

European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

ESCAPE Extended Discussion Day

WP4 + WP5 Discussion

28 September 2021



ESCAPE Goals for this session

- Refine our understanding of how ESAP should query data through the VO. Including, for example:
 - How does ESAP interact with the VO registry?
 - How does ESAP forward results from VO services to analysis systems?
 - What forms of VO query should ESAP support?
 - ...?
- Identify other VO technologies that could be used within ESAP. For example:
 - VO tools for visualization & analysis within notebooks or other IDA systems.
 - Dispatching compute jobs (UWS? others?).
 - Provenance.
 - ...?

(NB: we have ~40 minutes)





- This is a **discussion** session; we want to hear from you!
- As far as possible, we encourage you to engage on Zoom let's discuss your needs and ideas, rather than using it as just a Q&A.
- However, we will keep an eye on #1-general-discussion on Slack, so feel free to add thoughts there.







• <u>Slides by Yan & Hendrik</u>.

Funded by the European Union's Horizon 2020 - Grant N° 824064





• Are we properly making use of the VO registry, rather than simply hard-coding the use of particular services?





- What is the data flow for transferring query results from ESAP to analysis systems? Do we send the *query*, which is replayed on the analysis system, or the *result of the query*?
 - The former means the query needs to be run twice; the latter that the we have to store and forward data through ESAP, including preserving metadata etc.
 - Is it possible to subset queries VO? A standard ESAP workflow involves running a query then selecting a few rows from the output for further analysis; is that possible with a "send the query" model?
- Stelios suggests: async queries, the results of which can be retrieved multiple times.







- Do we need to support VO queries in ESAP at all?
 - If we have SAMP, can folks use external tools to run their query, then simply upload results to ESAP?
 - One imagines that tools like TOPCAT, pyvo, etc provide a more refined VO query experience than we're likely to achieve in a relatively short time with ESAP.
 - Does this preserve all necessary metadata?
 - What about provenance does SAMP let us track where the data came from originally?
 - Are there worries about data volumes?
 - Are there tools *other* than SAMP that would be relevant here?

Funded by the European Union's Horizon 2020 - Grant N° 824064





- Are there opportunities to use VO technologies in ESAP for particle physics applications? They deal with tables, etc, too...
 - "TAP for particle physics services"







- Is this simply an implementation issue for WP5? Are there WP4 participants who can get involved?
- Are there tools, code, other artefacts from WP4 that are relevant?
- Can we define some concrete goals?



