1. The UAT in the VO
Markus Demleitner
msdemlei@ari.uni-heidelberg.de

Reminders:
Virtual Observatory (VO) = a system of conventions and protocols to let machines work with astronomical data.
VO Registry: A set of metadata making that data discoverable for machines (humans can get an idea on https://dc.g-vo.org/WIRR).

(cf. Fig. 1)

2. Vocabularies in the VO
The VO needs a lot of hierarchically organised word lists: time scales, dataset component types, relationships...
We’re keeping them in negotiated RDF/XML, turtle, HTML and desise at http://www.ivoa.net/rdf.
Desise is a custom json serialisation that lets VO clients trivially consume the vocabularies.
The conventions governing this: Vocabularies in the VO 2 (currently under review).

3. The UAT and the Registry
When registering services, operators must give subject keywords.
VOResource 1.1 (2018) says: “Terms for Subject should be drawn from the Unified Astronomy Thesaurus (http://astrothesaurus.org).”
But: What does this mean?

4. Adoption Problems
VO Components have been relying on the subject element containing something human-readable.
Hence, we’d break a lot if we told people to put in UAT URLs.
Use preferred labels instead? But they’re designed to be changeable.
Also: We’d like to give our clients the UAT in desise anyway.
Solution: have an “IVOA mirror” of the UAT: http://www.ivoa.net/rdf/urat.

5. IVOA Mirror?
• Concept URIs are ...#preferred-label-at-first-map
• Stable, machine-readable mapping IVOA URI ↔ UAT URI
• Programmatic mapper, in principle executable as a github action.
• Details in Adopting the UAT as an IVOA vocabulary2.

6. Experiences: Mapping One Site
My data center3 has about 500 subject keywords; I’ve migrated them to the UAT (the "by subjects" tab, or see the Registry).
Experiences:
• Infrastructure services (e.g., IVOID validator) don’t really fit. VO Supplement? But for what? Meaning: Are there discovery use cases that would profit from VO-specific subject keywords?
• A few minor points on individual terms we might be missing (“multi-messenger”, “deals-with-cubes”).
• For all I can see: ready to go.

---

1 https://ivoa.net/documents/Vocabularies/20200612/index.html
2 https://ivoa.net/documents/uat-as-upstream/20201117/3 http://dc.g-vo.org
7. Experiences: Mapping the VO

I've also mapped all terms I found in the VO Registry. In summer, there were 1010 of these.

- 323 "unfixable" (i.e., subject keyword abuse or bad syntax)
- 54 keywords that might want representation in the UAT (in fewer concepts; mapped to ivoa:TryAgain)
- 633 plausible mappings
- Strongly non-Zipfian frequency-of-frequencies distribution suggests this is missing many specialised subjects. That's another talk I may give some other day.

Please review: mapping file

8. Sembarebro

Based on this mapping work, I've built the Semantics Based Registry Browser SemBaReBro. Also note how compact device-based vocabulary operations are (js source).

(cf. Fig. 2)

9. Lecture Notes With Links

http://docs.g-vo.org/ru.pdf (for a while)

---

4 http://svn.ari.uni-heidelberg.de/svn/gavo/hdinputs/sembarebro/res/mapping.tsv
5 http://dc.g-vo.org/sembarebro/q/ai/fixed
6 http://svn.ari.uni-heidelberg.de/svn/gavo/hdinputs/sembarebro/res/ai.shtml