Chatting with the services. VO Standards and NLP.

André Schaaff, Thomas Boch, Sébastien Derriere
Centre de Données astronomiques de Strasbourg

Alexis Guyot
IUT Dijon

IVOA Victoria, 28/05-01/06/2018
Apps Session 2

CENTRE DE DONNÉES
ASTRONOMIQUES DE STRASBOURG

IVOA

Astericcs
Astronomy ESFRI & Research Infrastructure Cluster

Purpose

- First presentation in Santiago Semantics WG: the aim was to relate a budding R&D work, Natural Language Processing applied to the querying of astronomical data services and to collect comments (maybe enthusiastic), ideas, recommendations, or to initiate collaborations
- We are not geeks and we would like to propose this way of interaction in … the future.
- Voice recognition ?
- … is it possible to reach query results satisfying professional astronomers ?

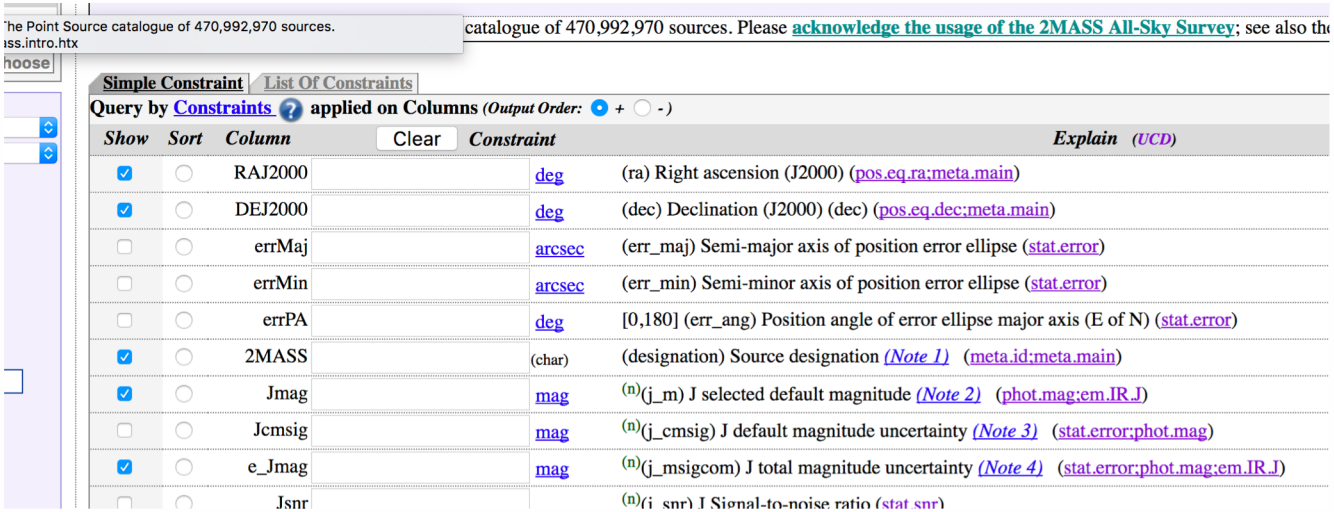
Comments

- Link with the VO ?
  - We use VO standards like TAP, UCDs, …
  - We use tools based on ML, a link with KDD.

- The VO enables the interoperability which is a mandatory backbone, helping us to query our services in NL and which will be useful in a step++ to query the whole VO through this way.

NL interaction is “common”

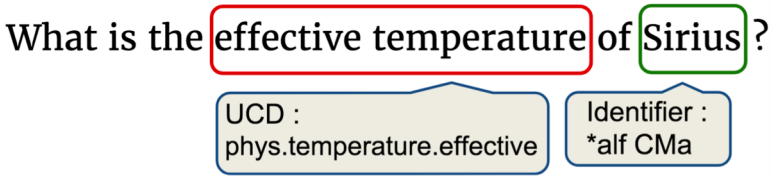
- Because it is natural (Apple Siri, Ok Google, Microsoft, etc.), assistants at your home to book a restaurant, find a medical appointment, ... in the every day life.
- We use forms (parameter fields, checkboxes, etc.) to query astronomical data services.



What we are doing

- Learning about NLP (basis, tools, examples, …)
  - Scope of the study: a first set of queries
  - Too large -> too much time and resources.

- A pragmatic approach (more R&D than R), we have
  - authors in Simbad, VizieR
  - missions and wavelengths in VizieR
  - DJIN to recognize identifiers in a text
  - UCDs
  - a name resolver
  - ADQL / TAP
  - …



A first set of queries in NL

1. **What is the redshift of 3C273? What is the redshift of the Virgo Cluster?**
2. **What is the parallax of Barnard's star? What is the distance of Barnard's star? What is the proper motion of Barnard's star?**
3. **What is the effective temperature of Sirius?**
4. **What are the galactic coordinates of Geminga?**
5. **Which galaxy interacts with NGC 4038?**
6. **Show me an image of the Pleiades in the K band**
7. **How many QSOs are there at redshift larger than 6? How many QSOs are there at z>6?**
8. **What is the redshift of galaxies members of the Virgo cluster?**
9. **Find globular clusters within 3° of M31. Find globular clusters in M31.**
10. **Query the latest Veron catalogue**
11. **What is the period of Algol? List of periods of Algol-type stars.**

From Natural Language to ADQL

- In the literature, many examples of translation from NL to SQL.
- With ADQL :

List the QSOs at Z > 6.

```
SELECT main_id, oid, rvz_redshift
FROM basic
WHERE otype = -14680064 AND rvz_redshift > 6;
```

Simbad, TAP query

What is the effective temperature of Sirius ?

```
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.)) ;
```

VizieR, TAP Query

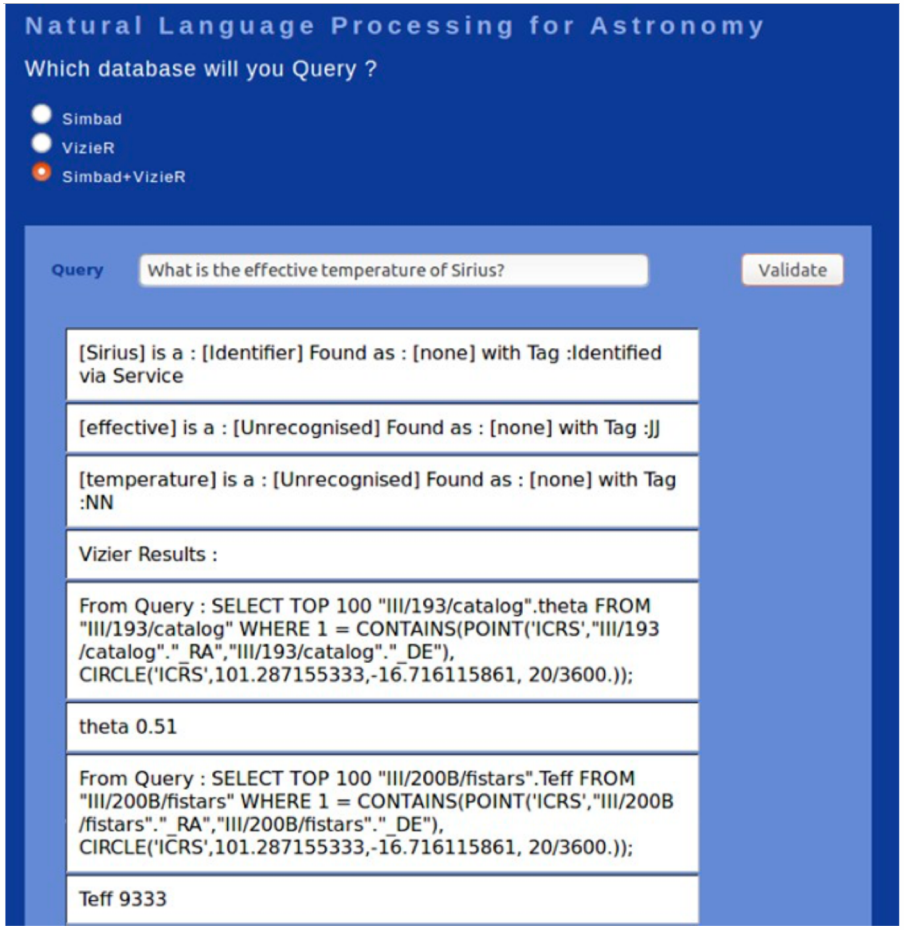
Comments…

Previous prototype

- Was based on Stanford NLP (POSTagger), DJIN, IVOA UCD and ADQL/TAP, …
- Far from a chatbot

What is the effective temperature of Sirius ?

How many planets orbit Kepler 20 ?

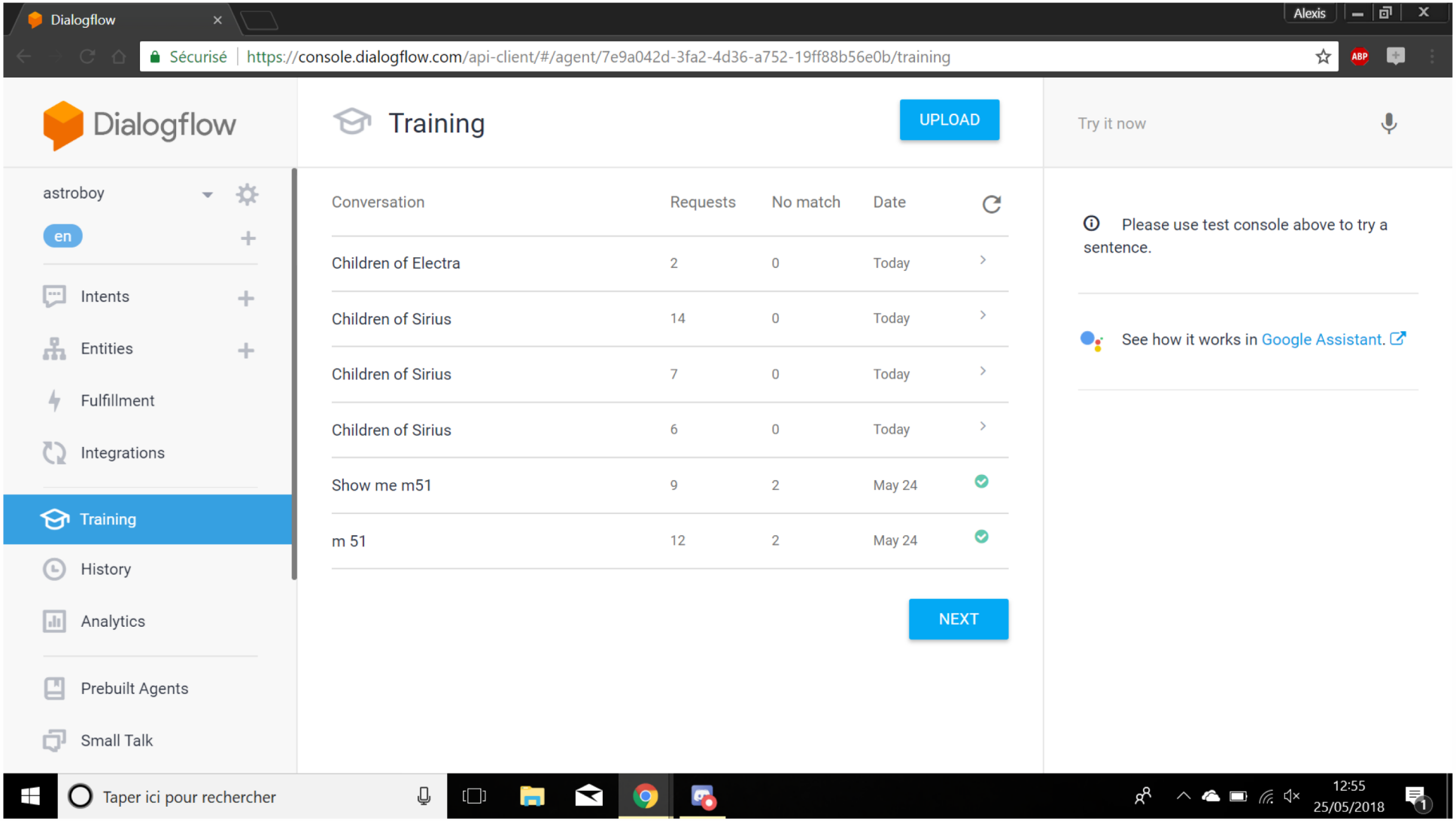


What is the redshift of galaxies members of the Virgo cluster ?

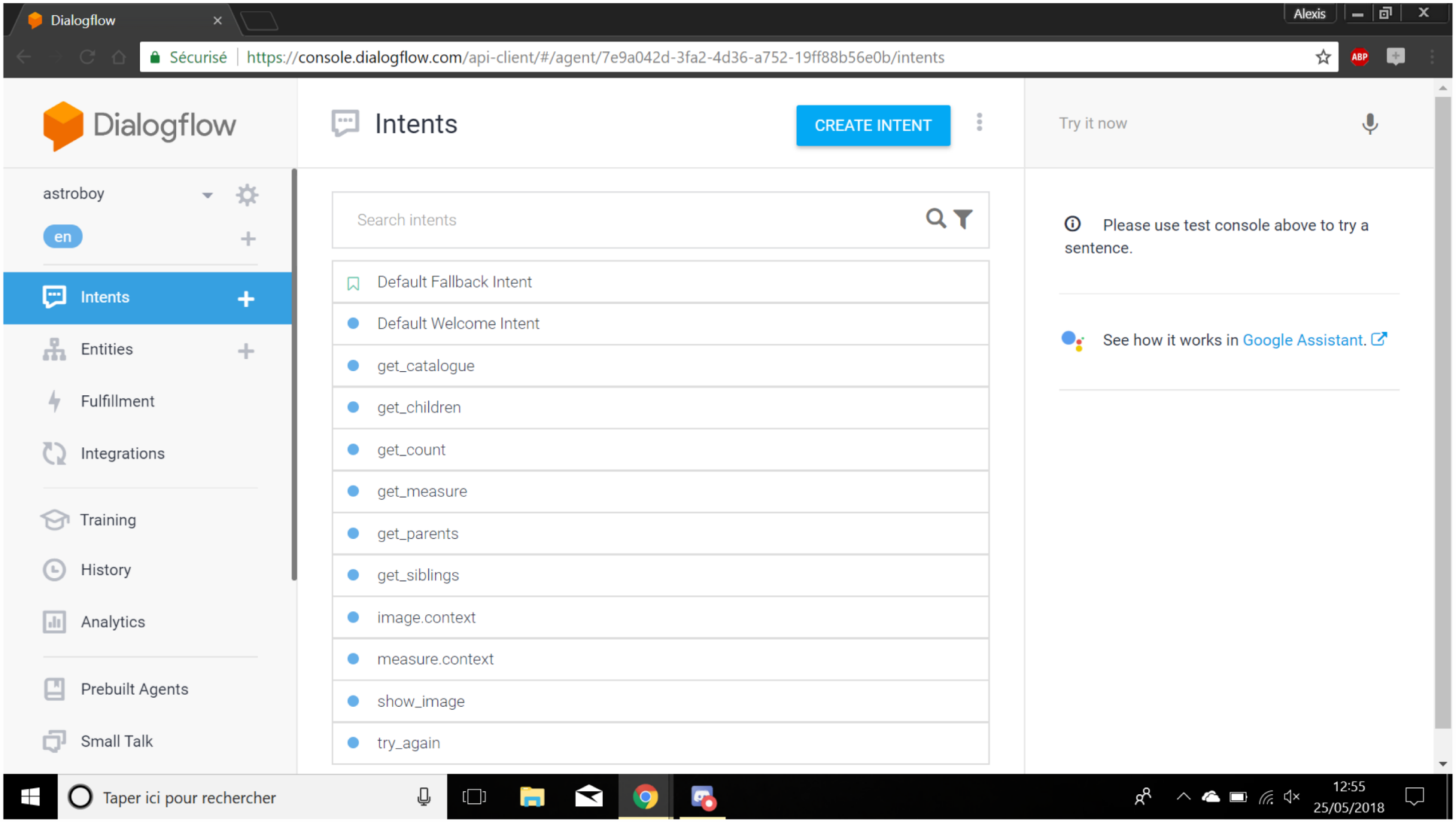
New prototype

- Chatbot approach to reduce the gap between “good” queries and unprecise / ambiguous queries.
- Based on Google Dialogflow
  - Many features
  - Machine learning
  - Engine is on the Google side, it is free, multiple users at the same time through a “Session” ID.

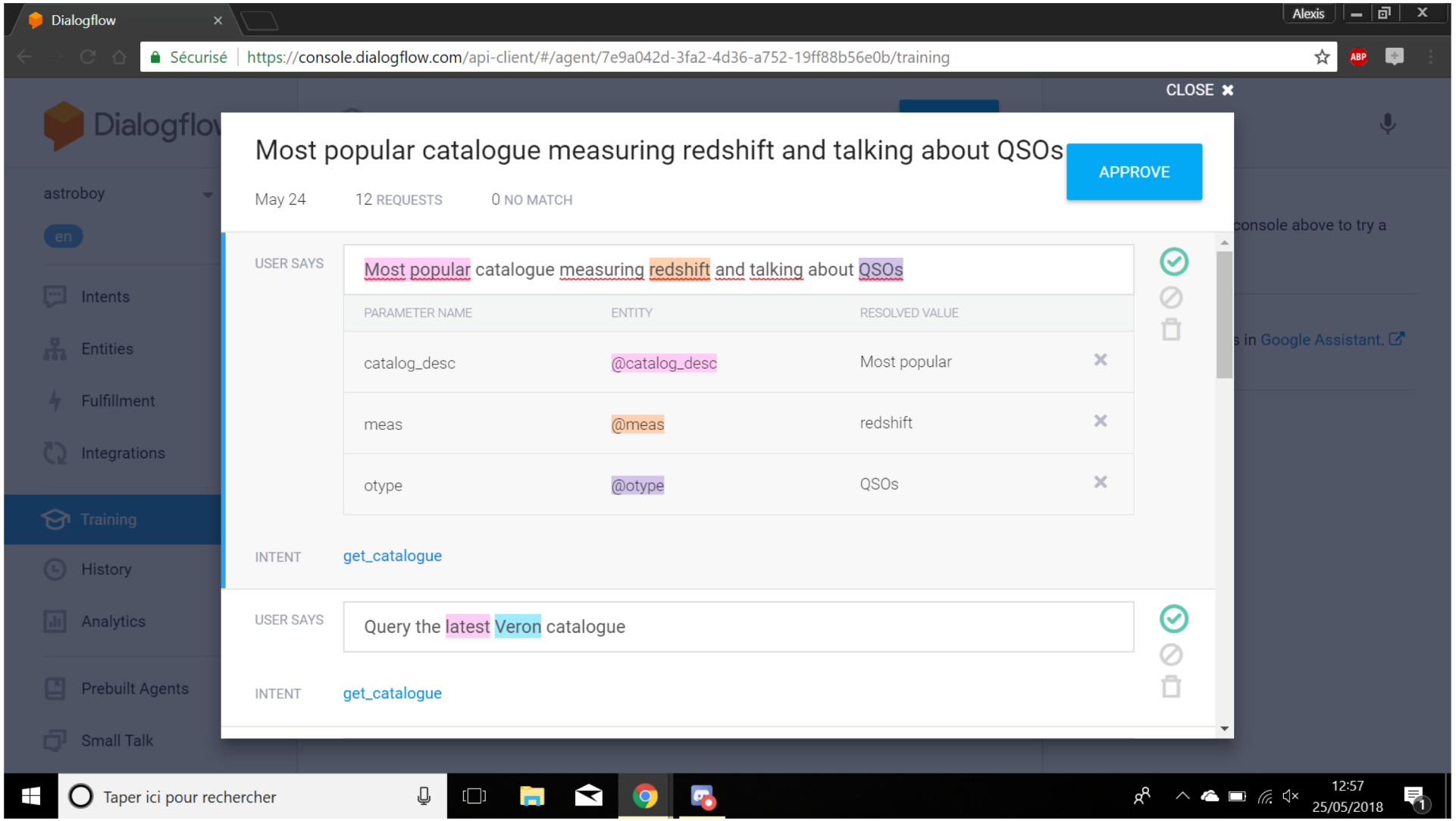
Training


Dialogflow
Alexis
Sécurisé | https://console.dialogflow.com/api-client/#/agent/7e9a042d-3fa2-4d36-a752-19ff88b56e0b/training
Dialogflow
Training
UPLOAD
Try it now
astroboy
en
Intents
Entities
Fulfillment
Integrations
Training
History
Analytics
Prebuilt Agents
Small Talk
Conversation
Requests
No match
Date
Children of Electra 2 0 Today
Children of Sirius 14 0 Today
Children of Sirius 7 0 Today
Children of Sirius 6 0 Today
Show me m51 9 2 May 24
m 51 12 2 May 24
NEXT
Please use test console above to try a sentence.
See how it works in Google Assistant.
Taper ici pour rechercher
12:55
25/05/2018

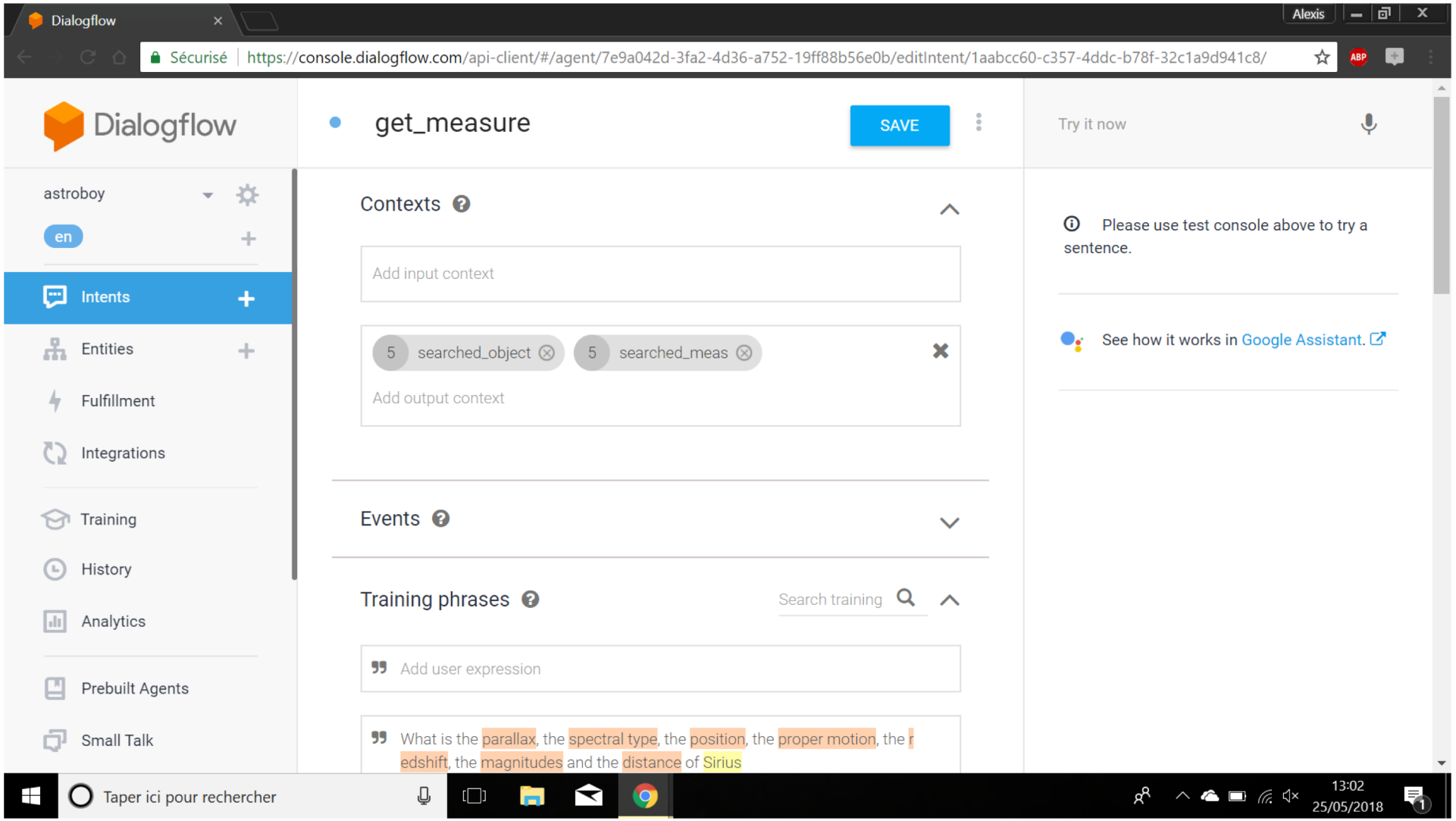
Intents


Dialogflow
Alexis
Sécurisé
https://console.dialogflow.com/api-client/#/agent/7e9a042d-3fa2-4d36-a752-19ff88b56e0b/intents
Dialogflow
Intents
CREATE INTENT
Try it now
astroboy
en
Search intents
Intents
Entities
Fulfillment
Integrations
Training
History
Analytics
Prebuilt Agents
Small Talk
Default Fallback Intent
Default Welcome Intent
get_catalogue
get_children
get_count
get_measure
get_parents
get_siblings
image.context
measure.context
show_image
try_again
Please use test console above to try a sentence.
See how it works in Google Assistant.
Taper ici pour rechercher
12:55
25/05/2018

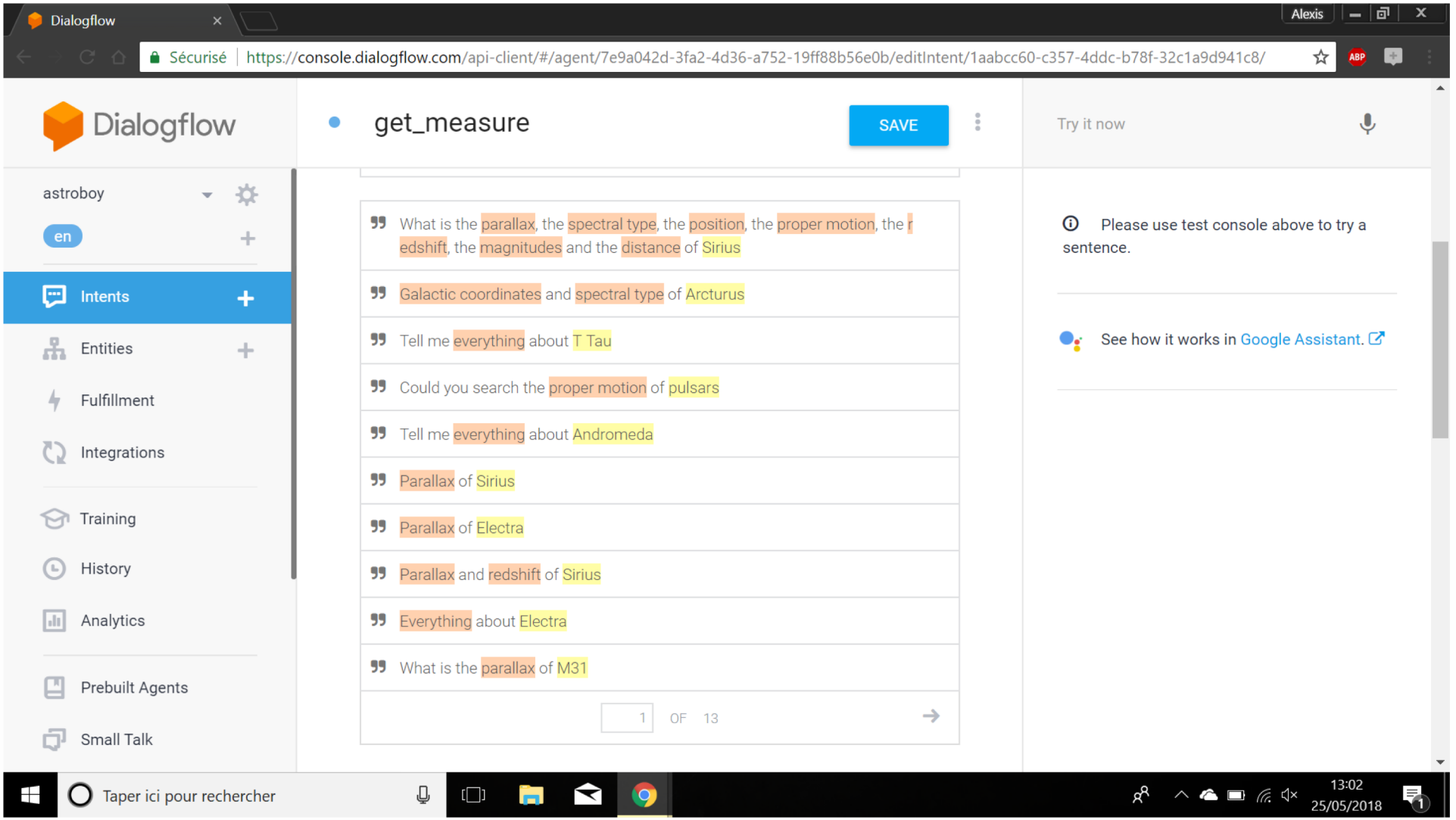
User query and intent



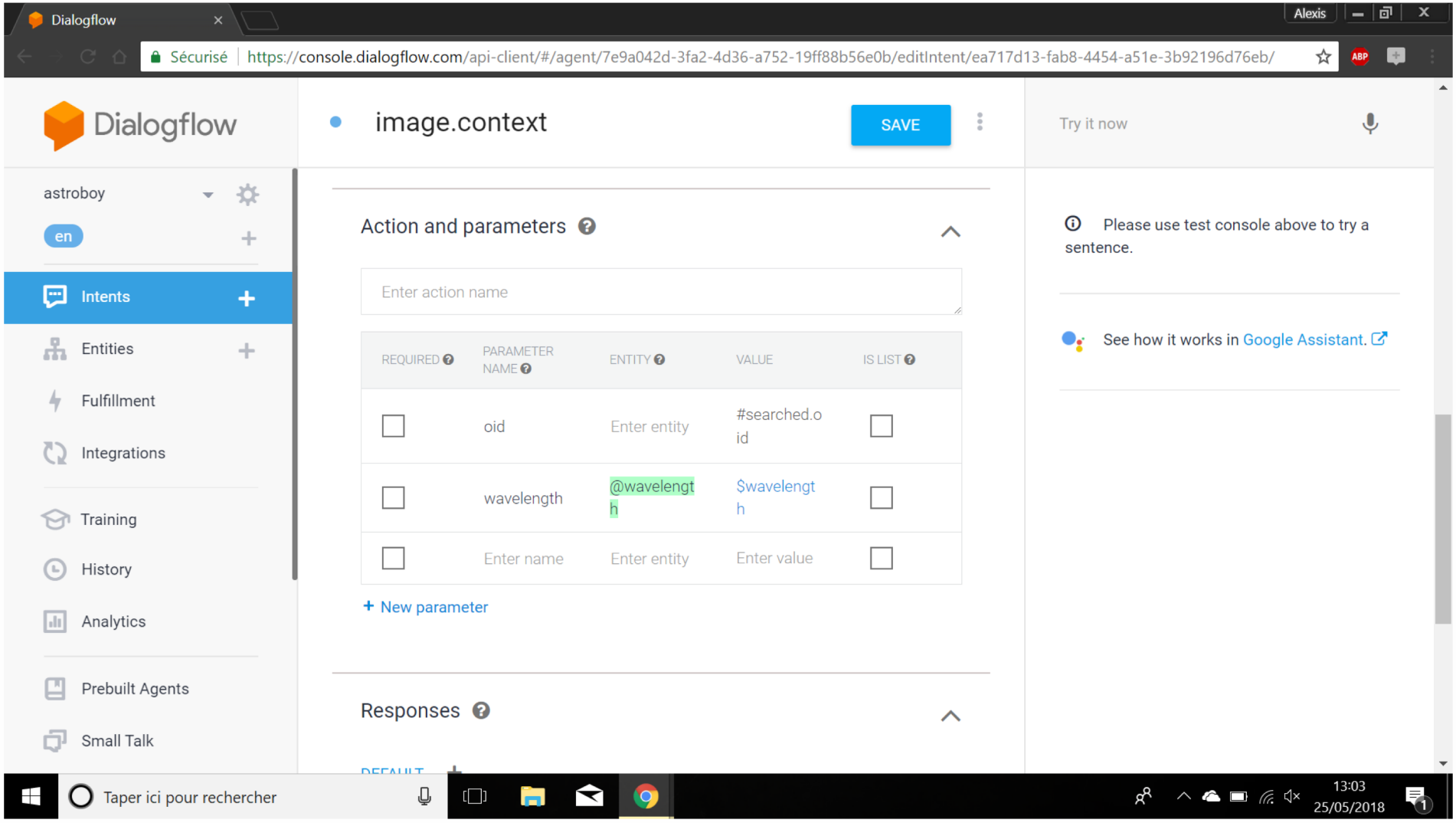
Examples of sentences linked to get_measure



Examples of sentences linked to get_measure (2)



Parameters


Dialogflow
Alexis
Sécurisé
https://console.dialogflow.com/api-client/#/agent/7e9a042d-3fa2-4d36-a752-19ff88b56e0b/editIntent/ea717d13-fab8-4454-a51e-3b92196d76eb/
Dialogflow
image.context
SAVE
Try it now
astroboy
en
Intents
Entities
Fulfillment
Integrations
Training
History
Analytics
Prebuilt Agents
Small Talk
Action and parameters
Enter action name
REQUIRED
PARAMETER NAME
ENTITY
VALUE
IS LIST
oid
Enter entity
#searched.oid
wavelength
@wavelength
$wavelength
Enter name
Enter entity
Enter value
+ New parameter
Responses
Please use test console above to try a sentence.
See how it works in Google Assistant.
Taper ici pour rechercher
13:03
25/05/2018

Before the demo

What is the parallax of Sirius ?

ML + Matching →

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

The parallax of Sirius is …

Demo

- Only the user interface is on my laptop

Please, give us a feedback

- Comments ?
  - Enthusiastic like in Santiago ?
- Ideas ?
- Recommendations ?
- Similar works ?

Conclusion

- We are testing and improving step by step.
- We are (somewhere) on a way to interact with data services differently.
- We will enjoy to have a few volunteers to try it and to give us a feedback.