions.

## Previous "support" activities (by TT-APmd of IPET-MDRD)

* WMCP Specification v1.2 (influenced implementations) –
* TEMPLATE records …. Skeletal only
* WMCP Specification v1.3 (not yet implemented by some)
* Wis wiki:   
   <https://wiswiki.wmo.int/tiki-index.php?page=WIS+up&structure=WIS+up>
* topic guidance (on using codelists, on licencing, on what an abstract should contain, etc)
* Development of "example records" per type (BUOY, SYNOP, CLIMAT, etc)
* Guidance docs, eg
* (2015) "CGMS" Guidance Documentation on WMO Core Profile Metadata Creation For Satellite Products" <http://www.wmo.int/pages/prog/sat/documents/SAT-GEN_ST-17-CGMS-TFMI-WMOCoreProfile-Satellite-Documentation-Aug2016.pdf>
* "Guidance for creating WMO Core Profile Metadata in version 1.3" [2016] <http://wis.wmo.int/file=3291> (now included in Manual on WIS)
* YOPP Guide on submitting data to the WIS
* Additional activities conducted by other WMO streams (training, etc)

## Usefulness / effectiveness of support :

The original workplan aimed to formally survey metadata creators, as well as to commence a methodical review of WIS metadata. However, due to other commitments, the team instead relied on opportunistic consultations with colleagues involved in metadata creation, and each also consulted colleagues (particularly novice users) trying to use the catalogue.

Its clear that while Guidance and support have evolved, good implementation hasn't kept pace.

Possible reasons:

Resourcing; timeliness of support; reduced communication to potential editors about support available; difficulty for editors in finding relevant guidance at their point of need; Inadequacy of support and guidance material and tools.

The adequacy of the content in support materials will need to be assessed separately.

Other possible solutions:

🡺 help at point-of-need: consider other solutions (a community help service);

🡺 make it easier to maintain:   
consider re-design of metadata structure /hierarchy, and store key information only once, with cross-references to either another umbrella record, or use of xlinks;

🡺 make it easier for metadata creators to know what is needs improvement, and what to do:  
 … assess (then periodically monitor) metadata, with logged reports back to metadata creators.  
The above considerations will form the basis of an "metadata editor" survey

## WIS Metadata content issues

### 4a. Metadata content and "searching".

The effectiveness of the metadata is related both to :

- the efficient findability of data, and

- a user's ability to efficiently understand and assess the suitability of the data for their purpose.   
Improved metadata content can also improve interoperability, including (for instance) providing seamless user exploration between OSCAR and WIS, etc.

TT-MDS: has focussed on deficiencies in both Catalog design and metadata content in relation to searching, and navigation within a catalogue. The implementation of facetted searching was recommended at an IPET-MDRD 2016 meeting; as was better and consistent use of keywords for data parameters (and use of Data Parameter terms consistent with those used in WIGOS metadata)

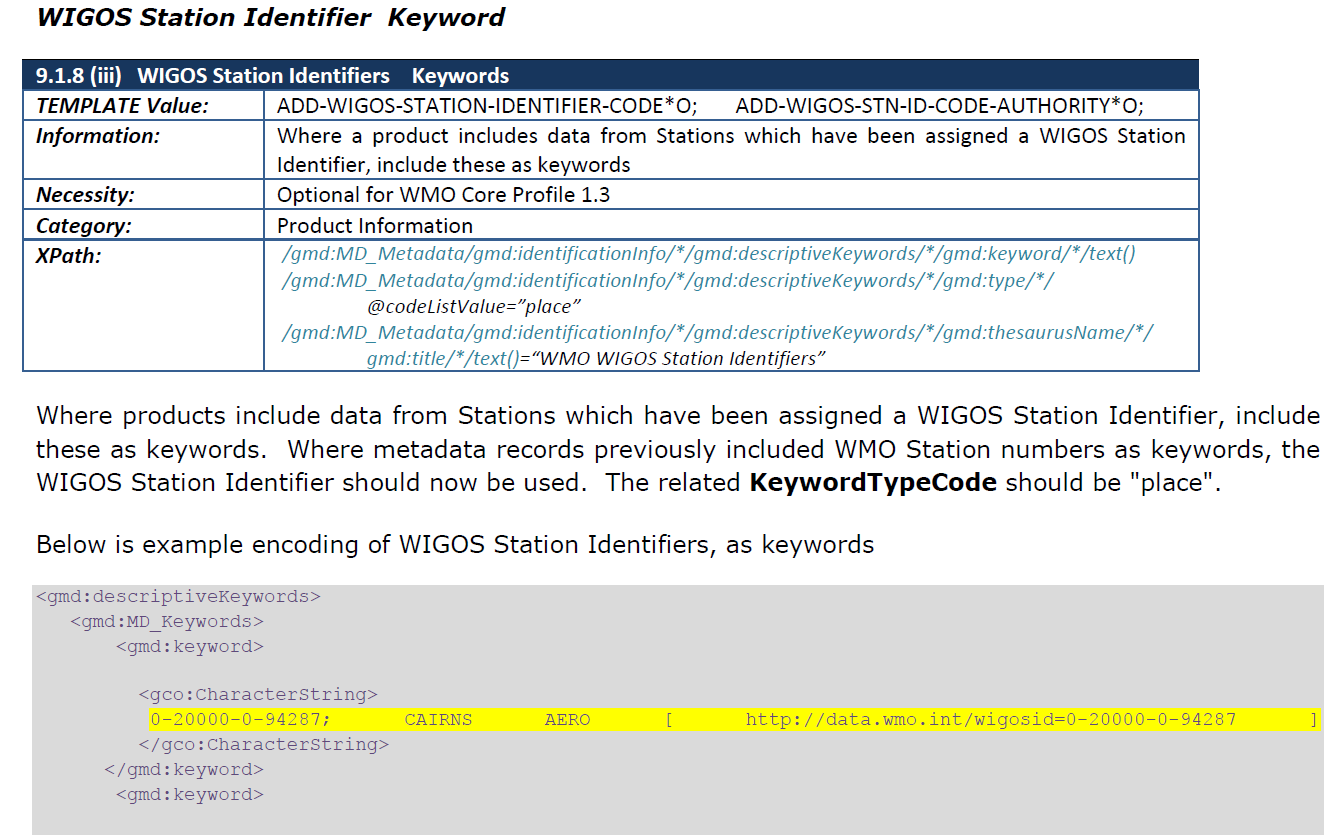
### 4b. Metadata content and "understanding the suitability" of the data

1. **Spatial/temporal**

Not all WIS catalogues offer temporal and spatial searching. However, if they did, the search results would not be ideal, due to that content being incorrect, or too sparse in detail, or inconsistently applied across records.

This is fairly important given that the spatial and temporal details of a dataset are 2 of the key what, where, when, who, how "discovery" aspects needed by users to easily understand /assess the data's suitability.   
As well as needing to be consistently applied, sufficiently detailed, and correct, spatial/temporal information also needs complementary/ additional information.  
For instance: it is useful to know whether the data is point or gridded data. If point data, its also useful have the related station id and station location. The WMCP guidelines don't currently request that detail. This additional information might also assist with metadata QC monitoring.

Spatial coverage:  
At a 2016 IPET-MDRD meeting, metadata inclusion of a region keyword was recommended.  
In the 2016 Guide (see above), use of the new Station Id was recommended.

  
  
Putting aside this more recent guidance : a quick survey of WIS records indicates that station id details are sometimes in keyword (type=place), and sometimes in Extent; and that the station id is sometimes appended by the station name. Station names are sometimes, but not always, in the abstracts. And the new station Number is in very few records.

A search of the Melbourne GISC <http://wis.bom.gov.au/openwis-user-portal/srv/en/main.home/> , for '0-20000-0\*' ……matched 10 records.

(To explore other GISC instances: SRU training Page: <http://www.wmo.int/pages/prog/www/WIS/wissruform.html> )  
  
Temporal coverage:   
Temporal coverage details in bulletin metadata is inconsistently implemented.   
Additionally, TemporalExtent should reflect the temporal coverage of data available to users at the time that they search …e.g. for bulletins, "now", rather than "1970+"  
If some GISCs couldn’t handle "now" in a temporal search, then it might be worth considering use of Xlink to automatically refresh that part of the record (eg for "latest month", the Xlink endpoint pointing to an xml fragment, could be changed monthly) ,

1. **GISC cache retention policy**

Users also aren't clearly advised that the GISC cache data disappears after 24 hours.

As a complement to improved "temporalExtent" content, the "Retention policy" of the cache should be clearly stated in bulletin records.

Ideally, records would also include a reference to where that same data is stored /accessible longer term.

1. **Formats / encodings:**

Non-WMO users would struggle to understand, if reliant of WIS metadata, that the bulletin data is encoded, and how they might decode that data.  Ideally all bulletin records for BUFR format data should include a url to the BUFR decoder.   
As well, the metadata typically don't contain a clear link to the relevant Data Structure/schema document  (the url is to the generic Codes webpage) ..see separate document reviewing the metadata record for CSAU01 .  Metadata would ideally:

* Refer precisely to the encoding used, and to the data structure and data parameters;
* Refer to Decoding tool(s)/programs;

1. **Understanding Data Access options**

Neither metadata records, nor catalogues ( at the top level) typically provide a clear pathway / or directions on how to access the bulletin data. There is sometimes a deadlink where a user would expect to link to the data.  
This applies even for WMOEssential data, for which access should possibly be ungated (but which typically is gated).   
   
To further compound "discovery" challenges, the related issue of "restricting" access rights to the data, based on the IP of the user would (in the absence of a clear explanation that they need to visit another GISC catalog) seem inexplicable to a general user.

🡺A clear statement about this, and a method of re-direct/ triage might be helpful   
 (See CSAU01 document, for further detail).

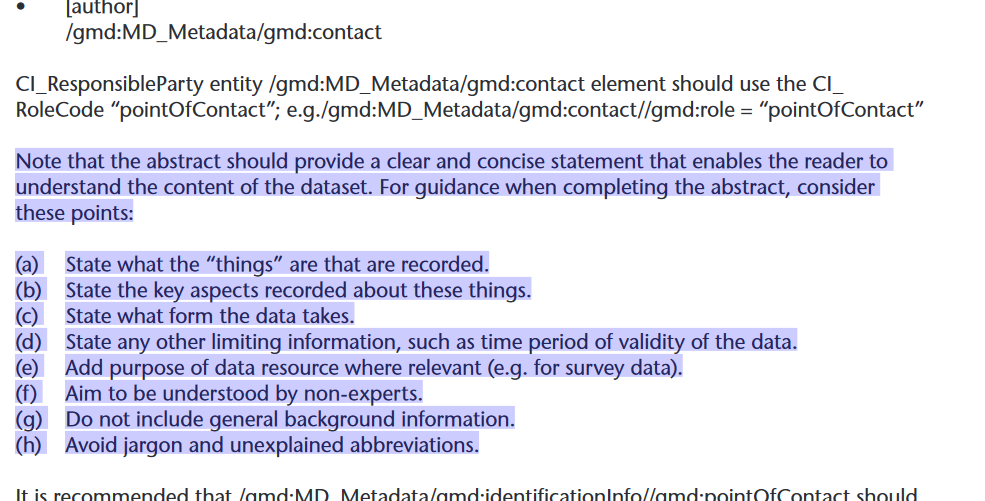
1. **Licencing**

The metadata typically doesn’t clearly advise the novice catalogue user of whether they are allowed access. For instance, while WMOEssential data is meant to be freely accessible, the absence of a link to definitions of WMOEssential obscures this knowledge.

1. **Other commonly heard comments** [from novice users] about WIS catalogue deficiencies:  
    "too much jargon in the records";   
    "no data parameter details included in records" ;   
    "no indication of quality";   
   and no sense of whats being "included, in the catalogue" / what sort of data is being covered, and likely access rights.

**General discussion:**

Although WMCP 1.3 pt.1 <http://wis.wmo.int/2012/metadata/WMO_Core_Metadata_Profile_v1.3_Part_1.pdf> states:

  
… the general feedback from users who are new to the WIS catalogue(s) is that the content of the metadata, for bulletins, is quite mystifying ( particularly with regard to describing what the data is, and how data can be accessed and used).

Recent WMO meetings have reported on 'negative' perceptions of the WIS catalogue(s). Rather than being about WIS software, or indeed about the metadata standard used, these perceptions are most likely, in addition to the need for improved Catalogue GUI design (being addressed by TT-MDS), related to the deficiencies in the metadata content itself. That is: the lack of key spatial / temporal details, lack of information about access, format, and data parameter content, as well as quite cryptic abstract content.   
(abstracts' content, for current GTS bulletin metadata, is typically what was in the TEMPLATE records created at WIS inception, which were only ever intended as a short-term solution.)

Additionally, there appears to be a perception, within the WIS implementing community, that the bulletin content, and formatting/encoding of that content is such that non-WMO users would struggle to use the data.

Some possible remediation to address this :

* Development of some rosetta-stone tools   
  (e.g. for each message type:  
   an example encoded file,   
   a reference to a decoder, and then   
   a plain English version of that file, with annotated descriptions);
* Reference to these, from each relevant metadata record;
* a more specific reference, within the metadata record, to the encoding specification
* inclusion of data params (with an explanation that not all stations can provide all values ) in the metadata record
* other?

**Other options to support improvement** of metadata content:

Xlinks: can be useful for maintaining common information in one place, and resolving the links regularly.

Previous explorations of ways to more easily improve metadata content have included exploration of use of Xlink (store information once, and refer to it many times).   
🡺While previous GISC responses have argued against use of Xlink, this might be worth further consideration.

## Monitoring to address "inadequate or inconsistent quality" across WIS records

A 2016 IPET-MDRD meeting discussed WIS monitoring – however this evolved to focus on software solutions to monitor and to provide a public "ranking" rather than focussing on what precisely should be measured, and with what ranking. Some current WIS monitoring systems might currently check for missing records   
(possibly JMA <https://www.wis-jma.go.jp/wcd/v1/detail_GISC-Tokyo-_JP_.html>).

TT-MDG need to review the scope of current WIS monitoring.

Existing validation tools measure the schema/schematron rules of 1.2 and 1.3;

However, most current WIS records would pass those rules.

We need to :

* define which aspects are most critical for "findability" and "understanding the suitability" of data
* devise a way to make it easier for METADATA creators to revise and improve their metadata   
  ( … possibly maintain common information in a collection-level metadata; and ensure that bulletin metadata refers to collection level record;   
  Or … advertise (to editors) an examplar bulletin record record, highlighting which parts need to be localised;)
* define the most effective ways to advertise support, and to advertise new metadata changes that are needed;   
  (noting that needs are always evolving. For instance, as new additional "guidance" is available, how can we efficiently integrate that into our examples, support materials, monitoring, etc, and ensure that target audiences are reached??
* define a way to check that busy GISC managers have become aware of those changes  
   (possibly survey .. with links to examples, etc)