Doc. 4(6)

WMO Discovery Metadata - How do we use it?

Gil Ross 3/5/02

Now that The Expert Team on Integrated Management are focussing on a working version of the discovery metadata, we have a number of further questions to answer, some of which may be more difficult than it might seem....

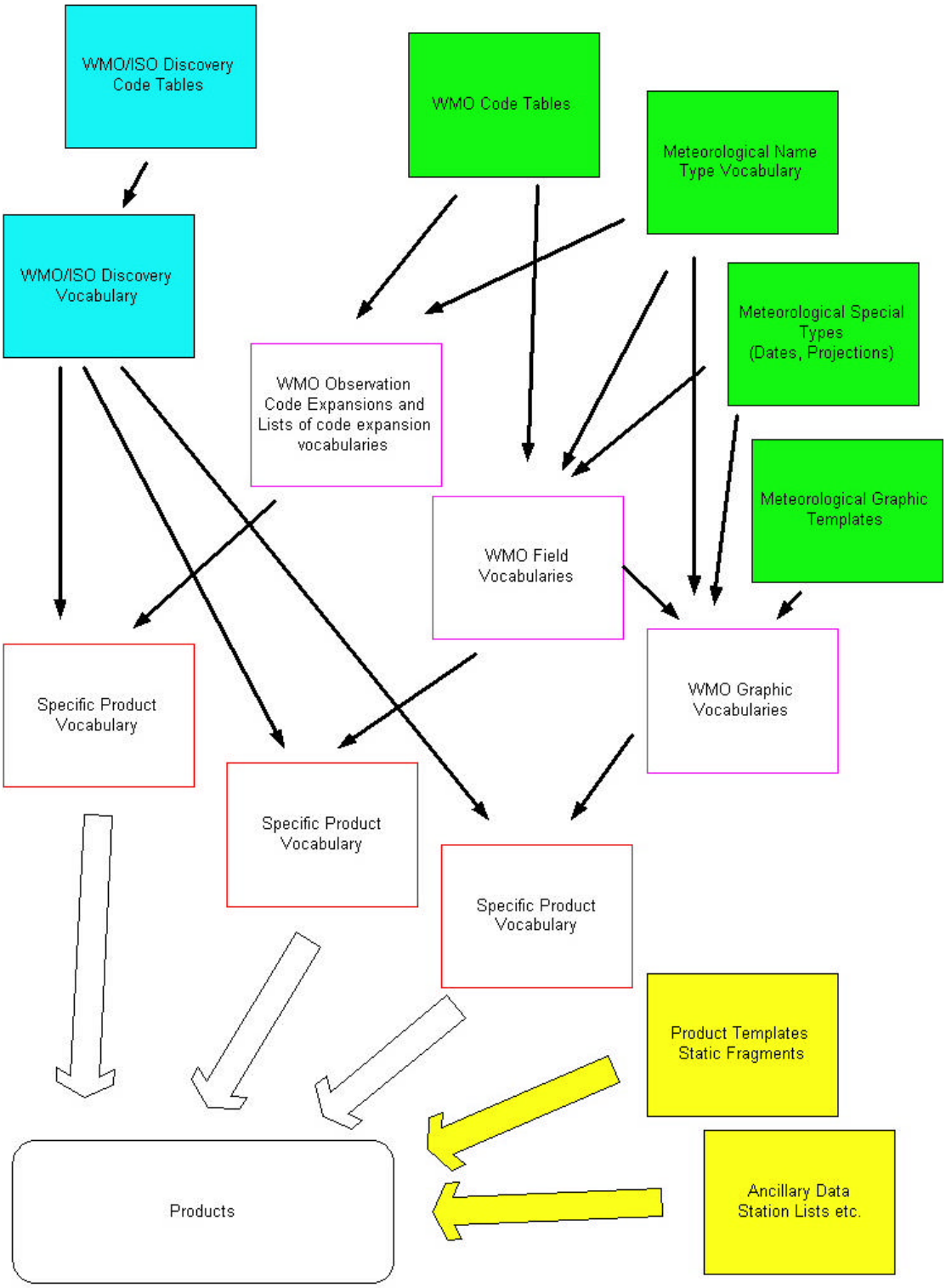
1) What changes should we make?
   a) To the metadata structure? (see Appendix A)
   b) To the metadata obligations (mandatory and conditional)? (Appendix A)
   c) How far can we deviate from ISO?
   d) Is the XML Schema workable? (Appendix A)
   e) How do we work with the ISO code lists? (e.g. we need a media code value to represent "available over WMO GTS")
   f) Do we follow ISO extension mechanisms?
   g) Is the keyword mechanism acceptable?
   h) How do we adopt new keywords?
2) How do we publish the standard and schema?
   a) Are we sure that XML and XMLSchema are the appropriate mechanisms to use?
   b) What should our "NAMESPACE" policy be?
   c) What documentation do we supply for users? (e.g. DIS19118??, UKGovernment guidelines)
   d) Do we replicate the UML documentation?
   e) What guidance do we give for developers?
   f) How do we maintain and develop the metadata standard?
   g) How do we operate versions? - Configuration management matters?
   h) Do we maintain reference links to ISO (Technical Committee (TC) 211, Geographic Information/Geomatics)? (Appendix B has a list of the ISO?TC program of work for Geographic information.)
   i) What about other links Dublin Core, FGDC (Federal Geographic Data Committee) Open GIS/ (OpenGIS Consortium) AskGiraffe (UK Geographic metadata repository).
3) How do use the metadata standard?
   a) Who gathers the instance data and how?
   b) What data do we gather? From whom?
   c) How do we store, search and retrieve the metadata?
   d) Do we want to use RDF summarisations? Or do we use the metadata directly?
   e) How do we link the metadata to the data?
4) How do we progress beyond the Discovery Metadata?
   a) We (WMO) need to consider a requirement.
   b) See Charles Sanders' paper to Expert Team On Data Representation And Codes (Inf 3 of this meeting arrangements)
   c) What Current WMO documents would benefit from XML conversion?
   d) Individual WMO Member may well progress down different routes. Should wee request information on best practice?
   e) Should we create a Roadmap?

Appendix A
Suggested Metadata modifications

1) Element 9 metadataCreationDate is type xs:dateTime - should it be xs:date?
2) Element 15 identificationInfo sub-element 33 descriptiveKeywords is WMO_KeywordsOrFreeKeywordsType - discuss
3) Element 15 identificationInfo sub-element 39 dataLanguage is type xs:language and is mandatory. What about WMO codes? What language are they?
4) Element 15 identificationInfo sub-element 41 topicCategory is MD_TopicCategoryCodeType with limited codes - expand?
5) Element 15 identificationInfo sub-element 45 dataExtent is of EX_ExtentType which is a choice (i.e.only one) of description geographicElement, temporalElement or verticalElement. Should it be a variable set of all which are all optional except for description?
6) Element 15 identificationInfo sub-element 45 dataExtent is of EX_ExtentType sub-elements geographicElement, temporalElement and verticalElement are inadequately described for WMO uses. How do we make them useful?
    a) geographicElement is EX_GeographicExtentType or polygon or bounding box. These are specified really as the true Earth. What about other co-ordinate systems?
    b) temporalExtent is EX_TemporalExtentType date "from" and date "to". We need a WMO temporalExtentType.
    c) verticalExtent is EX_VerticalExtentType and is also inadequate.
7) Element 17 distributionInfo sub-element 292 mediumName is type MD_MediumNameCodeType should have a code for availability on WMO GTS (rather than just onLine). Shouldn't it be allowed more than 1 type?
8) Element 19 dataQualityInfo is mandatory with dataQualityReport mandatory. Should it be?

We should also discuss the concepts of "mandatory" and "conditional" in the case of Xinclude'd static metadata.

Appendix B
An Outline Vocabulary Schema Hierarchy



Modular construction of vocabularies.

XMLSchema allows modular inclusions of other vocabularies by the "Include" and "import" mechanisms. Include can include and redefine, which allows re-use of higher level vocabularies. Import is similar, except it incorporates the vocabulary like a macro.
XSLT stylesheet Language Transform work on all XML documents and allow wholesale modification of XML (and therefore XMLSchema vocabularies). This can be done to create static XML or XMLSchema or it could be done in real time.
This might be used to replace or transform the ELEMENT names rather than Content, allowing language conversion to be performed, (assuming the vocabularies are available with the translation content). XSLT can also do units conversion of content numbers, and could translate WMO Code tables into SI Units. Similarly WMO station numbers could be replaced by Station names and positions from a table of station data.
XML Include mechanism can include XML documents within XML documents. For example, static discovery metadata can be included modularly within live documents or bulletins via this mechanism to reduce sizes of documents for data storage and transmission. Discovery Metadata for Met Office addresses, contact numbers, use constraints etc, need not be included with every document (unless the function is discovery). Basic identification metadata plus a link to the static data is an effective solution.

Blue box: The Discovery Metadata Vocabularies:
WMO/ISO Discovery Code Tables e.g. WMO19115CodeLists.xsd
WMO/ISO Discovery Vocabulary e.g. WMO19115Version2.xsd. this xs:include's the code lists in WMO19115CodeLists.xsd

Green box:
WMO Code Tables e.g. XML vocabularies describing GRIB, Table-Driven Data Formats (BUFR, CREX), SYNOP etc.
Meteorological Name Type Vocabulary e.g. multiple language dictionaries
Meteorological Special Types (Dates, Temporal sequence, Projections)
Meteorological Graphic Templates e.g. SVG Templates for station circles

Purple rim
WMO Observation Code Expansions and lists of Code Expansion Vocabularies expanded codes, perhaps in different languages
WMO Field Vocabularies describing words specific to NWP field data
WMO Graphic Vocabularies describing graphical product (e.g. topography, coastlines, contours, gridlines, labels, fronts, jets, centres etc.

Yellow box:
Product Templates, Static Fragments e.g. static metadata, pre-formed XML documents where only the synoptic content varies.
Ancillary Data, Station Lists etc. larger static documents e.g. from a large database, for example, where the database is queried for a small section of relevant data.

Red Rim:

Specific Product Vocabularies individually tailored for special customer types, e.g. conversion to Fahrenheit or US units; internationalisation for different languages; markup for flat files; database files or comma separated variable files

Products A wide range of different products derived from observation and or forecast data, in XML or XHTML

Not listed:

XSLT Stylesheet transformations
CSS Cascading stylesheets
Process sequence
Specific applications to process data