IVOA April 2022 Interoperability Meeting - DAL Session 1

Time: Tuesday April 26 20:30 UTC
Participants: (52) Marco Molinaro (MM), Robert Butora (RB), François Bonnarel (FB), Chloé Azria (CA) Baptiste Cecconi (BC), Markus Demleitner (MD), Pat Dowler (PD), Gregory Dubois-Felsmann (GDF), James Dempsey (JD), Grégory Mantelet (GM), Tim Jeness (TJ) et al.

Schedule

Marco Molinaro, Robert Butora - **VLKB VO plans to multi-cutout**
François Bonnarel - **SIA2/SODA upgrades and Extending SIA2 into a generic DataSetSAP**
C. Azria, B. Cecconi - **Updated heliophysics services in VESPA: science products, service design and capabilities**

Notes

VLKB VO plans to multi-cutout (*Marco Molinaro, Robert Butora)*

- VLKB = ViaLactea Knowledge Base
  - dedicated to milkyway astrophysics research
  - content: catalogues of sources, SEDs, distances
  - services: TAP and some custom services
- It has its own analysis tool: ViaLactera Visual Analytics. It can analyse several types of data (e.g. cubes, tables)
- Several VLKB datasets services: custom services to search, merge, cutout, ...
- On the VO view
  - search can be done with Obscore (already implemented but not available), SIAP-2.0 and ObsTAP
  - cutout with SODA interface
  - multi-cutout possible by combining SODA and UWS
- SODA does not provide a way to bind an ID to multiple parameters but would be possible with Datalink (see SODA-1.0 sec. 3.1)
- Current implementation uses a single UWS Job and and JSON array in input

Markus Demleitner: How does CADC do it (DaCHS refuses multiple ids and merges cutout regions)?

Pat Dowler: in soda, we reject multiple IDs; in soda-sync we reject multiple cutout params of the same type, while in soda-async we allow them and produce combination results

  MD: Use a combination of DALI-Uploads + require multiple identifiers

  Gregory Dubois-Felsmann: +1 on DALI UPLOAD

  Marco Molinaro: not always trivial to represent this in VOTable

  PD: Would need to have a mapping from the inputs to the result names

MD: The tricky part will be when you do sync…

  MM: yes, but we plan async only for this… up for discussion anyway

JD: Have you thought to name your results to associate a result with IDs?

RB: currently replicating the query row values in the response.

GDF: I agree with Marco's concerns about what current standards support for multi-cutouts. Both Rubin and SPHEREx are planning multi-cutout services for which one of the options will be to return a (potentially very large) MEF combining all the results. We'd be interested in figuring out how to index those back to the request.

SIA2/SODA upgrades and Extending SIA2 into a generic DataSetSAP *(François Bonnarel)*

- SIAP2:
  - Several PRs published
  - errata: POS=RANGE example inconsistency with spec
  - parameters: proposal to enforce the need to have self-describing parameters
  - One shot cutout from SIA2 - experiment at CDS to return a SODA URL as the access URL
  - new features planned: full metadata endpoint ; cube-dm perspective
  - current limitations to image and cube ; why not extending that to timeseries and spectra
    - with spectra -> SSA2 (obscore compatible)
    - catalogue/measurement more complex as there is no dataproduct_type (but requested by ASKAP)
    - Regarding this extension, proposal to rename into : Dataset Simple Access Protocol (DsSAP)

    •   - PD: Simple Data Access == SDA?

GDF: Rubin needs a protocol for discovering spatially sharded catalogs. We would be happy to use an SIAv2 derivative for this, but we have some concerns that the existing s_region may not be adequate to exactly document the shard boundaries.

James Dempsey: CASDA/ASKAP support the extension as we need a way to request visibilities, spectra and catalogs

MM: what happens in case of discovery research with positional parameters with a partial spatial coverage? Is the request rejected?

  François Bonnarel: don't think this is supported (intersection, partially/fully overlap, ..)

PD/FB: Will have a gathertown discussion on Friday about PRs

Updated heliophysics services in VESPA: science products, service design and capabilities *(C. Azria, B. Cecconi)*

- VESPA = Virtual European Solar and Planetary Access
  - Include a variety of data sets: images, spectra, events, cubes, time-series, ,,,,
  - Uses EPN-TAP (TAP + EPNcore dictionnary)
- Update of the services includes addition of many various data + move to DACHS-2.5
- New features:
  - new metadata ingestion method
  - datalink for progenitor, quicklooks, doc, data access API
  - Access API for time-series and dynamic spectra
  - catalogues of spectral-temporal features

- Technical aspects:
   - custom grammar extending EPN-TAP (NHS/OFREES) using ElasticSearch
   - odbcGrammar
   - s_region active region represented by chaincode (sunpy.net.helio.chaincode) and Shapely to deal with simple and multi-polygons

MD: limitations with polygons come from pgsphere, not from DACHS. Status about polygons in DALI?
PD: Have added multi-polygon to DALi 1.2 and have WD up. Have raised issue for VOTable to store arrays. Pgsphere has limitations but can work around for 99% of cases

Baptiste Cecconi: possibility to use PostGIS instead of PgSphere?
  MD: personal believe is with MOC, no need to support multi-polygone. We should support MOCs (much easier to work with than polygon).

  Tim Jenness: I asked on Slack, but is pgsphere alive?
  PD: yes, it's sort of alive
  TJ: We're on Google and we can't get them to add pgsphere into CloudSQL because it seems dead and they tell us to use PostGIS instead which they do support
  MD: Debian packaged so should be easy to use
  MD: Not highly active but working with Debian pro who are ensure it works with current versions. have a few PRs open.