Building a Smart Portal for Astronomy

S. Derriere (sebastien.derriere@astro.unistra.fr), **T. Boch** (thomas.boch@astro.unistra.fr)

Abstract :

The development of a portal for accessing astronomical resources is not an easy task. The ever-increasing complexity of the data products can result in very complex user interfaces, requiring a lot of efforts and learning from the user in order to perform searches. This is often a design choice, where the user must explicitly set many constraints, while the portal search logics remains simple.

We investigated a different approach, where the query interface is kept as simple as possible (ideally, a simple text field, like for Google search), and the search logics is made much more complex to interpret the query in a relevant manner.

We will present the implications of this approach in terms of interpretation and categorization of the query parameters (related to astronomical vocabularies), translation (mapping) of these concepts into the portal components metadata, identification of query schemes and use cases matching the input parameters, and delivery of query results to the user.

The CDS portal

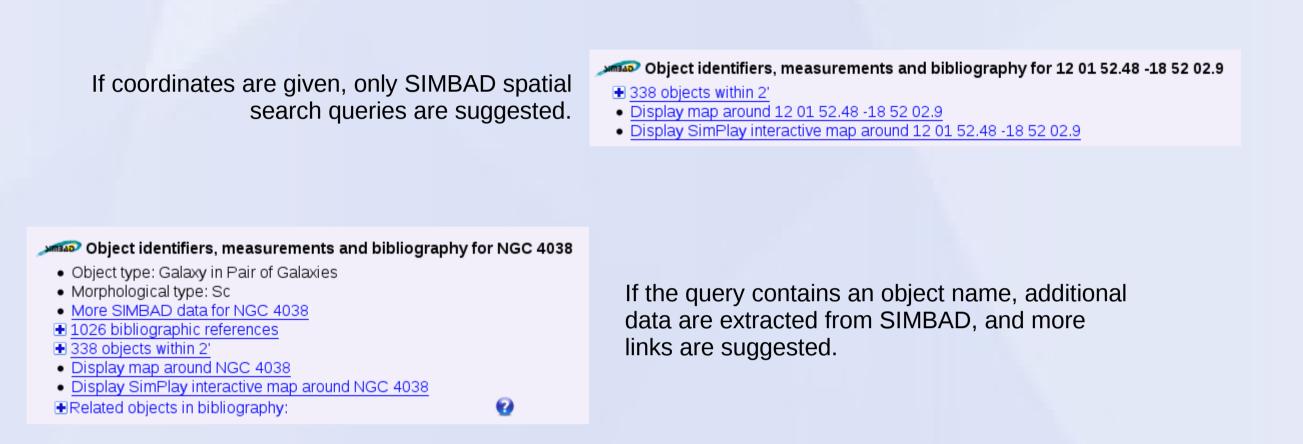**http://cdsportal.u-strasbg.fr**

The CDS provides 3 main services : SIMBAD, Aladin and VizieR, and each has one or several dedicated query pages. The CDS portal is an effort to provide simultaneous access to these services through a unique and simple query page.

The input form is a single text field, where users can type either an object name or sky coordinates :

Target: VV 245b GO
J2000 position for VV 245b (NGC 4038): 12 01 52.48 -18 52 02.9

If the input string is recognized as a valid object name, the corresponding object and its coordinates are explained just below the text input field. This is done on-the-fly using AJAX and the Sesame name resolver.

When the user performs a query, SIMBAD, Aladin and VizieR are queried simultaneously, and the different results are aggregated into a single result page. The results will be slightly different, depending on the input being an object name or sky coordinates.

If coordinates are given, only SIMBAD spatial search queries are suggested.

If the query contains an object name, additional data are extracted from SIMBAD, and more links are suggested.
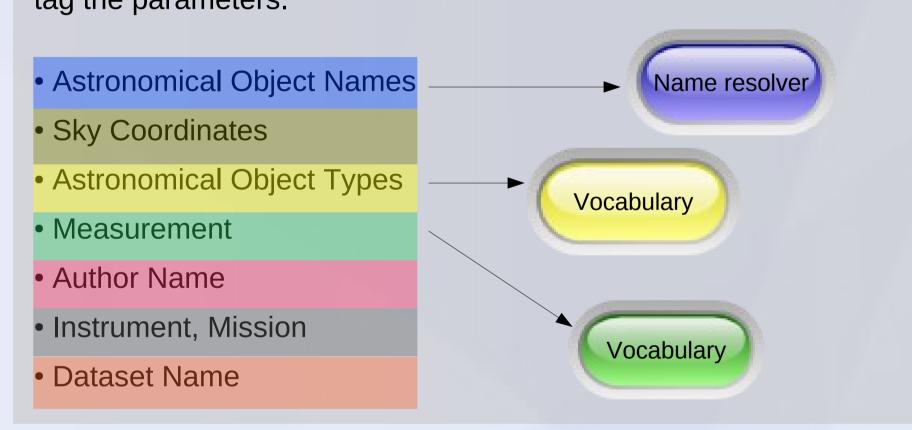
Towards a smart portal

The next step will be to add additional capabilities to the portal, allowing the user to input parameters other than object names or coordinates, and to use these parameters in the context of the various services.

In order to achieve this, we need 3 elements :
• parameters categorization (using vocabularies or services)
• mapping from reference vocabularies to service keywords
• query templates for each service

Parameters categorization

We identified a few categories of terms that can be used in queries, and for each category, a standard vocabulary or service allowing to tag the parameters.

• Astronomical Object Names → Name resolver
• Sky Coordinates
• Astronomical Object Types → Vocabulary
• Measurement
• Author Name → Vocabulary
• Instrument, Mission
• Dataset Name

Vocabularies mapping

The target services do not need to use internally the standard vocabulary. All that is needed is a mapping from the standard vocabulary to the services' internal vocabulary.

Example : the vocabulary of astronomical object types can be derived from the Ontology of Astronomical Object types[1]. Object types recognized by VizieR form a list of keywords. SIMBAD has its own classification scheme[2].
We need a mapping from "Globular Cluster" to "Globular_Clusters" (VizieR) or "GlC" (SIMBAD), or from "Galaxy Clusters" to "Clusters_of_galaxies" (VizieR) or "ClG" (SIMBAD).
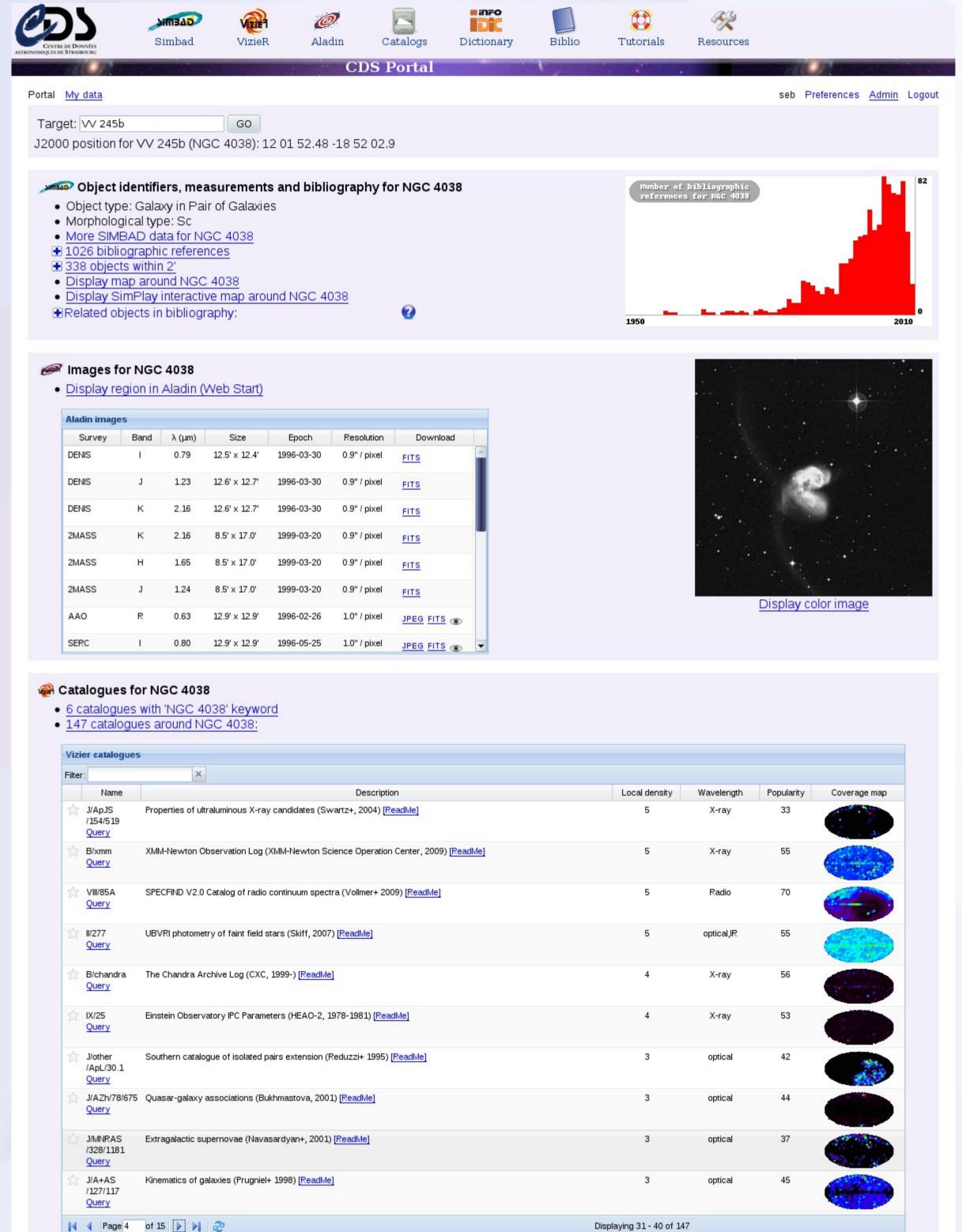
Query Templates

For each target service, we identify several query templates. A query template consists of :
• Categorized input parameter(s)
• A query link using these parameters

For example, if we have an object name and a measurement :
• In SIMBAD, we can search the value of this measure for this object
• In VizieR, we can search for catalogues containing a column corresponding to the measurement and covering the sky area containing our object

If we have a sky coordinate and an object type :
• In SIMBAD, we can search for object of this type around this sky location
• In VizieR we can give a list of catalogues dealing with this type of objects and covering this sky region

Use case

The user interface remains as simple as it can be : a simple text input.

Virgo cluster redshift GO

The input words are categorized... in this case, the system detects an astronomical object name and a measurement.

Virgo cluster redshift GO

The parameters are translated into each service's own keywords.

redshift -> UCD[3] src.redshift (VizieR)
redshift -> redshift or RV(Z) (SIMBAD)

All query templates involving these two categories of parameters are tested.

http://vizier.../VizieR?-ucd=src.redshift&-c=Virgo+Cluster

http://simbad.../%25RV(Z)"%0D%0Aquery+id+Virgo+cluster

The user is presented with details on how the query was interpreted, together with the result, for each query template, in each service.

List of catalogues containing redshift measurements around Virgo cluster

Redshift value for Virgo cluster   0.0038

This approach is generic enough to be extended to services other than the CDS ones, like ADS or others. The contextual description of service capabilities, together with the definition of standard vocabularies[4] will soon enable smart behaviors in astronomy portals, and aggregation of results from different independent services.

[1] http://ivoa.net/Documents/Notes/AstrObjectOntology/
[2] http://simbad.u-strasbg.fr/simbad/sim-display?data=otypes
[3] http://ivoa.net/Documents/latest/UCDlist.html
[4] http://ivoa.net/rdf/

CDS CENTRE DE DONNÉES ASTRONOMIQUES DE STRASBOURG