Common Archive Observation Model: VO-DML Description and Insights

Patrick Dowler
Canadian Astronomy Data Centre

National Research Council Canada   Conseil national de recherches Canada

- ## CAOM is a metadata model

  – describe astronomical observations

  – support user-facing data centre operations: data discovery & access

  – support internal and inter-data centre operations

- ## CAOM uses minimal features of UML

  – classes, minimal polymorphism, data fields

  – enumerations, moving to vocabulary-style in some cases

  – mostly composition: cardinality important

 – some references: loose coupling

NRC·CNRC

- creating VO-DML description of an existing data model

  - wrote a parser that reads the VO-DML document, and

  - performs XML Schema validation, and

  - performs ISO Schematron validation

  - http://github.com/opencadc/core → cadc-vodml

  - use cadc-vodml from command-line or in JUnit tests

- wrote the VO-DML description (XML) by hand!!

NRC·CNRC

- technical issue: schematron validation

  - vo-dml.sch.xml contains the scheatron rules

  - only works with XSLT-based validator due to use of XSLT-specific functions

  - XPath-based schematron implementation fails

  - was noted in RFC with request for clarification, not fix

- ## model issue: enumerations

  - – can specify a symbolic name

  - – cannot specify a standard value

  - – VO-DML defers the values to some serialisation spec

  - – cannot generate an XSD from just the VO-DML description

  - – cannot implement code from just the VO-DML description

NRC·CNRC
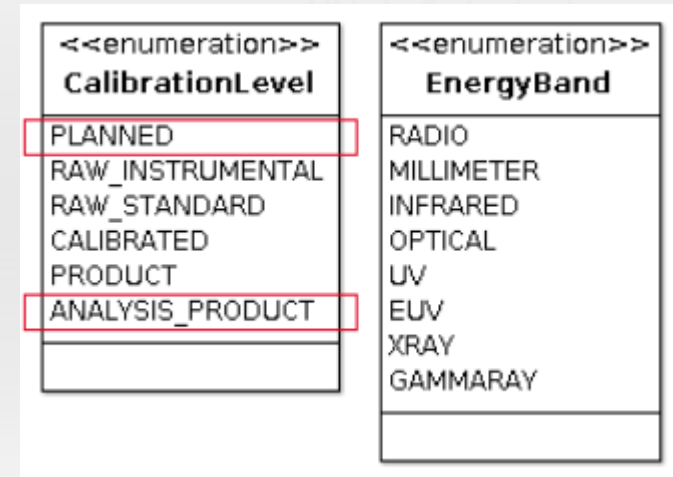
- model issue: composition cardinality for dataType

  - must specify minOccurs and maxOccurs

  - use of -1 for unlimited rather than null

  - schematron warning: "Attribute ... has multiplicity 3..666 which is STRONGLY DISCOURAGED."

  - warning discourages very common model construct: list or set of things
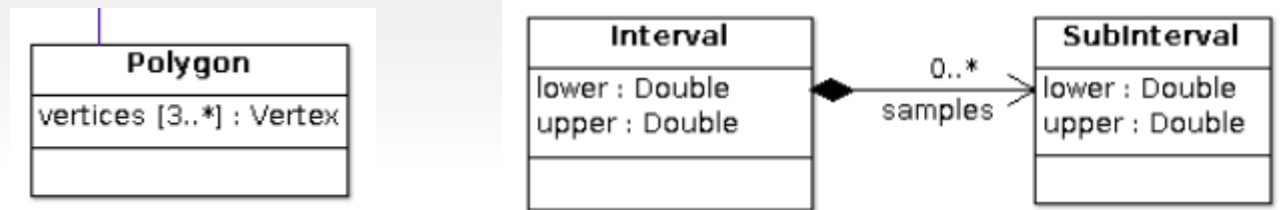
  - no such restriction/warning for objectType composition

NRC·CNRC

- experiences:

  – I quickly gave up trying to use the UML as the "source"

  – I quickly gave up on generating UML from model

  – I did not try to generate XSD from the model; already have one in use

  – model import downloads other vo-dml files; no offline build

  – enumeration lacking value seems incomplete; have to write something else

NRC·CNRC

- good experience:

  – for CAOM, the VO-DML description (XML document) serves as the normative definition; text and UML is part of documentation

  – model updates simple to do and easily validated

  – model changes easily tracked in version control

  – model updates easily validated

  – generated documentation for model

  – did use the nascent "ivoa" model for definition of primitive types

- VO-DML works for a real open-source data model

  – http://www.opencadc.org/caom2/

NRC·CNRC