

VODML Mapping: Parser Experience

IVOA Northern Fall Interop, 2016

Trieste

Tom Donaldson
22 Oct, 2016

Contents

- ST

VODML to VOTABLE Mapping

- Now that we can formally define data models (VODML), we will want to interoperably communicate instances of those objects.
- VOTABLE is the primary vehicle for VO data transport
- To communicate Data Model instances, we need an unambiguous mapping from those instances to the contents of a VOTABLE

- A mapping syntax was defined:
 - Schema:
 - https://volute.g-vo.org/svn/trunk/projects/dm/vo-dml/xsd/ext/VOTable-1.3_vodml.xsd
 - Document:
 - <https://volute.g-vo.org/svn/trunk/projects/dm/vo-dml/doc/MappingDMtoVOTable-v1.0-201607xx.docx>

Mapping Evaluation

Using the mapping document and schema:

1. Write a VOTABLE parser that “understands” the VODML parts, not just the schema.
 - I.e., it must be able to instantiate the objects described by the VODML in the VOTABLE.
2. Write VO service whose results include VODML.
3. Have a client demonstrate understanding of the serialized objects

Positives

- Only small changes to VOTABLE schema
 - GROUP, PARAM, PARAMref and FIELDref can have VODML child.
 - VODML can contain ROLE and/or TYPE.
- Easy to modify a parser to accept the new schema
 - Syntax only! Semantics much harder.
- Mapping document thoroughly explains the mapping cases

Challenges (1)

- Normative part of the document, while rigorous and mostly complete, is hard to read
 - Formal notation used requires a learning curve
 - The cases are complex (see next bullet)
- Semantics are context dependent
 - Whether VODML contains ROLE, TYPE, or both, depends on where it appears in a serialized object's structure.
 - Makes the mapping document descriptions complex
 - Makes it easy to create a VOTABLE that matches the schema, but is semantically invalid
 - Required many if/then statements in the parser

Challenges (2)

- Some confusion on how VODML types should be specified for PARAMs and FIELDrefs
 - PARAM and FIELDref values have VOTABLE primitive types
- Still unclear on the correct syntax for referencing objects defined per row in a TABLE.
- Multiplicity unclear unless more than one child of the same ROLE
- Need many more example VOTABLEs
 - The complexity of the mapping makes this slow going
 - Difficulty creating examples may indicate challenges for data providers

Nit Picks

- Objects (defined by GROUPs) know their ROLE within their containing object.
- PARAM must have “name” attribute, but that has no bearing on the VODML mapping and can be confusing.
- Scoping of instances is non-intuitive.
 - Any object instantiated outside a TABLE is global

VOTABLE/VODML Parser

- Runs as a web service described here:
 - <https://volute.g-vo.org/svn/trunk/projects/dm/vo-dml/mapping/MAST%20VODML%20Parser/VODML%20Parser.html>
- Side Goal: Recreate structure of serialized instances without needing to parse and understand the referenced models.
 - Encapsulate parser logic
 - E.g., parser could convert the VOTABLE to JSON for a client that has hard-coded knowledge of model(s)
 - Reduce network traffic and dependence during parsing

Alternative Syntax?

- Could some challenges could be addressed with an alternative syntax that adds several new elements to VOTABLE?
 - More explicit mapping with no overloading of existing VOTABLE concepts
 - Less context dependent
 - Simpler to describe and parse
 - Straw man schema:
 - https://volute.g-vo.org/svn/trunk/projects/dm/vo-dml/xsd/ext/VOTable-1.3_vodmlAlt.xsd
- If we were changing VOTABLE schema anyway, is adding many new elements any worse?
- Requires reworking mapping document
 - Much work was already invested in the first version

Recommended Path Forward

- Finish the experiment
 - Add many more examples and ensure parser handles all cases
 - Figure out in-table references
 - Refine alternative syntax for comparison
 - Finish a complete use case from data provider to client
 - Update mapping document
- Confirm that we want VODML in VOTABLEs
- Propose VOTABLE changes
 - We can assess general impacts now
 - How would it affect your parser?
 - How to ensure a smooth transition?
 - Need to coordinate with other VOTABLE changes?
 - COORDS
 - JSON
 - Removing unneeded elements

Discussion

- VODML Mapping
 - Questions
 - Concerns
- VOTABLE change process
- Alternative syntax
- Anything else