

Metadata development and implementation

Status of the WMO Core Profile of the ISO Metadata Standard

Version 1.0 of the WMO Core Metadata standard was agreed by CBS in November 2006. This provided a standard means of describing data for “discovery” purposes (that is, finding potentially relevant data). Key challenges remain in:

Deleted:

- i) creating supporting catalogues
- ii) tools to support implementation of the standard
- iii) publicising and training in the standard, including information about how to obtain data within the standard
- iv) extending the standard to allow more detailed descriptions of data sets.

A “virtual meeting” was held in July 2007 to progress the standard towards version 2.0 that is intended to address these issues.

Key developments to the core metadata profile

Creating supporting catalogues

For users to make sense of metadata the descriptions must be clear and use common definitions. For table driven codes this clarity is provided through the tables themselves. The Core Metadata Profile uses catalogues to achieve the same result.

Many different categories of information need to be converted into catalogues to support the metadata standard. Those that have already been created are station lists, and keywords (taken from the Met Vocabulary; this can be extended under change control to meet the needs of specific programmes) and.

Comment [jk11341]: As far as I am aware, there is still a need to expand this list to accommodate keywords relevant for GAW – is there a mechanism to do that?

Success of the metadata standard will rely on the contents of the catalogues being available simply and to automated processes. This means that not only the catalogues themselves need to be populated, but a standard method for publishing the contents of the catalogues needs to be created. An emerging standard for implementing registries is ebRIM (a registry is a more general extension of a catalogue); this is being used in demonstration projects by members of IPET-MI and is likely to be used by the European INSPIRE standard.

Comment [j2]: comment on Joerg's comment: are the keywords required for GAW (and similar) part of the Core Profile or do they belong to the programme-specific extensions, in which case they would be "out of scope"? (Not sure about this myself).

In addition to developing the catalogues and methods of publishing these, a robust, operational site for hosting is required ahead of GISCs being available. This should/will be provided in the context of implementing WIS.

Comment [jk11343]: How is this done for INSPIRE? It seems, WMO could take advantage of that.

Comment [j4]: Longer-term, wouldn't all the GISCs do this (again I am not at all sure, just wondering)?

Tools to support implementation of the profile

The full ISO metadata standard is complex and lengthy and although the WMO Core Profile only requires a small subset of the potential information there is still a need to ensure that the data are created correctly and to the right standards. Much of the data passing through WIS will be suitable for “metadata harvesting” such as data transferred currently in SYNOP format. This has already been demonstrated using standard XML editing tools, but even these would benefit from a tool that could create the baseline metadata description into which the data-dependent aspects could be inserted.

There are many tools available for creating ISO compliant metadata, from simple public domain editors to full GIS systems. Taking advantage of surveys that have been done by host institutions, the IPET-MI is recommending that the GEONETWORK tool (<http://geonetwork-opensource.org>) will provide a suitable base for entering WMO Core Profile metadata, although a development effort will be required to add in the additional information required by the WMO Core Profile and to enforce the WMO Core Profile interpretation of items in the ISO standard.

Publicising the profile and training in its use

Not only is the concept of metadata abstract, but the ISO metadata standard (and, as a result, the WMO Core Profile derived from it) is complex. The virtual meeting of IPET-MI carried out an initial training needs analysis and identified the needs shown in Table 1. So far as possible, the same base materials should be used to deliver the information required to as many audiences as possible to reduce both the time to develop the material and to reduce the workload of making sure that changes to the material are implemented consistently.

Comment [j5]: purely stylistic comment: this might be a bit repetitive - seems to repeat what is said 2 paras earlier.

Table 1. Training and documentation needs. Topics and methods of delivery for metadata training and documentation that are appropriate for the user, technical and management communities. Green highlighting is used for highest priority items, Orange for medium priority and pink for the lowest priority for initial implementation. Unshaded combinations will not be targeted in the initial implementation.

Key to target U=User T=Technical M=Manager

	Web site	News-letter	Manual & Guide & Guidelines	FAQs	Examples	Slides	Workshops	1-1	Conferences	CBT
Medium →										
Content ↓										
Why metadata	UTM		UT	M		UTM		UM	UTM	U
Why Standard	UTM		UT	M		UTM		UM	UM	U
ISO relations	T		T			TM				

How to use	UM	U	UT	U	U	U	U	U	U	U
How to create	U		UT	U	U	U	U	U	U	
Impacts on systems	T		T	T	T	TM	T	T	TM	
Impacts on users	UM			UM		UM	U		UM	
Impacts on organisations	M			M		M	T	M	M	
Technical details	T	T	UT	T	T	T	UT	T		UT
Software infrastructure	T	T	T	T		T	T			
Governance rules	UTM		UT	UTM		UTM	U		T	T
Best Practice	UT	U	UT	UT	UT	U	UT	UT	UT	UT

Comment [jkl1346]: It seems to me that this cell should be green to generate some impetus and provide support for the 'willing'.

Comment [Z7]: I support the change to green

Comment [jkl1348]: By the same token, this cell should be yellow. Folks like myself learn by reviewing and adapting other folks' examples.

Service metadata – how to access the data

The ultimate aim of WIS is to allow both humans and automated systems to discover, request and use WMO data. To achieve this, not only does the metadata need to describe the data themselves, but it also needs to describe how those data can be requested. Although this is a problem that is shared with many other communities, the ISO metadata standard does not yet address in detail how this need can be met. It follows that any definition of such “service metadata” in the WMO Core Profile at version 2.0 will be made obsolete when the ISO standard is updated. Proposals for implementing service metadata for version 2.0 will, therefore, be very restricted and address only the urgent requirements.

The four types of generic user access that the WMO Core Profile will attempt to describe using service metadata will be:

- i) Fully manual request (e.g. requesting data by email)
- ii) Completion of a web form (that is, the location of the web form and the fact that it is a web form but without enough information to automate the process of requesting data)
- iii) Parameter-based retrieval using web feature services and web coverage services as defined by the OGC (Open Geospatial Consortium). In developing the service metadata for this, IPET-MI considers that it will be necessary to restrict the flexibility of access in order to make the metadata description manageable.

- iv) Univariate time series with fixed/variable geolocations (i.e. stationary site/in-situ observation(s); stationary site/remote observation(s); moving platform/in-situ observation(s); moving platform/remote observation(s))

Adding more detail to data descriptions – features

One of the challenges facing IPET-MI has been to decide on which “features” need to be described in the WMO Core Profile. In essence, a “feature” is a way of describing the information, and it is perfectly valid to describe information in more than one way. For example, a climatologist might choose to describe a set of observations as a collection of time series for different locations, whereas a synoptic meteorologist might describe the same data as a collection of global observations for a series of times. In order to make progress, IPET-MI will develop feature descriptions for data routinely exchanged on the GTS that describe:

- i) data grouped in bulletins as exchanged on the current GTS
- ii) regular grids of data, such as from models, in two and three dimensions
- iii) two dimensional imagery (such as satellite or radar images).

In creating these definitions, IPET-MI will make use of the definitions in the GRIB and BUFR tables to generate the required catalogues and registries.

Action Plan for developing the core metadata standard

An action plan is being developed to identify how these requirements will be met.