



ISO 19115-1 – Questions and Answers

The ISO Standards are evolving.

Metadata has been revised from 19115 to 19115-1

Data Quality metadata has moved from ISO 19115 to ISO 19157.

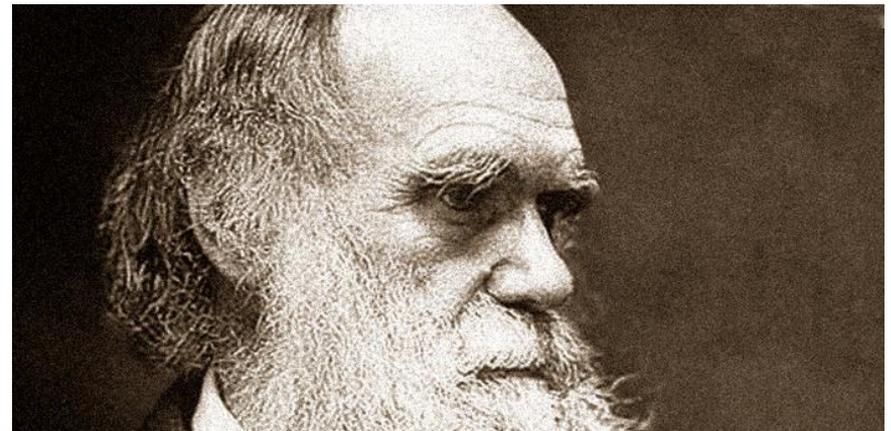
Ted Habermann

Director of Earth Science

The HDF Group

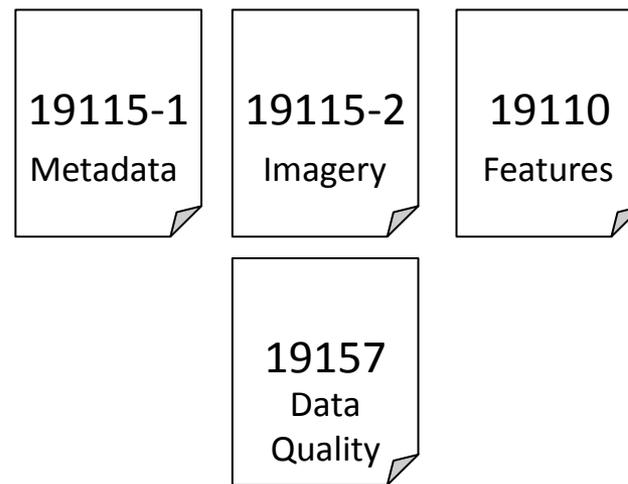
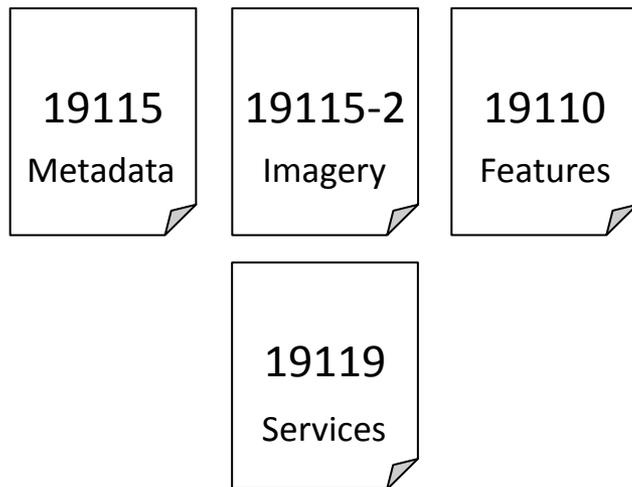
thabermann@hdfgroup.org

Evolution is good.



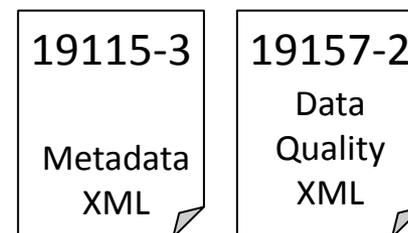
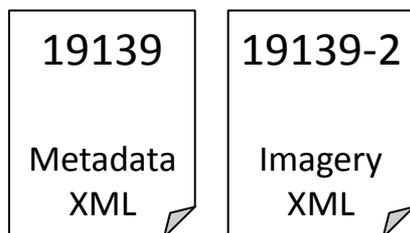
"I am confused by all of these numbers!"

Conceptual Models (UML)



Then Now

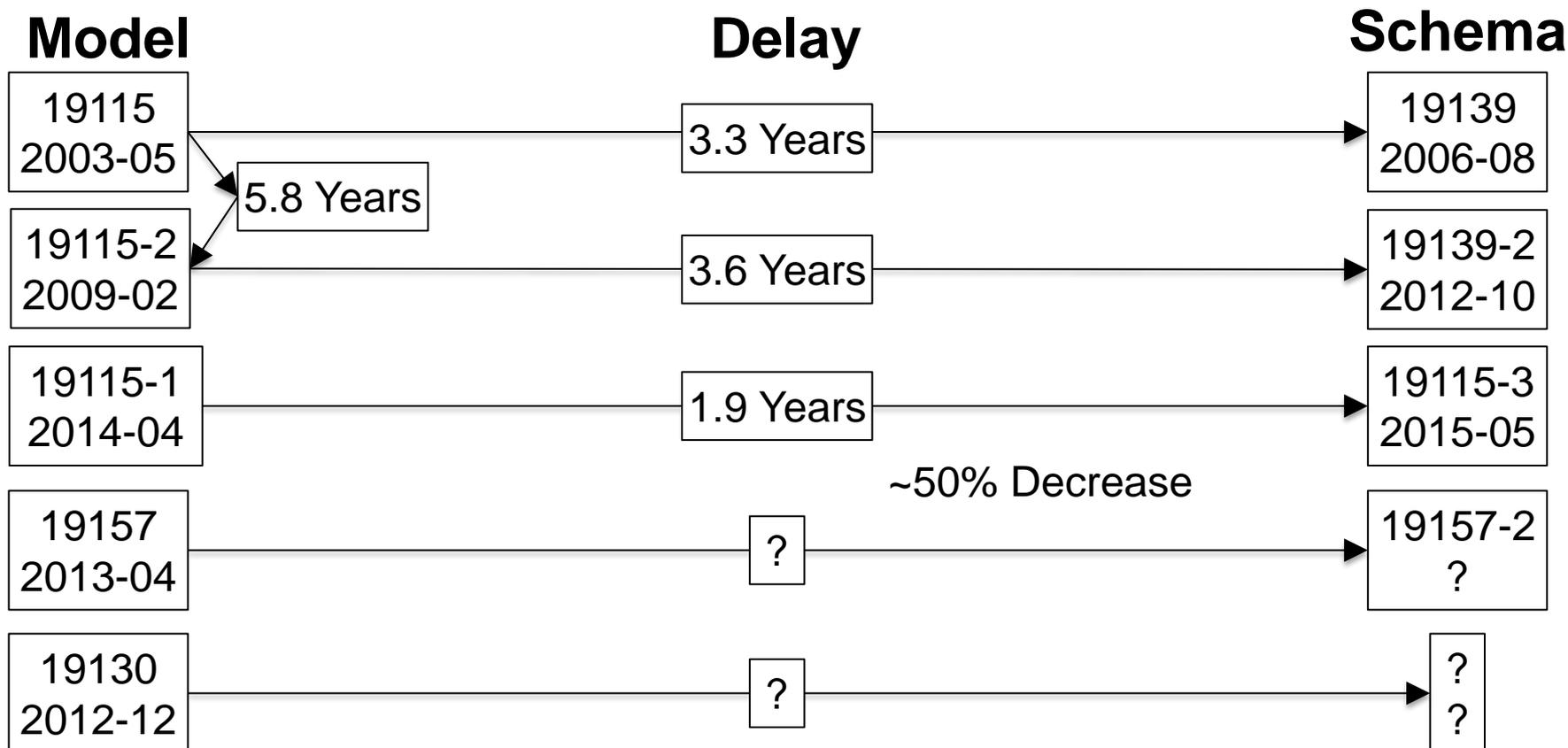
Implementations (XML)





Conceptual Models and Implementations

“Conceptual models are useful, but I need XML implementations to get going.”

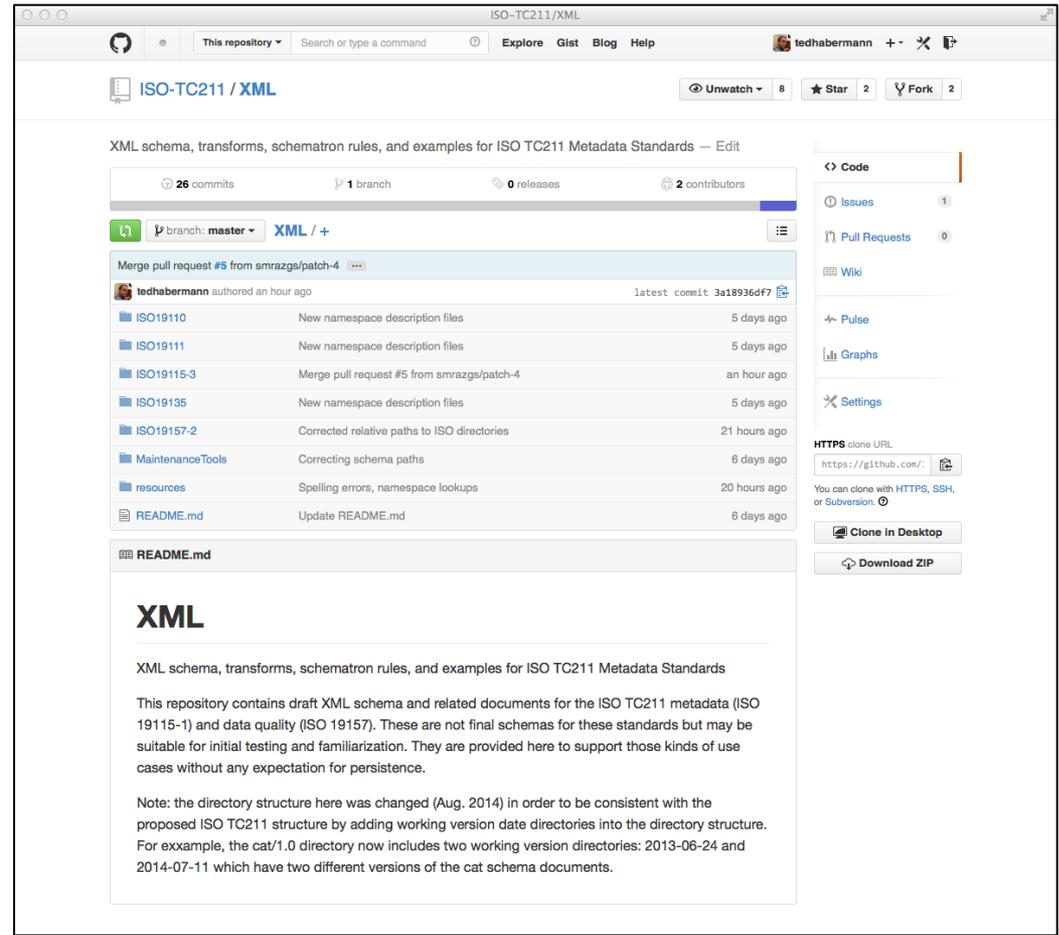


Now creating XML schemas directly from UML models using software developed in OGC Testbed (ShapeChange). This capability is also being added directly into the tool by the vendor (Sparx Systems).

“How do I become familiar with and help test the new implementations?”

The XML schemas and other resources and RDF/OWL implementations are available at:

<https://github.com/ISO-TC211>



The screenshot shows the GitHub repository page for ISO-TC211/XML. The repository is titled "XML schema, transforms, schematron rules, and examples for ISO TC211 Metadata Standards". It has 26 commits, 1 branch, 0 releases, and 2 contributors. The repository is currently on the master branch. A merge pull request #5 from smrazgs/patch-4 is highlighted. The repository contains several directories: ISO19110, ISO19111, ISO19115-3, ISO19135, ISO19157-2, MaintenanceTools, resources, and README.md. The README.md file is open, showing the title "XML" and the following text:

XML schema, transforms, schematron rules, and examples for ISO TC211 Metadata Standards

This repository contains draft XML schema and related documents for the ISO TC211 metadata (ISO 19115-1) and data quality (ISO 19157). These are not final schemas for these standards but may be suitable for initial testing and familiarization. They are provided here to support those kinds of use cases without any expectation for persistence.

Note: the directory structure here was changed (Aug. 2014) in order to be consistent with the proposed ISO TC211 structure by adding working version date directories into the directory structure. For example, the cat/1.0 directory now includes two working version directories: 2013-06-24 and 2014-07-11 which have two different versions of the cat schema documents.

“Can I migrate my existing metadata to 19115-1?”

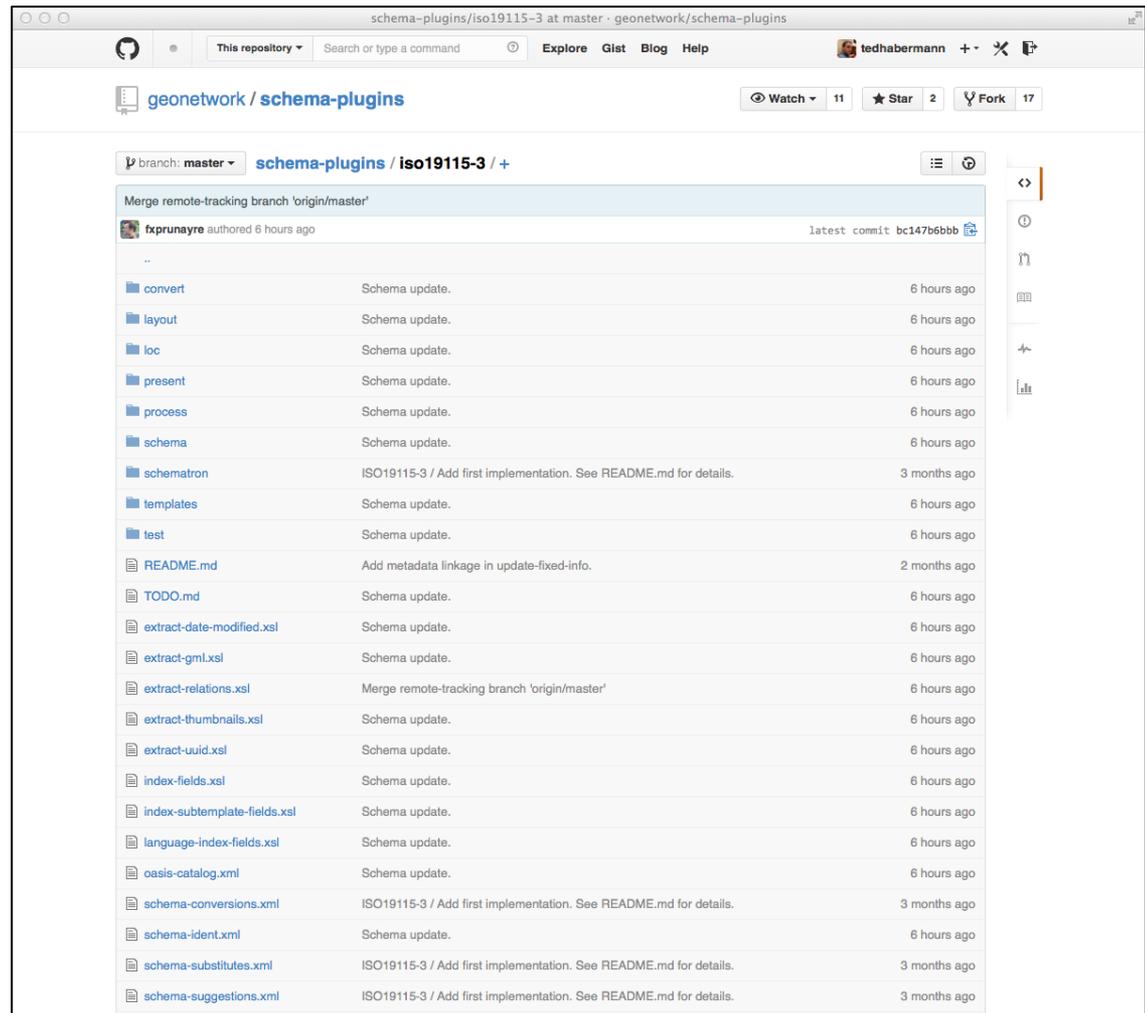
The screenshot shows an XML Editor window with two files open: '19115to19115-1.xsl' and 'Just_MI_MetadataPaths.xml'. The left pane shows the source XML with a red box highlighting a section of the `<gmd:cellGeometry>` element. The right pane shows the target XML structure. A red arrow points from the highlighted section in the left pane to the right pane. The text 'Transform Available for Testing' is overlaid on the red box.

```
<gmd:cellGeometry gco:nilReason="inapplicable"/>
```

<https://github.com/ISO-TC211/XML/blob/master/resources/transforms/19115to19115-1.xsl>

"I use GeoNetwork to manage metadata. How can I try 19115-1?"

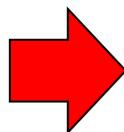
GeoNetwork 19115-3 resources are available at:
<https://github.com/geonetwork/schema-plugins/tree/master/iso19115-3#iso-19115-3-schema-plugin>



“I need to unambiguously identify metadata records in multiple repositories”

ISO 19115 identified metadata records using a single character string that often times had to be overloaded to include the information required for unambiguous identification.

05f314b8-bffe-cb8d-418e-744613aa4f01



nice opaque identifier, but who owns the metadata record or the identifier?

ISO 19115-1 brings the advantages of the MD_Identifier (+codeSpace) to the identification of the metadata record itself.

```
<mcc:code>
  <gco:CharacterString>05f314b8-bffe-cb8d-418e-744613aa4f01</gco:CharacterString>
</mcc:code>
<mcc:codeSpace>
  <gco:CharacterString>nz.govt.geodata</gco:CharacterString>
</mcc:codeSpace>
```



Tracking the Metadata Life Cycle

"I need to track when changes in my metadata happen"

ISO 19115 includes a dateStamp with the creation time for the metadata. Many other kinds of times are also important in the life-cycle of metadata.

ISO 19115-1 includes any number of CI_Date objects for the metadata which allows tracking of the metadata throughout its life-cycle. 19115-1 includes many more dateTypes than 19115.

```

<mdb:dateInfo>
  <cit:CI_Date>
    <cit:date>
      <gco:DateTime>2111-11-11T11:11:11</gco:DateTime>
    </cit:date>
    <cit:dateType>
      <cit:CI_DateTypeCode
        codeList="codeListLocation#CI_DateTypeCode"
        codeListValue="lastUpdate">lastUpdate
      </cit:CI_DateTypeCode>
    </cit:dateType>
  </cit:CI_Date>
</mdb:dateInfo>

```

19115
included 3
dateTypes

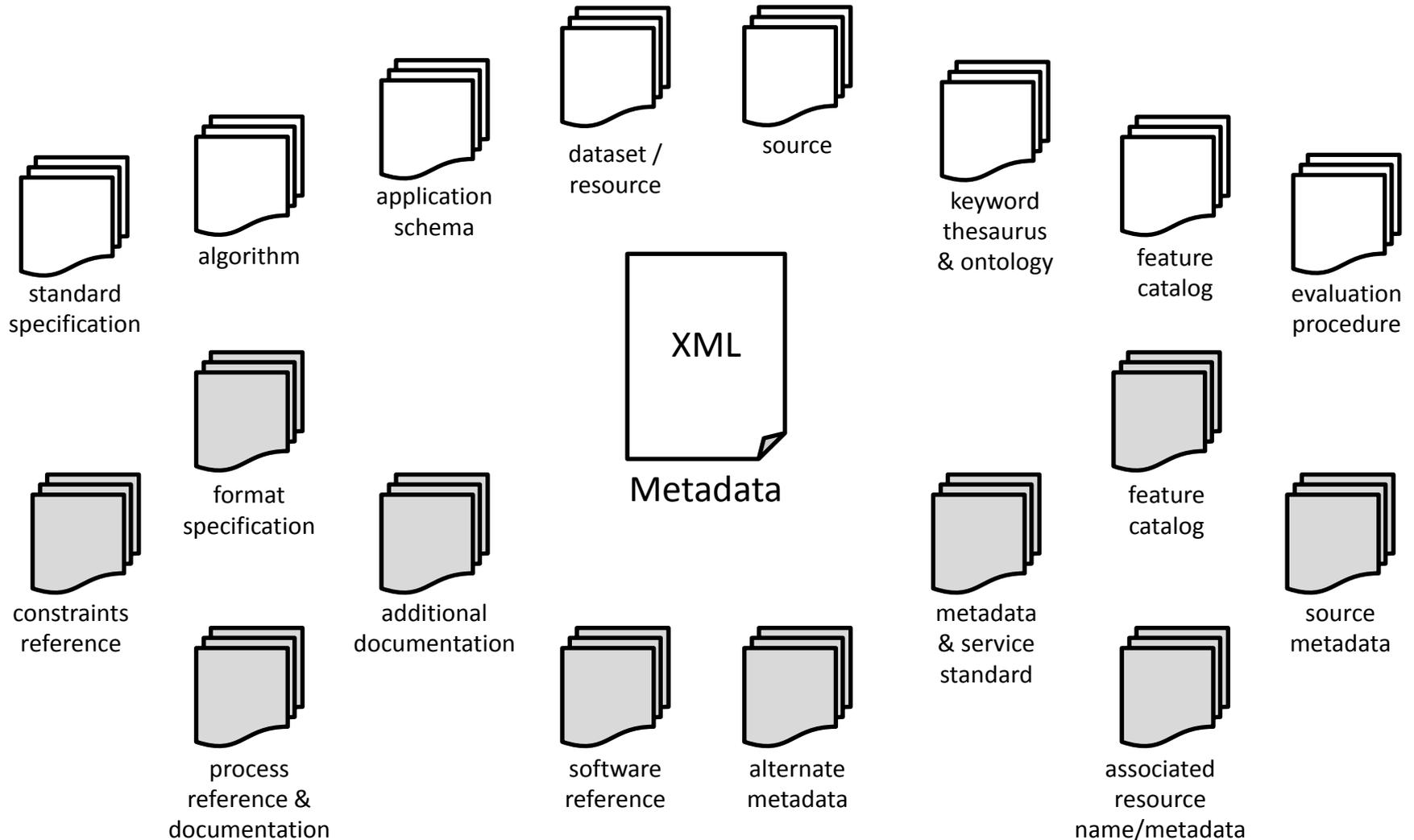
19115-1
adds 13
new
dateTypes

| <<CodeList>> MD_DateTypeCode | |
|---------------------------------|-------------------|
| + creation | + inForce |
| + publication | + adopted |
| + revision | + deprecated |
| + expiry | + superseded |
| + lastUpdate | + validityBegins |
| + lastRevision | + validityExpires |
| + nextUpdate | + released |
| + unavailable | + distribution |



Connecting Other Documentation

"I have many existing documentation resources that can help users"

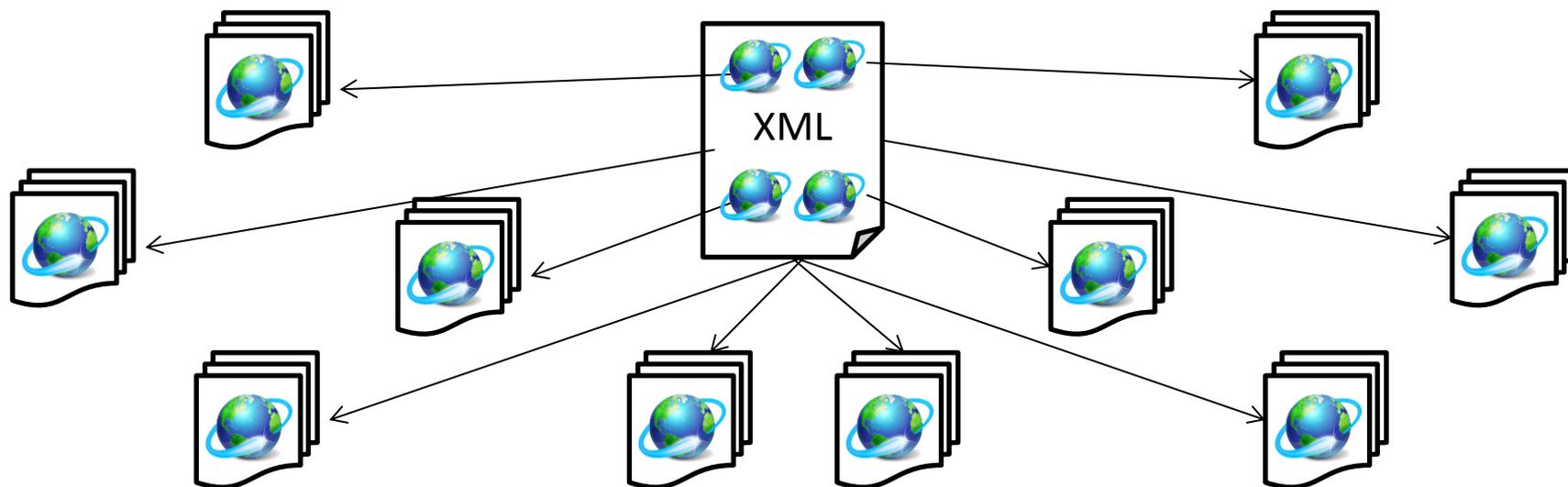


"I have many existing web resources that can help users"

The ISO 19115 CI_Citation worked well for citing books, journal articles, and other physical resources.

ISO 19115-1 added two important new elements to the CI_Citation:

1. onlineResource provides a web address for the cited resource
2. graphic provides a graphic that can be used for display of the cited resource.





Using Common Open Source Licensing

“I use a Creative Commons License for my data”

ISO 19115 included limited descriptions of constraints related to the data (useLimitations) or imposed by organizations (legal and security constraints). This made it difficult to describe commonly used open source licenses.

The Marine Community Profile extended 19115 to include Creative Commons License.

ISO 19115-1 included the necessary elements.

| MD_Commons |
|--|
| + useLimitation [0..*] : CharacterString |
| + jurisdictionLink : URL |
| + licenseLink : URL |
| + imageLink : URL |
| + licenseName : CharacterString |
| + attributionConstraints : CharacterString |

| MD_Constraints |
|--|
| + useLimitation [0..*] : CharacterString |
| + constraintApplicationScope [0..1]: MD_Scope |
| + graphic [0..*] : MD_BrowseGraphic |
| + reference [0..*] : CI_Citation |
| + MD_Releasability [0..1] : MD_Releasability |
| + responsibleParty [0..*] : CI_Responsibility |

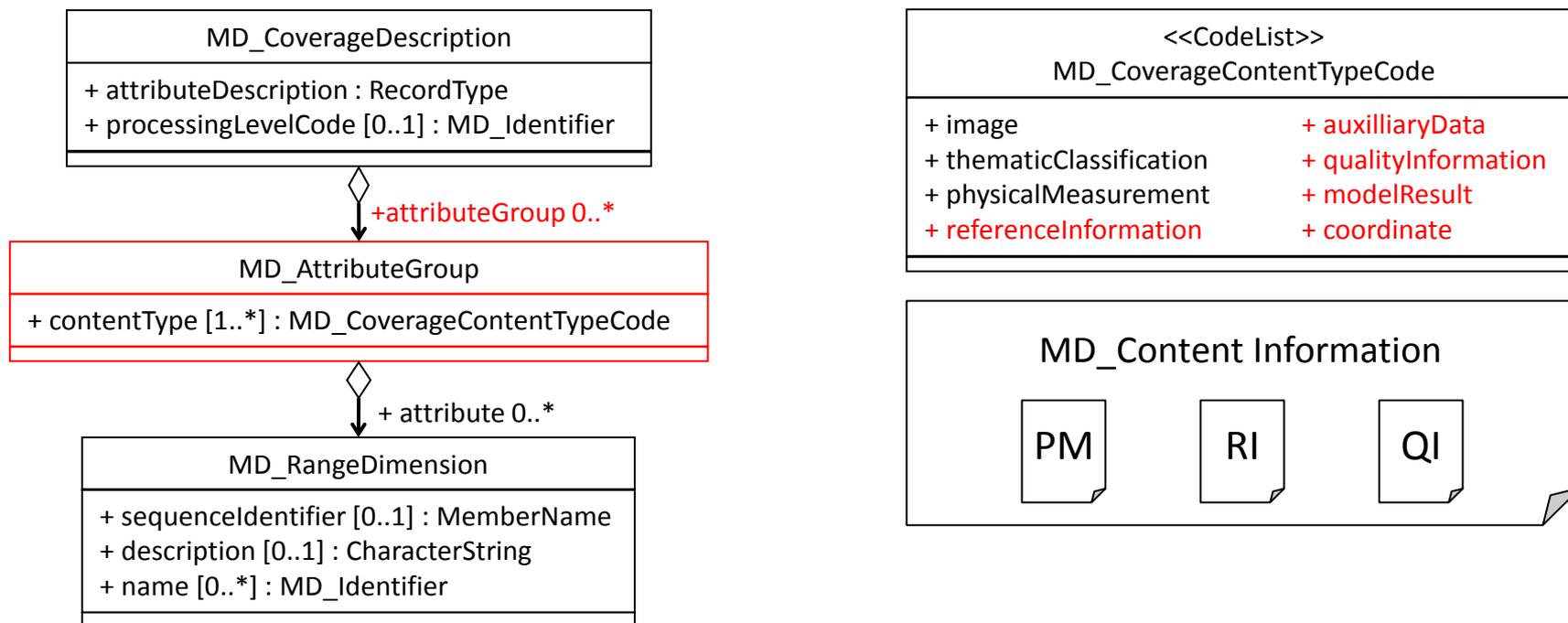
jurisdictionLink : URL -----> responsibleParty
 licenseLink : URL -----> reference
 imageLink : URL -----> graphic
 licenseName : CharacterString -----> reference
 attributionConstraints : CharacterString -----> reference

Organizing Parameters

“My datasets include measured parameters, reference and quality information”

The ISO 19115 allowed only one type of information in each contentInfo section.

ISO 19115-1 adds the capability to group similar coverages and introduces more coverage types.

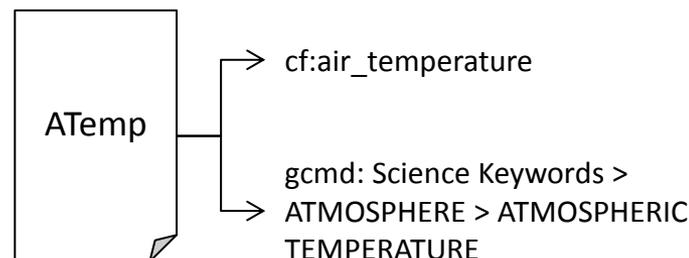
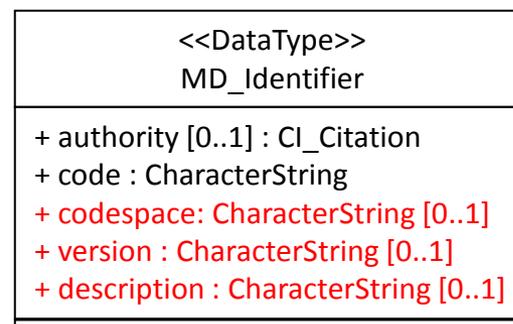
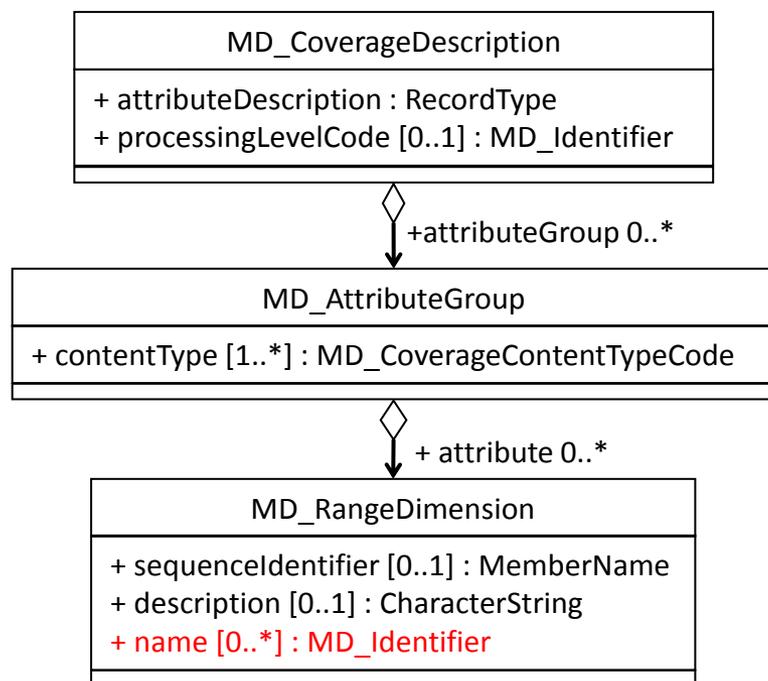


HDF Naming Parameters

“My group uses local parameter names but we need standard names to share”

The ISO 19115 sequenceIdentifier only allowed one local name for parameters

ISO 19115-1 adds the capability to add multiple names for parameters and to identify the sources for those names.





Stand Alone Quality Reports

“There are papers and web pages that describe the quality of my data.”

Papers and reports that describe data quality are StandAloneReports. Metadata can include brief descriptions of the results (abstracts) and references to any number of these (citations).

Abstract: The fire training-set may also have been biased against savanna and savanna woodland fires since their detection is more difficult than in humid, forest environments with cool background temperatures [Malingreau, 1990]. There may, therefore, be an under-sampling of fires in these warmer background environments.

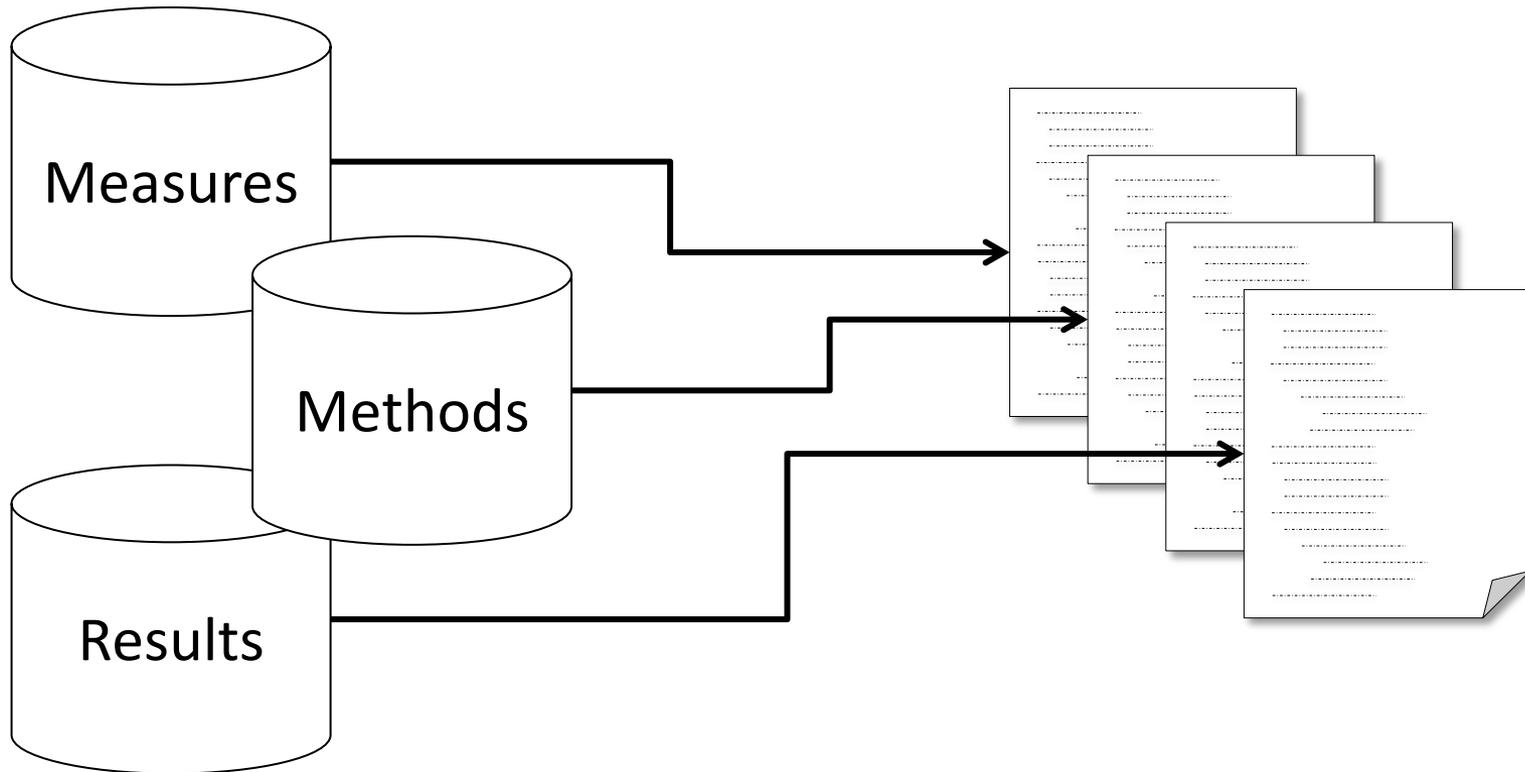
Citation: Malingreau J.P, 1990, The contribution of remote sensing to the global monitoring of fires in tropical and subtropical ecosystems. In: *Fire in Tropical Biota*, (J.G. Goldammer , editor), Springer Verlag , Berlin: 337-370.

DQ_StandaloneQualityReportInformation

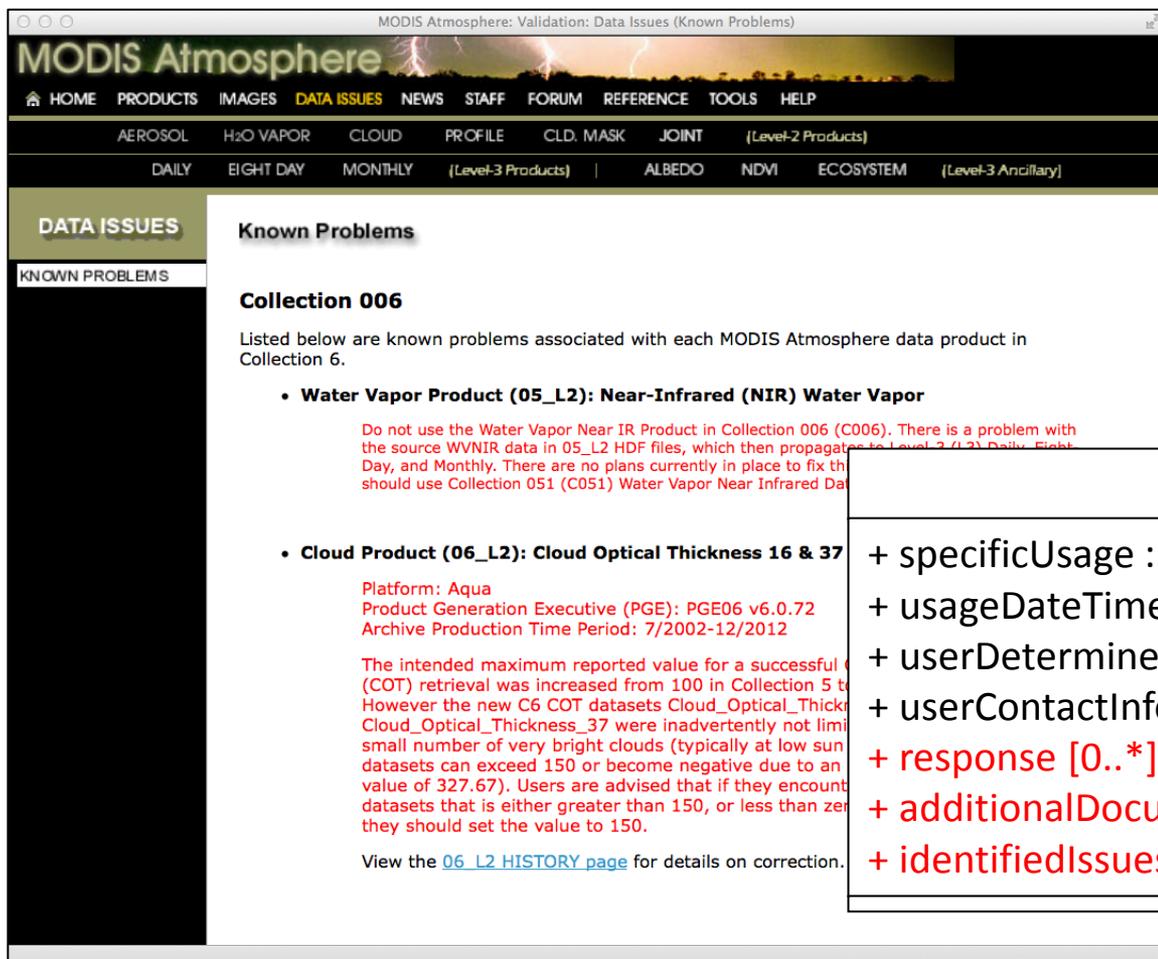
+ reportReference: CI_Citation
+ abstract : CharacterString

“My data quality information exists in databases or web services.”

Major elements of the 19157 conceptual model are separate components that can be independently connected to the metadata and reused in multiple records.



“Users increase our understanding of data quality. We need to keep them in the loop.”



The screenshot shows the 'MODIS Atmosphere' website with a navigation menu including HOME, PRODUCTS, IMAGES, DATA ISSUES, NEWS, STAFF, FORUM, REFERENCE, TOOLS, and HELP. Under 'DATA ISSUES', there are sub-sections for AEROSOL, H₂O VAPOR, CLOUD, PROFILE, CLD. MASK, JOINT (Level-2 Products), and DAILY, EIGHT DAY, MONTHLY (Level-3 Products), ALBEDO, NDVI, ECOSYSTEM (Level-3 Ancillary). The 'DATA ISSUES' section is expanded to show 'Known Problems' for 'Collection 006'. The text states: 'Listed below are known problems associated with each MODIS Atmosphere data product in Collection 6.' Two problems are listed:

- Water Vapor Product (05_L2): Near-Infrared (NIR) Water Vapor**
Do not use the Water Vapor Near IR Product in Collection 006 (C006). There is a problem with the source WVNIR data in 05_L2 HDF files, which then propagates to Level-2 (L2) Daily, Eight Day, and Monthly. There are no plans currently in place to fix this problem. Users should use Collection 051 (C051) Water Vapor Near Infrared Data.
- Cloud Product (06_L2): Cloud Optical Thickness 16 & 37**
Platform: Aqua
Product Generation Executive (PGE): PGE06 v6.0.72
Archive Production Time Period: 7/2002-12/2012
The intended maximum reported value for a successful Cloud Optical Thickness (COT) retrieval was increased from 100 in Collection 5 to 327.67 in Collection 6. However the new C6 COT datasets Cloud_Optical_Thickness_16 and Cloud_Optical_Thickness_37 were inadvertently not limited to 327.67. A small number of very bright clouds (typically at low sun elevations) in some datasets can exceed 150 or become negative due to an error in the retrieval algorithm (value of 327.67). Users are advised that if they encounter datasets that is either greater than 150, or less than zero, they should set the value to 150.
View the [06_L2_HISTORY](#) page for details on correction.

| MD_Usage |
|--|
| <ul style="list-style-type: none"> + specificUsage : CharacterString + usageDateTime [0..1] : DateTime + userDeterminedLimitations [0..1] : CharacterString + userContactInfo [1..*] : CI_ResponsibleParty + response [0..*] : CharacterString + additionalDocumentation [0..*] : CI_Citation + identifiedIssues [0..1] : CI_Citation |



Questions?

tedhabermann@hdfgroup.org



Acknowledgements



This work was partially supported by contract number NNG10HP02C from NASA.

Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author and do not necessarily reflect the views of NASA or The HDF Group.