1. A Vocabulary of Object Types

Markus Demleitner with Anaïs Oberto,
Cécile Loup, Sébastien Derriere
and the Simbad Crew
msdemlei@ari.uni-heidelberg.de

- What's it for?
- Building principles
- Usage examples
- Next steps

(cf. Fig. 1)

2. What's it for?

Originally, Simbad have had an object classification since... forever. See `otype` and `otype_txt` in `basic` as well as the `otypedef` and `otypes` tables.

Later, standards like SSAP (the `ssa:Target.Class` output field) and Obscore (the `target_class` column) needed object types, too.

Obscore: "It is a string with possible values defined in a special vocabulary set to be defined: list of object classes (or types) used by the SIMBAD database, NED or defined in another IVOA vocabulary."

So: how do I look for a Radio-loud quasar?

3. A Case for RDF Classes

If I'm looking for, say, evolved stars, I'd like to find white dwarfs as well as AGBs – but when I'm looking for AGBs, I don't want to see white dwarfs.

A classic for a taxonomy. In our terms: Should this be an RDF class vocabulary?

As usual, trouble: E.g., cataclysmic stars are both variable and multiple.

This is a problem since in our RDF class vocabularies, a concept can only have one parent concept. Whether or not this is a problem that persists in actual usage remains to be seen.

Alternative: In SKOS vocabularies, cataclysmic stars can be narrower than both variable and multiple stars. But there, relationships are not transitive (i.e., just because X is a variable star you cannot conclude that it is a star), which is a hindrance in many interesting applications.

Fig. 2

4. Building the IVOA Mapping

Largely re-using Simbad's hierarchy is an obvious choice.

Terms there are identified using quite a bit of punctuation (`Ae*`, `Radio(cm)`), which is trouble in our URIs. We hence rewrite them using a few rules:

- CamelCase → words-with-dashes (this is because mixed case is trouble in many contexts).
- * → star
- parentheses, slashes and underscores become dashes (this helps keeping things unescaped and predictable)
- plus a handful of custom rules, e.g., ** → multiple-star
- lowercase what's left.

These rules do not always lead to ideal identifiers, because astronomical nomenclature is not case-insensitive (e.g., Ae as a spectral class vs. AM as a variable star name). However, it's only identifiers. Humans should in general see the case-preseving labels, so there shouldn't be any confusion.

Result: http://www.g-vo.org/rdf/object-type[1]

5. HTML rendering

(cf. Fig. 2)

This is a small piece of the vocabulary. In the HTML rendering, you see the concept identifier, the label (i.e., short, human-oriented name), the definition (which could be a bit less terse here), possibly a parent concept, and relationships. #assoc-star has its narrower terms folded out (#stream-star and #mouv-group), #agn has its synonyms folded out. One of those points to the original Simbad concept, the other into the UAT.

---

[1] http://www.g-vo.org/rdf/object-type

6. In Desise

In addition to standard RDF, you can retrieve the vocabulary in the IVOA-specific Desise:

```
"agn": {
  "label": "AGN",
  "description": "Active Galaxy Nucleus",
  "preliminary": "",
  "wider": ["galaxy"],
  "narrower": [ "seyfert", "radio-g", "liner", "qso", "seyfert-1",
    "seyfert-2", "blazar", "bl-lac", "bl-lac" ]
},
"seyfert": {
  "label": "Seyfert",
  "description": "Seyfert Galaxy",
  "preliminary": "",
  "wider": [ "agn" ],
  "narrower": [ "seyfert-1", "seyfert-2" ]
},
```

This is designed for maximally painless consumption; nobody should need to show identifiers because they don't want to pull in an RDF library, and nobody should forego exploiting hierarchies because it seems algorithmically complex.

7. Intended Usage

Suppose you look for spectra of young stellar objects in the vicinity of M42:

```
select obs_title, target_name, target_class, access_url
from ivoa.obscore
where
  1=gavo_vocmatch('object-type', 'yso', target_class)
  and distance(point(s_ra, s_dec), point(83, -5))<5
  and dataproduct_type='spectrum'
```

This will also return rows for T Tauri stars, Herbig Ae stars, Outflows, Orion variables, and whatever else object-type declares as narrower than YSO.

You can run the query on the TAP service at https://dc.g-vo.org/tap.

Full disclosure: This doesn't return anything at the moment, partly because I don't think I have spectra anyone has identified as being of such an object in any way, but also because I've not yet ported the few target class annotations I have to the object-type vocabulary. But this is what object-type should enable in the end.

8. Next Steps

- Everyone: Review! Is your favourite object class missing?
- Perhaps extend definitions? As people outside of the CDS want to use the vocabulary for annotation, it may be necessary to be a bit more specific on what is and is not a, say, evolved star.
- Figure out whether our use cases are better served by a deep tree or a flat-ish SKOS structure.
- Ensure there are no gross contradictions with the UAT.
- Perhaps briefly think about annotating "candidates"?

Help wanted!

Thanks!