Scope definition

- Query schema of VO registry
- Steps of VO registry query

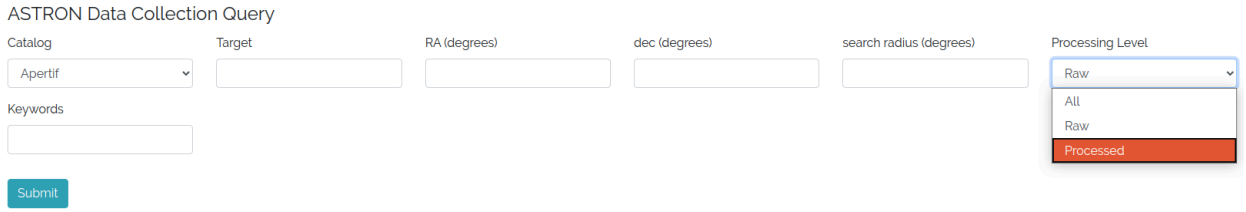
Example Query Schema

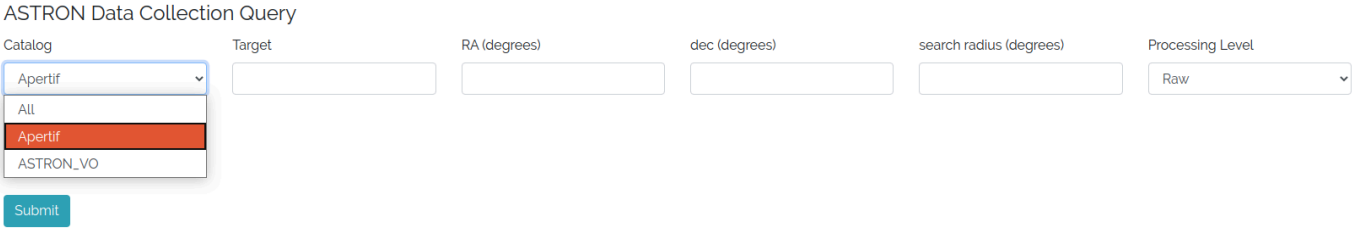
Please suggest editing the example query schema below:

```
# definition of the query fields and values
query_schema = {
        "name": "adex",
        "title": "ASTRON Data Collection Query",
        "type": "object",
        "properties": {
                "catalog": {
                "type": "string",
                "title": "Catalog",
                "default": "apertif",
                "enum": ["all","apertif", "astron_vo"],
                "enumNames": ["All","Apertif", "ASTRON_VO"]
                },
                "target": {
                "type": "string",
                "title": "Target"
                },
                "ra": {
                "type": "number",
                "title": "RA (degrees)",
                },
                "dec": {
                "type": "number",
                "title": "dec (degrees)",
                },
                "fov": {
                "type": "number",
                "title": "search radius (degrees)",
                },
                "level": {
                "type": "string",
                "title": "DataProduct Level",
                "default": "raw",
                "enum": ["all","raw","processed"],
```

```
                "enumNames": ["All","Raw","Processed"]
                },
                "category": {
                "type": "string",
                "title": "Keywords",
                },
        }
}
```

Rendering this query schema gives the query form that looks as follows:





VO Registry Query steps

The following describes some use cases that would use the VO Registry to discover data and data services.

The VO Registry allows you to query for certain for certain service types, including but not limited to: *sia, ssa, scs, tap*

- **TAP**: Tables
- **SCS**: Cone Search
- **SSA**: Spectra
- **SIA**: Images

Use Case Scenarios:

Scenario 1:

User arrives at the search page. They are given a list of service types to select from, including catalog data, images, spectra, cone search services, all of which correspond to equivalent VO Protocols (listed above). Along the service types there is a textfield/textarea for searching based on a keyword.

Based on the type of service they have selected, there is an optional set of parameters displayed, which are checkbox options corresponding to the keyword search.

For example:
If the catalog data option is selected (TAP), users can select whether the keyword should match: *Table Names, Table Descriptions, Service Name.*

There are also a number of checkbox options that are default and shown for all service types: *Short Name, Title, Subjects, ID, Publisher & Description*

Similarly, all these options define whether the keyword search should return results that match these parameters to the keyword.
An optional parameter that can be used in this search is a *waveband* (i.e. waveband = "x-ray").

Aside from the above, an option for a user to manually enter a service URL can be provided, for users who are already aware of the service they want to access.

Once the user has selected the options and entered a keyword to search for, they can hit the "Search" option, which will return a list of services that correspond to that search.
The result list should at the very least contain a *short name,* a *title,* the *subjects* (short description of service), and an *access URL*.

The user can select the service they want to access, and launch a second (query) page, where they can run a query to access data from that service.

The query page would most likely depend on the service type defined above.

For a TAP service a user can:

- Display list of schemas/tables & columns for the selected service
- Create SQL query manually

- Generate SQL query through form (i.e. select [cols as checkboxes] from [table name as checkboxes] where [where criteria as text?] and ...)

For a Cone Search access service:

- Provide text field options for defining *RA, DEC, Object name & Radius*

For an Image access service:

- Provide text field options for defining *RA, DEC, Object name, Angular Size & Image format*
- Provide Image format dropdown with the following options: *image/fits, GRAPHIC, ALL*

For either of these options, a query should probably display a new page (stateless, i.e. can copy paste URL and get the same results) with a list of rows in the case of tables, or a list of images in the case of an image search. From here on they can either be presented with a data-cart functionality for exporting the data, or an interactive way for navigating through the results, and saving the results locally etc..

Scenario 2:

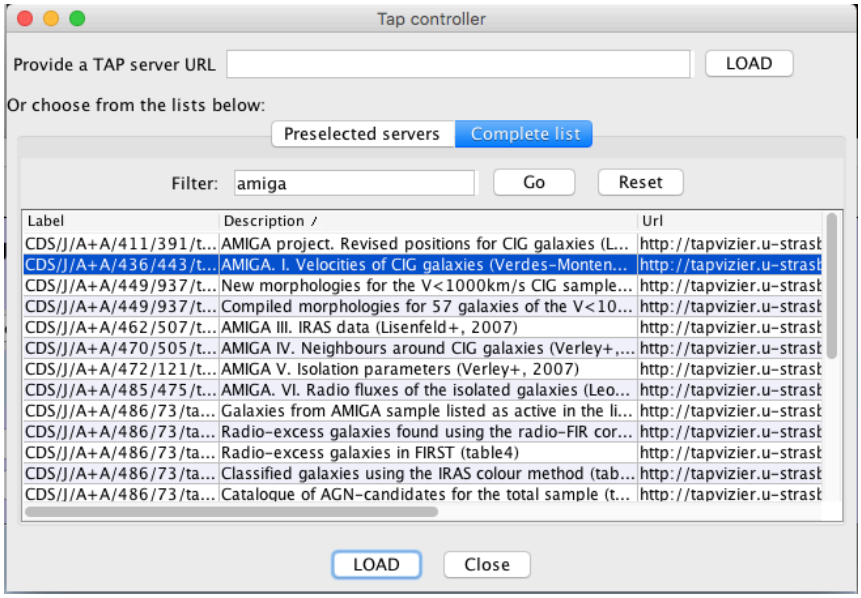
User arrives at the search page. Using e.g. a TAP query they discover a set of objects of interest. It is likely that they would then like to search for and retrieve any imaging and spectral data that are available for the objects that they are interested in. If it does not “come for free” as part of the VO protocol, then it might be a nice value-added feature of the ESAP VO query interface to enable query chaining or joins between related queries for different data related to the same objects.

Concretely this might involve extracting RA/Dec coordinates and object angular extensions from the results of a TAP query and using these to execute image access queries for the corresponding objects.

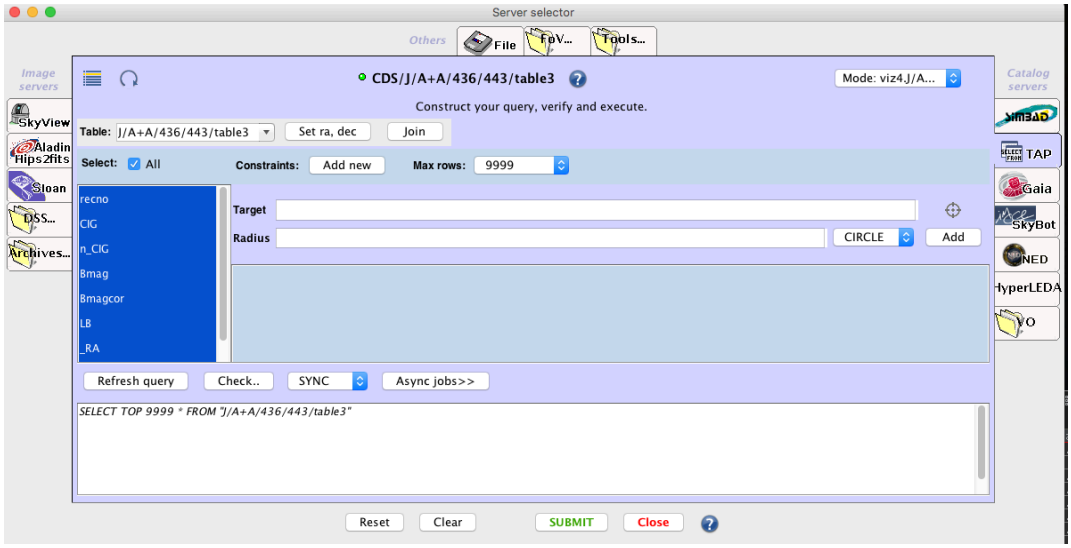
Scenario 3: connecting ESAP to other VO tools through SAMP

SAMP Use Case 1: TAP services and Aladin

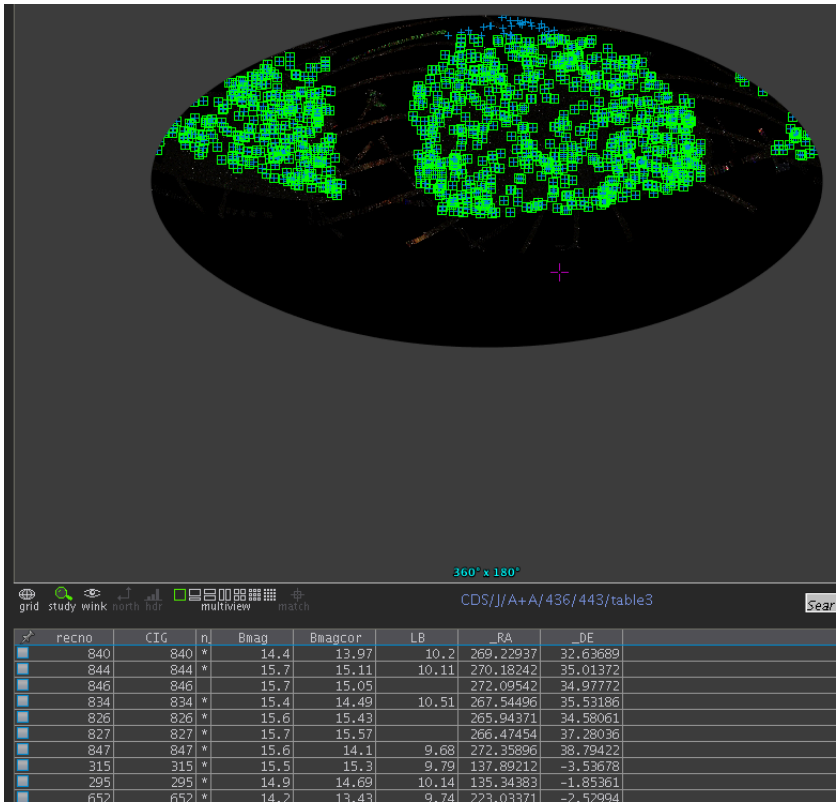
- An user opens the server selector window in Aladin, and using this GUI, he/she filters the complete list of TAP services using some keywords. Then select (LOAD button) the TAP service that he/she needs.



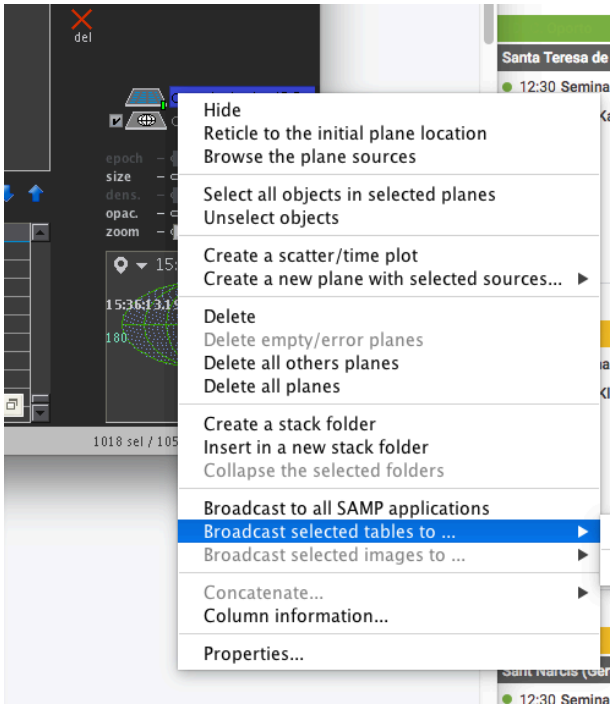
- Then the user creates a TAP query using the ALADIN GUI.



- TAP query results are loaded in ALADIN



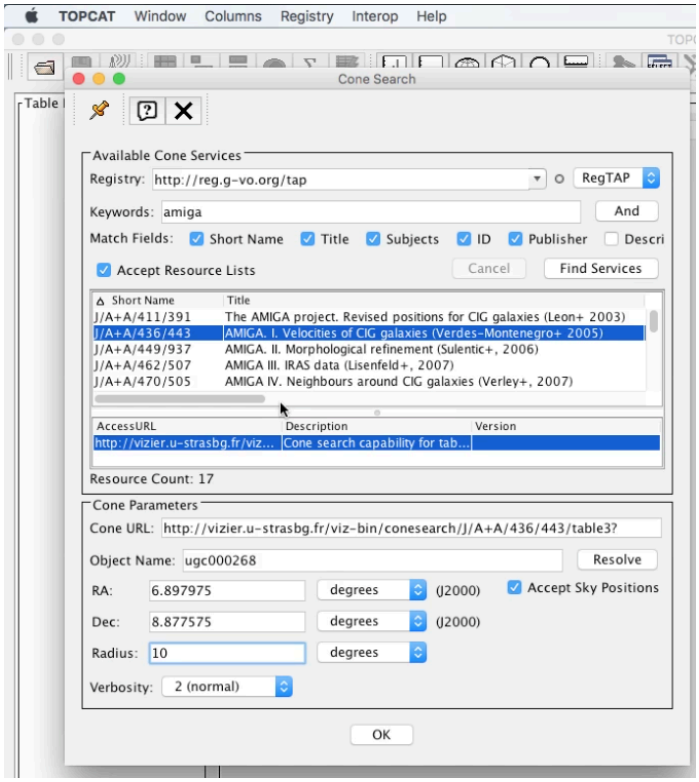
- Once the data results is loaded in ALADIN the user can e.g. load the SLOAN DR8 to better identify the data he/she needs
- Once she/he has selected the required data, select the ALADIN option "Broadcast selected tables to" > ESAP platform.



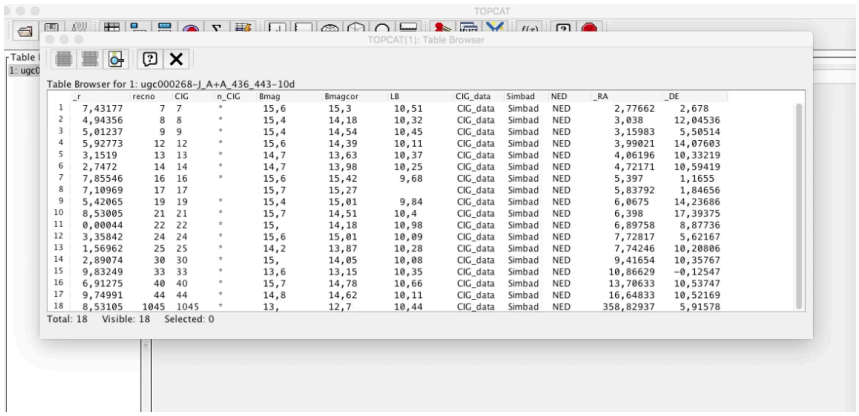
- The VO Table resulting from the TAP query is sent to ESAP via SAMP protocol
- The VO Table is shown in the ESAP Front End and then the user presses a button to start the staging process to transfer the data from the catalogs to the data lake.

SAMP Use Case 2: Cone Search services and Topcat

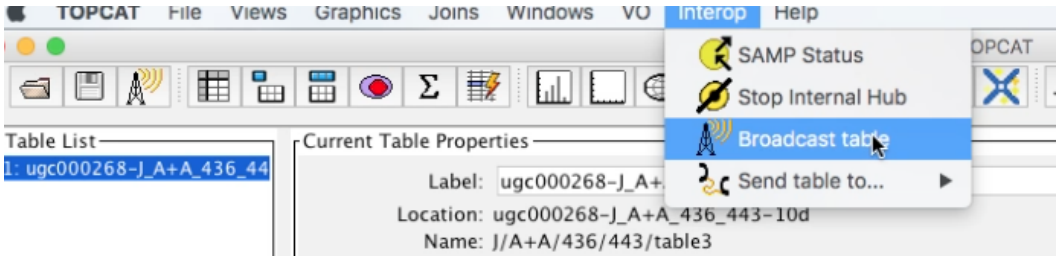
- The user opens TOPCAT and from the ConeSearch GUI search for a coneSearch service in the VO Registry. Then he/she provides the input values for RA, DEC and Radius



- The VOTable resulting from the ConeSearch query is loaded in TOPCAT



- Then, the user sends this table to the ESAP Front End, by the TOCAT functionality “Broadcast table” (or “Send table to”)



-

About SAMP and HTTPS

Although SAMP's Web Profile works well with HTTP-based web applications, it cannot be made to work for web applications hosted on servers using the HTTPS protocol.

This problem is fully explained here:

- https://wiki.ivoa.net/twiki/bin/view/IVOA/WebSampHttps
- https://arxiv.org/pdf/1912.00917.pdf
- https://wiki.ivoa.net/internal/IVOA/InterOpOct2019Apps/tlsamp.pdf
- https://github.com/astrojs/sampjs

Obscore Search

Several libraries including PyVO allow you to find TAP services that support the Obscore data model.

Obscore lets people publish observational datasets through TAP tables.
For example with PyVO this can be done using:
```
vo.regsearch(datamodel="obscore")
```

Obscore uses generic metadata so that any Obscore service can be used in the same way.
Here is an example use case that uses the Obscore protocol:

Show me a list of all data which satisfies:

I. DataType=TimeSeries
II. RA includes 16.00 hours
III. DEC includes +41.00
IV. Time resolution better than 1 minute
V. Time interval (start of series to end of series) > 1 week
VI. Observation data before June 10, 2008
VII. Observation data after June 10, 2007

Query

```
SELECT  TOP 100  ivoa.ObsCore.access_url,ivoa.ObsCore.obs_id
FROM ivoa.ObsCore
WHERE          (CONTAINS(POINT('ICRS', ivoa.ObsCore.s_ra, ivoa.ObsCore.s_dec),
CIRCLE('ICRS', 10.684667, +41.268750, 0.016666666666666666)) = 1)
          AND ivoa.ObsCore.dataproduct_type = 'timeseries'
          AND ivoa.ObsCore.t_resolution > 60
```

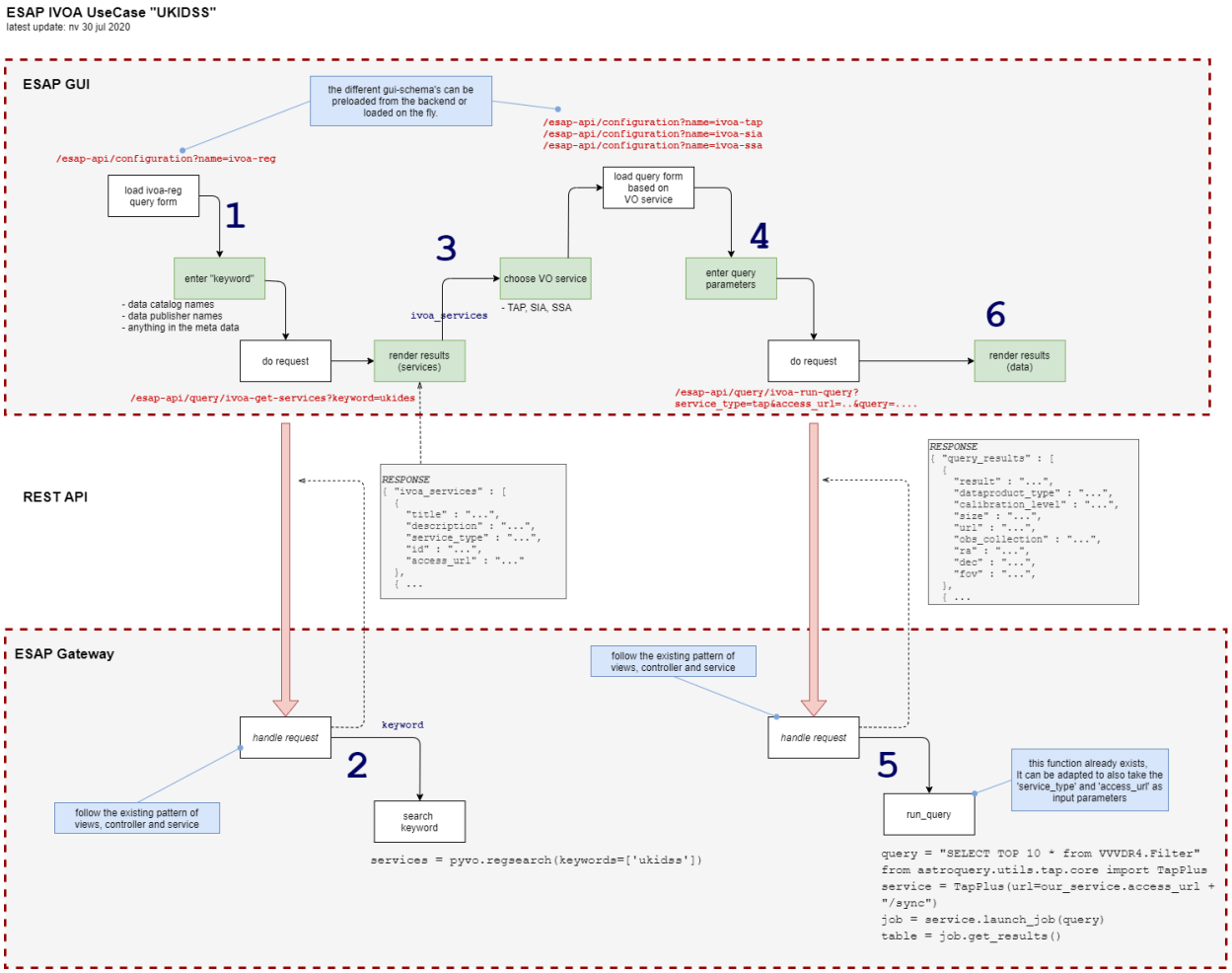
AND ivoa.ObsCore.t_min BETWEEN 54261 AND 54627

An additional user data flow diagram could be used for the Obsore protocol, where a search returns a list of services that support the Obscore protocol, and then the user provides a set of parameters which generate a query on every service in the list. The user can then select from the result what they want to checkout.

more...

User Data Flow & UKIDSS VO Example

NV: This is a model and design for the current implementation of ESAP based on what Stelios wrote. There is one main difference; the description had not taken into account that the workflow is handled by the GUI, the business logic by the ESAP Gateway and that all communication is through a REST API.



GUI page 1: Enter Keyword (input)

Users arrive at the Search page, where they are given a "keyword" textfield (Google-like), where they type a keyword to search for. This keyword may refer to Data Catalog names, Data Publisher names, or anything else in the registered description or metadata of the service.

Once they click the search button, the ESAP platform would call a service which could use PyVO to search for the given keyword:

```
services = pyvo.regsearch(keywords=['ukidss'])
```

There is an alternative to this simple keyword based search, which is to allow users to create a more advanced search, where they are also able to define the type of the service that they want to search for. This could be in the form of a dropdown list populated with the most common service types (TAP (tables), SIA (images), (SCS) Cone search, SSA (Spectra)). If this is provided the registry search would add an extra parameter as follows:

```
services = pyvo.regsearch(keywords=['ukidss'], servicetype='tap')
```

Once this search is done the user is provided (probably in a new page) with the list of results from this search.

The info for each result that show for each user can be as detailed as we want, but the results from the registry search above would return the following fields:

ivoid
cap_index
intf_index
intf_type
intf_role
std_version
query_type
result_type
wsdl_url
url_use
access_url
mirror_url
authenticated_only
security_method_id
*ivoid_*
*cap_index_*
cap_type
cap_description
standard_id
ivoid__
res_type
created
short_name
res_title
updated
content_level
res_description
reference_url
creator_seq
content_type
source_format
source_value
res_version
region_of_regard
waveband
rights
rights_uri
harvested_from

Out of this list *res_title, res_description, content_type, standard_id, access_url* are probably the most important for the user to decide which to use, as well as for us to gather the necessary metadata to find out how to get data from that service.

NV: *"The fields (in the json response) will have to have 'standard esap names' that the GUI will recognize. Like "title, description, type, id, url". These 'standard' names can be translated by specific 'parameter_mapping' for this catalog".*
See:
https://git.astron.nl/astron-sdc/esap-api-gateway/-/wikis/Service-Category:-Query

As an example, we could grab the relevant information, and show it as a list to the user as such:

NV: *"This cannot really work this way. Because communication with the user is through the frontend (GUI). Which is ReactJS, not Python. All communication between frontend and (Django) backend is through http requests (a url with parameters) and responses (a json structure containing the desired information).*
So the pyvo functionality should be implemented in the backend, but workflow is implemented in the frontend"

```
import collections
class RegResults(object):

    SERVICE_IDS = collections.defaultdict(str,
{"ivo://ivoa.net/std/tap" : "tap", "ivo://ivoa.net/std/sia" : "sia"
})

    def __init__(self, results):
        self.title = results["res_title"]
        self.description = results["res_description"]
        self.content_type = results["content_type"].decode("utf-8")
        self.servicetype =
self.SERVICE_IDS[results["standard_id"].decode("utf-8")]
        self.access_url = results["access_url"].decode("utf-8")

results = [RegResults(svc) for i in results]
```

GUI page 1: Enter Keyword (input) + Results (output) + Select VO Service

```
for i in results:

    # Some formatting here..
    print (i.title)
    print (i.description)
    print (i.content_type)
    print (i.servicetype)
    print (i.access_url)
    print ("")
```

NV: *"The results should be returned to the frontend in the http response in json format"*

A user is then able to select the service they want to use with a click which should then direct them to the next ("Query") page.

From the above list in Python, we could just select the one that they've chosen like:

```
 our_service = results[0]
```

GUI page 2: Enter query parameters for selected VO Service (input)

In the next page, the query parameters will depend on the service type.
Ignoring SIA and SSA for now, let's focus on TAP & Cone Search (SSA)

For the TAP Query UI, at the very basic form, we'd provide the user with a textarea element, which is labeled "query".
Here the user would populate an ADQL query that they want to send to the selected service.

NV: *"This is very VO specific. In ESAP the esap query parameters are now mapped to VO query parameters and combined into such a ADQL query. Manually typing in a ADQL query by the user seems a bit like a step back from that?"*

Here is an example of how to query a TAP service, assuming the user has input a query:

```
query = "SELECT TOP 10 * from VVVDR4.Filter"
from astroquery.utils.tap.core import TapPlus
service = TapPlus(url=our_service.access_url + "/sync")
job = service.launch_job(query)
table = job.get_results()
```

In a more advanced UI we could also provide a way for the user to navigate through the list of schemas, tables & columns, using the TAP_SCHEMA tables, which are available for every TAP service.

This could be done by sending TAP queries (see above) that select from the TAP_SCHEMA Schema.

e.g.

```
query = "select+*+from+TAP_SCHEMA.schemas"
query = "select+*+from+TAP_SCHEMA.tables"
query =
"select+*+from+TAP_SCHEMA.columns+where+table_name='II/336/apass9'"
```

In the case of Cone search the UI presents the user with options for ra, dec and radius.

NV: *"Currently a TAP based cone search is already onboard. Does the specific pyvo.conesearch functionality offer something extra?"*

We can then either use PyVO to execute the query:

```
import pyvo
url = "https://irsa.ipac.caltech.edu/SCS?table=fp_psc"
objects = pyvo.conesearch(url, pos=(180, 0), radius = 0.05)
objects.table
objects.fieldnames
objects.getdesc('k_m').description
```

NV: *"The results will need to be written into the json body of the response in the structure that the GUI can understand"*

GUI page 3: Query parameters (input) + Results (output)

Or we can do it manually by generating a Cone Search URL:

NV: *"This would probably not fit in the current model. The query parameters are translated per 'catalog' from ESAP specific parameters to catalog specific parameters.*
VO_REG would be 1 (ESAP) catalog), with 1 translation from ESAP to catalog.
ESAP knows 'ra, dec, fov' and this is currently translated for the VO (ESAP) catalog to 's_ra', 's_dec', 's_fov' as per obscore standard.

This can be a different translation for the VO_REG catalog, this could indeed be 'RA, DEC, SR' as in the example. But if it would have to be different for every underlying service then that would be hard to model."

curl -o out.tbl "https://irsa.ipac.caltech.edu/SCS?table=fp_psc&RA=180.0&DEC=0.0&SR=0.05&format=ipac_table"

NOTE: The above does not describe the ObsCore use case, where we may have a different flow, for example generating a query around a certain ra, dec & sending it to many services.

Tools:

http://saada.unistra.fr/taphandle/?url=http://voparis-tap-astro.obspm.fr:80/tap#

NV: 'Enhanced' version of the UKIDSS use case, with interactive field selection based on the service chosen by the user:

https://app.diagrams.net/#G1zMQADmY3UgI_oJVvAjTwdQIEbyrU1k9d

ESAP IVOA UseCase "UKIDSS" - enhanced
latest update: nv 24 aug 2020