

World Meteorological Organization

**IPET-MDRD TASK TEAM ON
AVIATION XML SECOND
MEETING**
TOKYO, 25-28 JUNE 2013



TT-AvXML-2/Doc 3.5(2)

Submitted by: Toyoda

Date: 08.V.2013

Original Language: English

Agenda item: 3.5

FEEDBACK ON GAW METADATA RECORDS



IPET-MDRD's review of WIS Discovery Metadata made by GAWSIS

gawsis.md

Markdown

summary

- no problem
 - well-formed XML
- some issues in
 - validity against XSD of ISO 19139
 - non-XSD requirements of ISO
 - recommendations in WCMP (WMO Core Metadata Profile)
 - use of "googleMaps" as geographicElement

details

metadata set

- 8896 XML files linked from <http://gaw.empa.ch/gawsis/xml/> (<http://gaw.empa.ch/gawsis/xml/>)
- filenames containing '+' seem to be inaccessible
- Toyoda checked 779 samples, which is 1/10 of the whole set excluding inaccessible files

well-formedness of XML

xmllint gave no warning on ill-formed xml

validation against XSD of ISO/TS 19139

Needs correction for following three types of error:

invalid character for ID type

Attributes id and gml:id are both type xsd:ID where characters comma (,) or equal sign (=) are not allowed.

gco:CharacterString cannot repeat

XSD doesn't allow elements gco:CharacterString repeated:

NG:

```
<gmd:voice>
<gco:CharacterString>+421 25 4776196/9415342</gco:CharacterString>
<gco:CharacterString>+421 (7) 59415342</gco:CharacterString>
</gmd:voice>
```

Could be:

```
<gmd:voice>
<gco:CharacterString>+421-25-4776196/9415342 +421-7-59415342</gco:CharacterString>
</gmd:voice>
```

elements typed empty

Elements `gml:descriptionReference`, `gml:usesVerticalCS`, `gml:usesVerticalDatum` are sometimes given text content "NA", which is not allowed by XSD. These GML elements can have attribute `nilReason`, which should be used instead to indicate missing value due to "not applicable" or "unknown".

NG:

```
<gml:descriptionReference>NA</gml:descriptionReference>
...
<gml:usesVerticalCS>NA</gml:usesVerticalCS>
```

Could be:

```
<gml:descriptionReference nilReason="unknown" />
...
<gml:usesVerticalCS nilReason="inapplicable" />
```

non-XSD requirements by ISO

Several types of requirements in ISO 19115/19139 cannot be written in XSD. I've wrote schematron to test it. <http://toyoda-eizi.net/2013/04/15gawsis/iso19139taba1.sch> (<http://toyoda-eizi.net/2013/04/15gawsis/iso19139taba1.sch>)

Conditional Elements

OK: ISO 19115/19139 contains a number of elements labeled "conditional", which is mandatory with some condition. GAWSIS samples has no problem on them.

please use `gco:nilReason`

Similar to above GML case, I see empty elements in `gmd:MD_Format` (for example http://gaw.empa.ch/gawsis/xml/ch.meteoswiss.gawsis.20201000030_UVMultiband_c3m.BND_USDA_CSU.xml (http://gaw.empa.ch/gawsis/xml/ch.meteoswiss.gawsis.20201000030_UVMultiband_c3m.BND_USDA_CSU.xml) has). There should be `gco:nilReason` attributes.

NG:

```
<gmd:MD_Format>
  <gmd:name />
  <gmd:version />
</gmd:MD_Format>
```

Could be:

```
<gmd:MD_Format>
  <gmd:name gco:nilReason="unknown" />
  <gmd:version gco:nilReason="unknown" />
</gmd:MD_Format>
```

Bounding Box Latitudes

ISO 19115 says latitude must be between -90 and 90. Unfortunately the test sometimes fails. For example the file http://gaw.empa.ch/gawsis/xml/ch.meteoswiss.gawsis.20100000010_Diffusesolarradiation_c1d.BBN_x.xml (http://gaw.empa.ch/gawsis/xml/ch.meteoswiss.gawsis.20100000010_Diffusesolarradiation_c1d.BBN_x.xml) contains following bounding box.

NG:

```
<gmd:EX_GeographicBoundingBox>
  <gmd:westBoundLongitude>
    <gco:Decimal>112.91</gco:Decimal>
  </gmd:westBoundLongitude>
  <gmd:eastBoundLongitude>
    <gco:Decimal>-43.648</gco:Decimal>
  </gmd:eastBoundLongitude>
  <gmd:southBoundLatitude>
    <gco:Decimal>153.62</gco:Decimal>
  </gmd:southBoundLatitude>
  <gmd:northBoundLatitude>
    <gco:Decimal>-10.711</gco:Decimal>
```

```
</gmd:northBoundLatitude>
</gmd:EX_GeographicBoundingBox>
```

Could be: some appropriate values (please check source information).

WCMP recommendations

naming conventions of gmd:fileIdentifier

Currently gmd:fileIdentifier values starts with "ch.meteoswiss.gawsis."

WCMP recommends a naming convention (<http://redmine.toyoda-eizi.net/projects/wcmp/wiki/FileIdentifier#WMO-naming-conventions>) (<http://redmine.toyoda-eizi.net/projects/wcmp/wiki/FileIdentifier#WMO-naming-conventions>) which in this case recommends "urn:x-wmo:md:ch.meteoswiss.gawsis:." as prefix to varying part.

role of metadata contact

Currently the role of metadata contact (gmd:MD_Metadata/gmd:contact*/gmd:role) is set to "custodian".

WCMP recommends using "pointOfContact" for this place. The gmd:role field is provided for indicating various roles of responsible parties *if* there are many contact points on dataset (gmd:identificationInfo*/gmd:pointOfContact), while it is natural to assume there is one primary contact on metadata documentation (i.e. who actually created the metadata).

If we understand right, each station data in GAW has three responsible parties:

- GAWSIS
- World Data Center
- data originator (actual observer)

It's up to the arrangement among the programme and WIS centres, but it seems like all three parties are useful for those who see the metadata. And it is not necessary to use special value of gmd:role, since the three roles can be indicated by locations in the metadata:

- gmd:MD_Metadata/gmd:contact - contact on metadata
- gmd:MD_Metadata/gmd:distributionInfo*/gmd:distributor - contact on dataset
- gmd:MD_Metadata/gmd:identificationInfo*/gmd:pointOfContact - contact on distribution

email address of gmd:pointOfContact

Current WCMP recommends documenting email address of gmd:identificationInfo*/gmd:pointOfContact.

But it looks like it is not always appropriate to publish contact on dataset, especially in programmes like GAW. The intention of above recommendation is not to affect working structure of each programme; we simply intended to provide at least one online contact point for each dataset.

If it is deemed necessary to relax the regulation, please let us know.

suggestions

metadataStandardName

Currently metadata standards are indicated this way:

```
<gmd:metadataStandardName>
  <gco:CharacterString xmlns:srv="http://www.isotc211.org/2005/srv">ISO 19115:2003/19139</gco:CharacterString>
</gmd:metadataStandardName>
<gmd:metadataStandardVersion>
  <gco:CharacterString xmlns:srv="http://www.isotc211.org/2005/srv">1.0</gco:CharacterString>
</gmd:metadataStandardVersion>
```

It's no problem and perhaps necessary for some application (though namespace srv is unnecessary), but if you wish to indicate WCMP conformance, it might be:

```
<gmd:metadataStandardName>
  <gco:CharacterString>WMO Core Metadata Profile of ISO 19115 (WMO Core), 2003/Cor.1:2006 (ISO 19115), 2007 (ISO/TS 19139)</gco:CharacterString>
</gmd:metadataStandardName>
```

```
<gmd:metadataStandardVersion>
  <gco:CharacterString>1.3</gco:CharacterString>
</gmd:metadataStandardVersion>
```

use of "googleMaps"

I find some instances of contain xlink to "googleMaps" (without domain name) with query parameters. That validates against xs:anyURI type of XML Schema, but apparently it is not possible to reference the URL. Also some people have concern on interoperability on having both xlink:href and content (gmd:EX_GeographicDescription).

```
<gmd:geographicElement xlink:href="googleMaps?name=Matorova&latitude=68&longitude=24.2397&altitude=340" xlink:title="Matorova">
  <gmd:EX_GeographicDescription>
  <gmd:geographicIdentifier>
  <gmd:MD_Identifier>
    <gmd:code>
      <gco:CharacterString>Matorova</gco:CharacterString>
    </gmd:code>
  </gmd:MD_Identifier>
  </gmd:geographicIdentifier>
  </gmd:EX_GeographicDescription>
</gmd:geographicElement>
```

resources fyi

- Inspire validator <http://inspire-geoportal.ec.europa.eu/validator2/> (<http://inspire-geoportal.ec.europa.eu/validator2/>) (please note INSPIRE requests more than WMO, and not all GAW WDC is located inside EU)
- NOAA NGDC provided a validator called Rubric which is available at http://toyoda-eizi.net/xmlvw/wmoRubricReport/http://gaw.empa.ch/gawsis/xml/ch.meteoswiss.gawsis.100000000010_C6Cl6hexachlorobenzene.HCB.inair_x10 (http://toyoda-eizi.net/xmlvw/wmoRubricReport/http://gaw.empa.ch/gawsis/xml/ch.meteoswiss.gawsis.100000000010_C6Cl6hexachlorobenzene.HCB.inair_x10)