



Fig. 1

1. Towards Vocabularies 2

Markus Demleitner
msdemlei@ari.uni-heidelberg.de

- Requirements for Vocabularies 2
- What Vocabularies 1 said
- What I'd put into Vocabularies 2

[Update after talk: In case you're looking for a gentle introduction into RDF and its thinking, my predecessor as Semantics chair Norman Gray produced a nice piece: <http://eprints.gla.ac.uk/101484/>]

(cf. Fig. 1)

2. Requirements

- How to define common IVOA vocabularies (e.g., Datalink, Reposition). This is against Vocabularies 1's stance of letting everyone define their own vocabularies and mapping between them.
- Fix deployment rules (ivoa.net/rdf, content negotiation). This basically reflects age-old W3C best practices also reflected in Vocabularies 1.
- Recommend a small set of relationship properties (SKOS and RDFs). Vocabularies 1 only talked about SKOS. We want to be a bit more inclusive here because it turned out many of our vocabularies really define properties.
- Rules for adding terms, changing definitions. That's really the most pressing problem; there are terms for datalink/core that have been waiting for four years now. In contrast, I don't expect anyone will do actual semantics with what we're doing here any time soon.

Anything else?

3. Vocabularies 1.19

- Main use case mentioned: VOEvent <Why> and <What>. Does anyone know if Sem1 has been used for this? Is there perhaps still something like this going on?
- Focus on general-purpose vocabularies. This seems less of an application in today's VO, so I'd certainly tone this part down a lot.
- Limits itself to SKOS.
- Much on mapping terms from different vocabularies. That I'd like to drop entirely – it has simply not happened in the last 10 years, and if it happens, I think we'd want to have rules on how to discover such maps and then update the standard anyway.
- Recommends localised terms (e.g., Moon@en, Mond@de and Lune@fr). I'd like to drop that, too – does anyone see a realistic scenario where this might develop into anything but a trap?
- Suggests hash-style term URIs and date-based versioning.
- Sample vocabularies to A&A keywords, AVM keywords, UCDs and the IAU thesaurus. These are still around; see, for instance UCD in SKOS¹ (but note that subsequent UCD changes have not been reflected there).

4. Reminder: SKOS

SKOS is an RDF application for thesauri and similar “soft” knowledge organisation.

It organises `skos:Concepts` in *broader/narrower* hierarchies (there's also *related*).

These are loose properties; for instance, *House* is *broader* than *Door*.

(ok, transitive versions of these exist, too).

Use for “weak reasoning” (e.g., subject classification).

5. Meanwhile: Datalink

2014, Datalink needed a formal vocabulary of dataset fragment types:

“A dark frame *is* a piece of calibration data.”

The vocabulary should reflect IVOA consensus and be managed by the IVOA. This was the point when we started to leave the Voc1's concept of roll-your-own-and-map-later.

It's also talking about properties rather than `skos:Concepts`. RDFS is a better match here.

Norman Gray built a simple system for authoring such simple vocabularies. 2015: datalink/core² is REC.

¹ <http://www.ivoa.net/rdf/Vocabularies/UCD>

² <http://www.ivoa.net/rdf/datalink/core>

6. Meanwhile: More RDFS use

A derivative of Norman's vocabulary manager has been used for vocabularies of

- content levels,
- content types,
- relationship types,
- date roles,
- reference frames,
- reference positions,
- timescales.

See also <http://www.ivoa.net/rdf>³.

7. Meanwhile: UAT

Outside of the IVOA, the journal keywords and the IAU/IVOA thesaurus were developed into the Unified Astronomy Thesaurus UAT⁴.

That's SKOS territory, so the UAT's primary format is SKOS.

8. Meanwhile: Theory

The theory IG has built SKOS vocabularies for

- Object types (based on CDS work; it's different from Simbad's otypes, though)
- Physical processes and quantities
- Datatypes in simulations
- "Algorithms" (rather: simulation methods)

They followed Vocabularies 1, and their subject matters mostly are more SKOSable, so they're all in SKOS.

9. Lessons learned

- Non-IVOA folks are doing "the" Astronomy Thesaurus
- For most purposes, we'd like "strong", consensus vocabularies
- There are places of SKOS, too
- Nobody wants to map between vocabularies

So: Vocabularies in the VO needs an update.

³ <http://www.ivoa.net/rdf>

⁴ <http://astrothesaurus.org/>

10. Voc2: Properties

- In SKOS: skos:Concept and skos:narrower.
- For properties ("predicates"): rdfs:Property and rdfs:subPropertyOf
- For thing-like terms: rdfs:Class and rdfs:subClassOf
- Use owl:equivalentProperty and owl:DeprecatedProperty for properties we no longer want.
Todo: see how this extends to rdfs:Class and skos:Concept..

11. Properties vs. Classes

Properties are what we need for datalink and friends, essentially:

(URL, is-calibration-file-for, dataset)

For spatial reference frames, timescales and the like we'd much rather have:

(coordinates, in-frame, ICRS)

This means that we're talking about objects (or subjects) of RDF triples, not the predicates we have in datalink. That's why we want, in addition to Norman's work and SKOS, vocabularies defining rdfs:Classes-es and accordingly using rdfs:subClassOf.

Plan: Don't mix vocabularies for classes and properties. Also note that what's in [ivoa.net/rdf](http://www.ivoa.net/rdf) uses properties throughout at this point. That's a bug I'll fix soon.

12. Voc2: Deployment

As (essentially) in Voc1:

- Vocabulary URI: [http://www.ivoa.net/rdf/\(vocname\)](http://www.ivoa.net/rdf/(vocname)).
- Redirects to .../YYYY-mm-dd/(vocname).(something).
- (something) is determined by HTTP content negotiation (via the accept header).
- Have at least HTML and RDF/X (SKOS counts), perhaps turtle.

13. Voc2: Updates 1

To add a term (or change its meaning), put up a VEP on the wiki:

Vocabulary: <http://www.ivoa.net/rdf/datalink/core>

Author: msdemlei@ari.uni-heidelberg.de

New Term: isMetadataFor (child of auxiliary)

Label: Metadata

Description: additional documentation for this dataset

(e.g., observatory logs, provenance information)

Rationale: This would annotate material such as observation

[...]

This then essentially follows the errata process.

14. Voc2: Updates 2

VEPs are announced on semantics and concerned WG mailing lists. Discussion should take place on the WG and TCG mailing lists.

Meanwhile, the term is added to published vocabulary as a candidate. Adding a candidate would *not* change a vocabulary version.

For each TCG meeting and each open VEP, the semantics chair produces a discussion summary and a proposed decision.

The TCG (in consensus) has the final word. When terms are added, the vocabulary version is updated to the date of the addition.

I plan to also strongly recommend to maintain all vocabularies within a single repository (currently, this would still be *volute*).

15. Voc2: Timeline

I'll start working on Voc2 after the interop.

First WD: August

First VEPs plus discussions in September

First TCG involvement by Groningen

PR after the first terms have passed the review?

Any help and feedback is highly appreciated!