Chatting with the astronomical data services

André Schaaff[1], Alexis Guyot[2], Thomas Boch[1], Sébastien Derriere[1]

Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, F-67000 Strasbourg, France[1]

IUT Dijon, France[2]

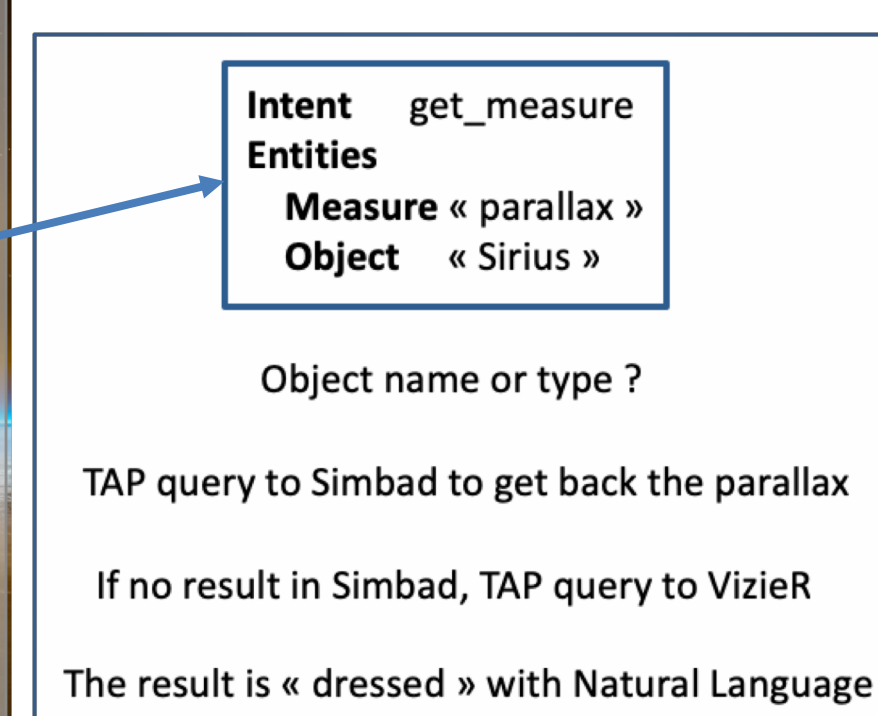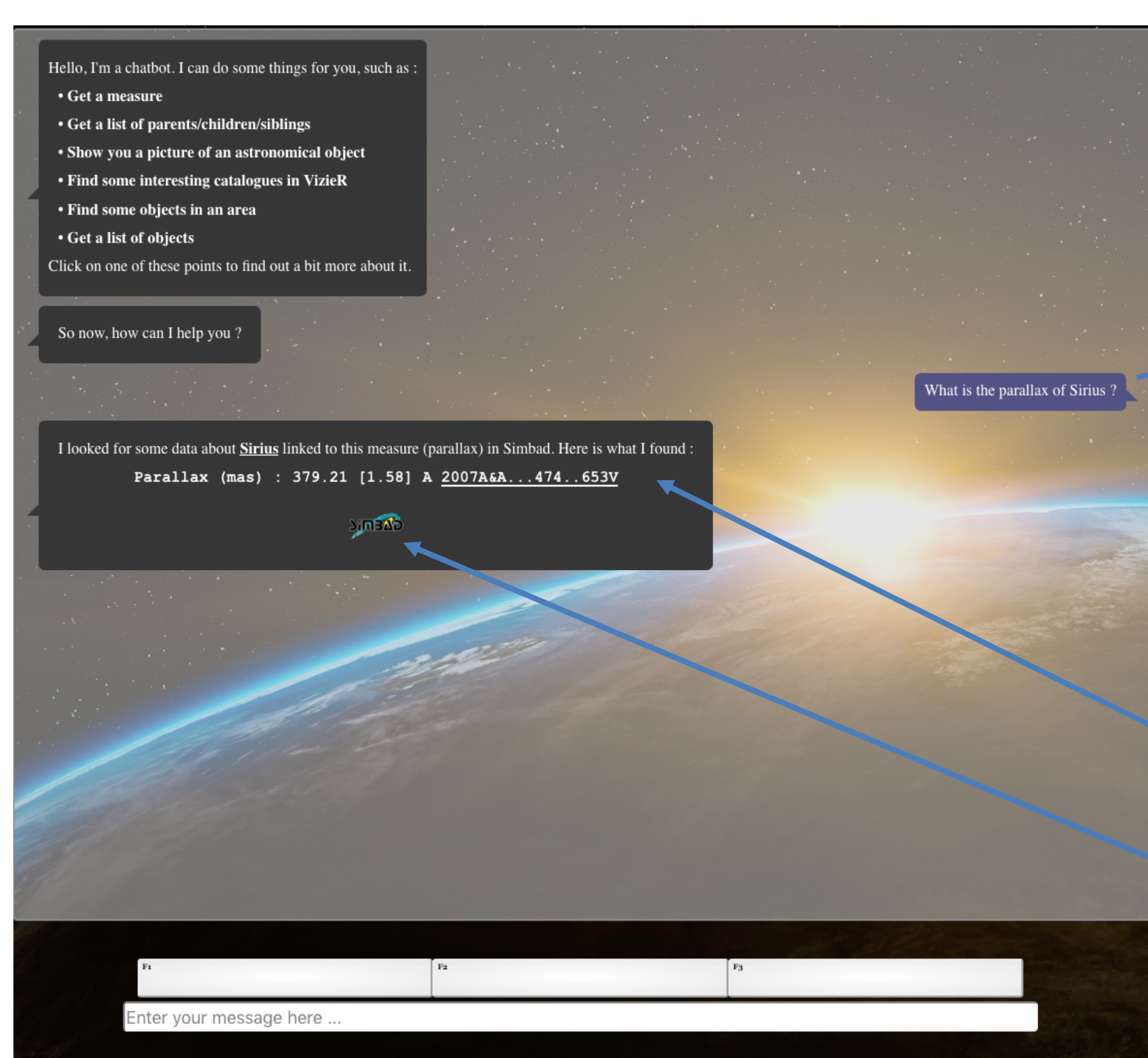*Demo on demand ! During the poster sessions or at the CDS booth*

At the beginning…

We use more and more natural language to interact with assistants during heterogeneous requests (weather, booking, shopping, etc.). Chatting with various devices is becoming common. By following this observation is it realistic to propose it as a new way of interacting with our data services, as an alternative to the traditional forms that expose parameter fields, check boxes, etc.?
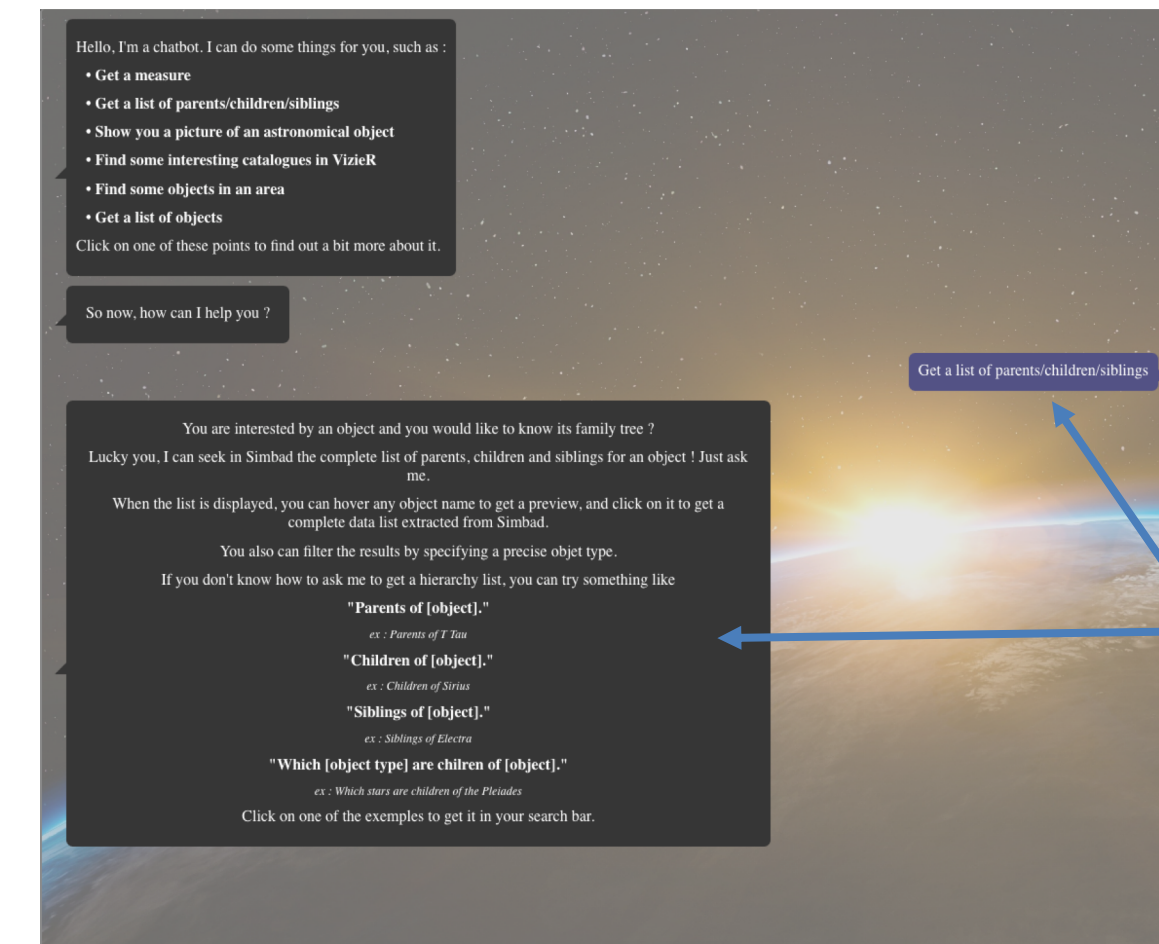
To answer this question, it is necessary to first answer the most fundamental question: is it possible to satisfy in this way the needs of professional astronomers ?

We present here how we try to answer this question in the frame of the CDS data services.

Overview

Our Chatbot has a user interface written in Javascript which manages the design and the link with the CDS services. The natural language recognition is delegated to a customized Dialogflow agent. The user writes a request in natural language. The request sentence is then analyzed by the agent which returns its structure. We use then this structure to translate it in one or more queries to our services. In the last step, we "dress" the result(s).


```
Intent     get_measure
Entities
    Measure « parallax »
    Object    « Sirius »

Object name or type ?

TAP query to Simbad to get back the parallax

If no result in Simbad, TAP query to VizieR

The result is « dressed » with Natural Language

I looked for some data about Sirius …
```

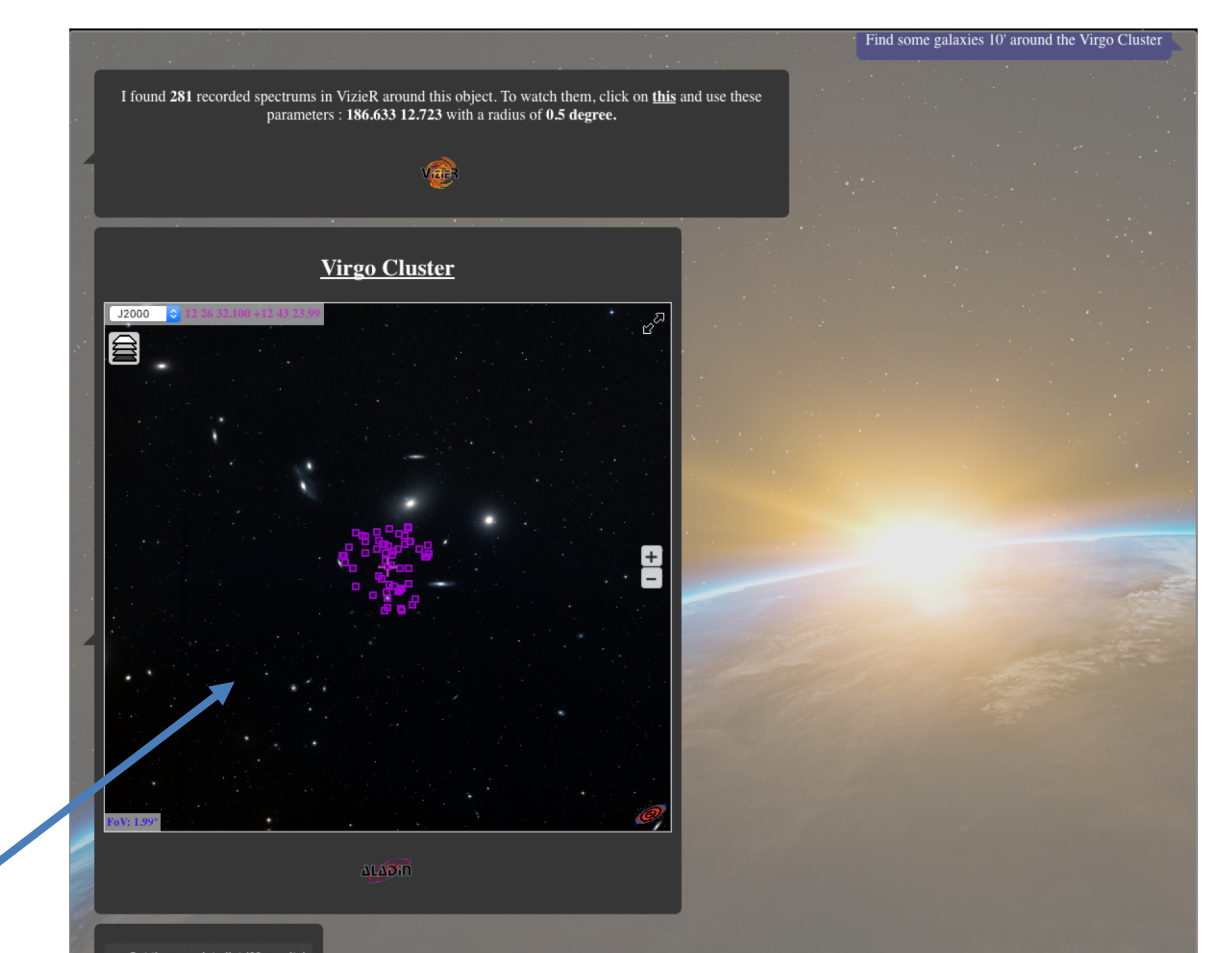*The source of the result is specified*

More information

Help

Result sources

Arcturus is implicitly in the context

Aladin Lite (easy widget embedding)

We did not start from scratch !

We have useful tools and resources (name resolver, authors in Simbad, missions and wavelengths in VizieR, etc.) and the Virtual Observatory (VO) brings us standards like TAP/ADQL, UCDs, …, implemented in the CDS services. We have defined a first list of "typical" queries.

UCD: phys.temperature.effective

*Identifier: *alf CMa*

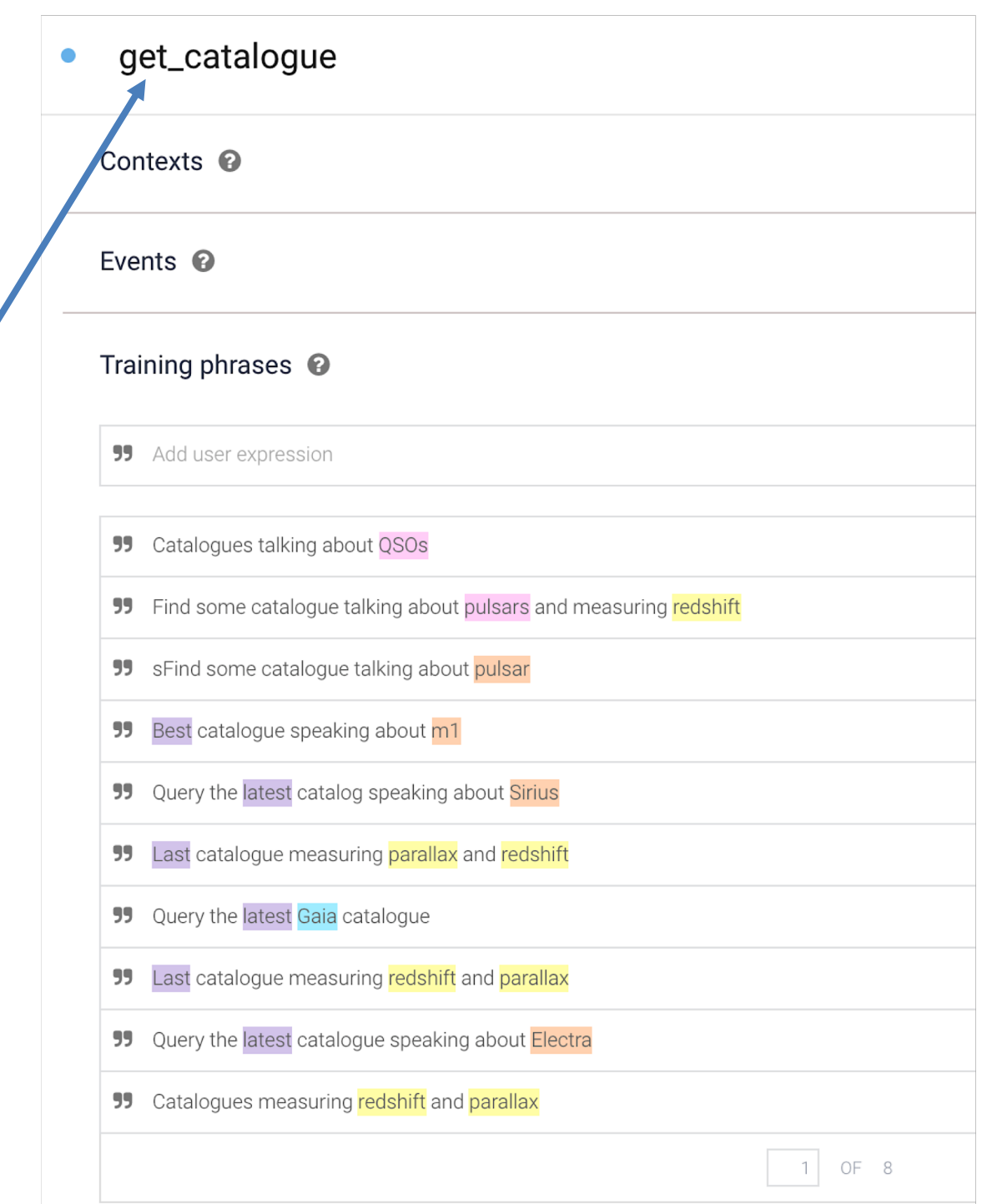*What is the effective temperature of Sirius ?*

Translation in ADQL

```
SELECT "VI/137/gum_mw"."Teff
FROM "VI/137/gum_mw"
WHERE 1 = CONTAINS(POINT('ICRS', "VI/137/gum_mw"."RAJ2000",
"VI/137gum_mw"."DEJ2000"), CIRCLE('ICRS', 101.287155333,
-16.716115861, 20/3600.)) ;
```

The natural language recognition

We made a first experiment with Stanford NLP libraries but it required too much time to improve it. In a second step we chose Dialogflow to manage the recognition side. With Dialogflow we delegate a greater part of the tool than with the first approach and we can thus spend more time to better the quality of the user query translation. The link to the services and de facto to the data is the key part and our development is designed to be able to switch to other NLP tools. The Dialogflow agent must be customized (Intents, etc.) and we evaluate the user requests (Machine Learning).

Status

The interoperability standards, enabled by the IVOA, are a mandatory backbone providing us a part of the mechanism useful to translate a natural language request to a query understandable by our services.

The first presentations of the prototype encouraged us to continue and to improve the fitting of the results with the astronomers requests.

Next steps

- Open it to the community to improve it.
- More sophisticated chat with the user to refine a request.
- Evaluate the extension to other VO services, outside the CDS.
- Add a voice recognition.

Perspectives

There are many potential applications.

- The most important challenge is probably to integrate it into the CDS portal, provided that sufficient accuracy and quality criteria are reached.
- In the field of education we plan to couple it to a virtual reality device (for example, to browse HiPS) to offer direct voice control.

A best effort development to explore a new way to access the astronomical resources.
Comments, necessarily enthusiastic, are welcome. Collaborations too.