Space-Time Coverage in the Virtual Observatory Registry

Markus Demleitner

Universität Heidelberg, Astronomisches Rechen-Institut, Mönchhofstraße 12-14, Germany

msdemlei@ari.uni-heidelberg.de

Abstract

Probably the most severe deficiency of the VO Registry is that so far it has been impossible to write discovery queries in terms of time, space, and spectrum.

In 2018, a roadmap for remedying this deficiency [1] was published. The first steps on the roadmap are now falling into place: VODataService 1.2 [2] gives a reasonably straightforward way to declare a resource's coverage, and a small, so far nonstandard, extension to RegTAP already lets VO clients use the information available for discovery.

Takeup of VODataService 1.2 on the side of the data providers, however, so far has been lackluster. This poster is an attempt to change that.

Declaring Space-Time Coverage

To declare where data within a resource is located in the VO phase space, VODataService 1.2 defines three new child elements of the pre-existing `coverage` element:

- `spatial` – zero or one allowed. Contains an ASCII MOC, written in the ICRS.
- `temporal` – zero or more allowed. Contains a space-separated pair of floats giving an MJD interval.
- `spectral` – zero or more allowed. Contains a space-separated pair of floats giving an energy interval (so it works nicely for non-EM messengers; it is given in Joule so we do not favour a specific electromagnic waveband).

While it should not matter much for discovery, everything should be given for the solar system barycenter, and times should use TDB.

An Example

```
<coverage>
  <spatial>3/282,410
    4/40,323,326,329,332,387,390,396,648-650,1083,1085,1087,
    1101-1103,1123,1125,1132-1134,1136,1138-1139,1144,1146-1147,
    1173-1175,1216-1217,1220,1223,1229,1231,1235-1236,1238,1240,
    1597,1599,1614,1634,1636,1728,1730,1737,1739-1740,1765-1766,
    1784,1786,2803,2807,2809,2812</spatial>
  <temporal>37190 37250</temporal>
  <temporal>38776 38802</temporal>
  <temporal>41022 41107</temporal>
  <temporal>41387 41409</temporal>
  <temporal>41936 41979</temporal>
  <temporal>43416 43454</temporal>
  <spectral>3.01e-19 6.02e-19</spectral>
  <waveband>Optical</waveband>
</coverage>
```
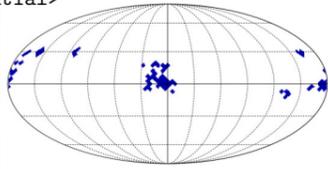
By the way, this is the actual coverage of the Palomar-Leiden Trojan surveys. We admittedly could get a bit more precise on the spatial side. The recommended MOC maximum order for the Registry is 6.

Reality

Unfortunately, not many registry resources declare their STC coverage yet. The tally is (as of 2019-09-23):

- 15131 spatial, but over 90% of these are actually havested from their footprint URLs (which is a pain and does not scale).
- 80 temporal (from 5 authorities)
- 75 spectral (from 4 authorities)

If you are publishing VO resources: **It is up to you to improve this!**

Also: pgsphere currently only has POINT vs. MOC operations, so you are currently limited to points (rather than regions) of interest, and there is no support for indexing MOCs yet, so queries will scale with $O(n)$ as the Registry grows. We are working on fixing both problems.

The Discovery Side

To enable discovery on these coverages, the Roadmap for STC in the Registry [1] suggested an extension to RegTAP that, adapted for the change of spectral modelling, is already implemented on the RegTAP service at http://dc.g-vo.org/tap. Three tables are added:

- `rr.stc_spatial` with columns `ivoid`, `coverage`, and `ref_system_name`; coverage is a MOC-typed geometry which so far only supports CONTAINS. `ref_system_name` is NULL for ICRS. Non-celestial data will have non-empty values in that column.
- `rr.stc_temporal` with columns `ivoid`, `time_start`, and `time_end` (with MJD values).
- `rr.stc_spectral` with columns `ivoid`, `spectral_start`, and `spectral_end` (in Joules of energy).

With this, you could look for services carrying data for messengers around the rest mass of electrons for the center of the LMC like this:

```
SELECT ivoid
FROM rr.stc_spatial
NATURAL JOIN rr.stc_temporal
WHERE
  1=CONTAINS(POINT(80.9, -69.8), coverage)
  AND ref_system_name IS NULL
  AND ivo_interval_overlaps(
    in_unit(spectral_start, 'keV'), in_unit(spectral_end, 'keV'), 500, 550)
```

One slightly silly thing you can do with this data is a density map of VO resources; with a DaCHS-specific extension of ADQL that generates integer series, here is how you can do this (12287 is the largest HEALPixes index on level 5):

```
SELECT hpx, COUNT(*)
FROM generate_series(0, 12287) AS hpx
JOIN rr.stc_spatial
ON (1=CONTAINS(ivo_healpix_center(5, hpx), coverage))
GROUP BY hpx
```

The result currently looks like this in equatorial Aitoff projection (can you spot SDSS and Kepler?):
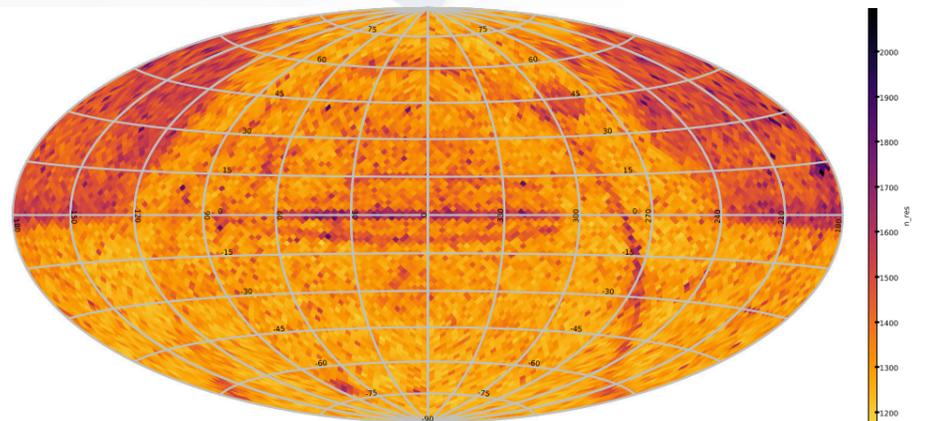
Extending the Scheme

- Solar system data: Right now, coverage is always in ICRS. That should be ok for celestial data, but solar system resources need something else. We hope `ref_system_name` will be sufficient.
- Redshift/Distance: STC1 offered redshift as another axis, and indeed one could make a point that distance is part of the spatial location. Various practical considerations made us drop the axis for now.
- Non-EM messengers: There is no way to look for "Neutrino" or "Gravitational Wave" by this scheme yet (but you *can* tell GeV-neutrinos from eV-neutrinos). Perhaps subject keywords are enough? Or perhaps expand the waveband vocabulary?

[1] Markus Demleitner. A roadmap for space-time discovery in the VO registry. IVOA Note, 8 February 2018, January 2018.

[2] Raymond Plante, Markus Demleitner, Aurélien Stébé, Kevin Benson, Patrick Dowler, Matthew Graham, Gretchen Greene, Paul Harrison, Gerard Lemson, Tony Linde, and Guy Rixon. VODataService: A VOResource schema extension for describing collections and services version 1.2. IVOA Proposed Recommendation, 15 July 2019, July 2019.