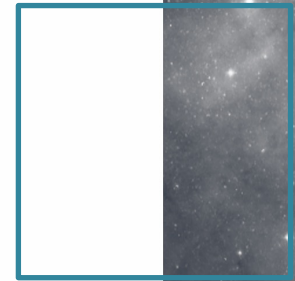# R&D and IT internships@ CDS

## …and various developments

André Schaaff on behalf of the CDS Team

**CDS Scientific Council 2023**

# Foreword

- Fast & various technological evolutions are explored each year
  - interactions, visualization, mobility, components, Big & Open Data, Clouds, etc.), and now AI with its realities and fantasies.

- The R&D activity is well identifed since 2 decades, structured and involving several persons of the staff with the help of (10 to 12 per year) interns and short contracts.

- By the way, a continous training for many people.

# Internship programme

- After a slowdown in 2021, the return to the normal level was confirmed in 2023.

- It is becoming usual: after his internship, Julien Abid was hired as apprentice for one year from September 2023.

- IT Job Market still tight and hiring is not easy. Interns and apprentices are potential candidates for our IT positions.

# illustrations

- More to show the diversity, from experiments / proof of concept to production work.

- IVOA standards are often used (like VOTable and many others) but also developed (like MOC).

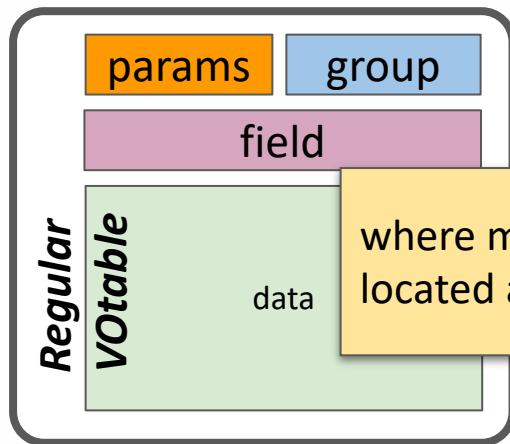- Not possible to enter too much into details.

- Neither exhaustive.

- Aladin Desktop is regularly used for preparing and checking observations.

- It allows the overlay of a large set of instrumental Fields of View represented in VOTable XML.

- Such FoV can also be created by drawing with a graphical editor.

- Update with recent IVOA standards needed for better interoperability.

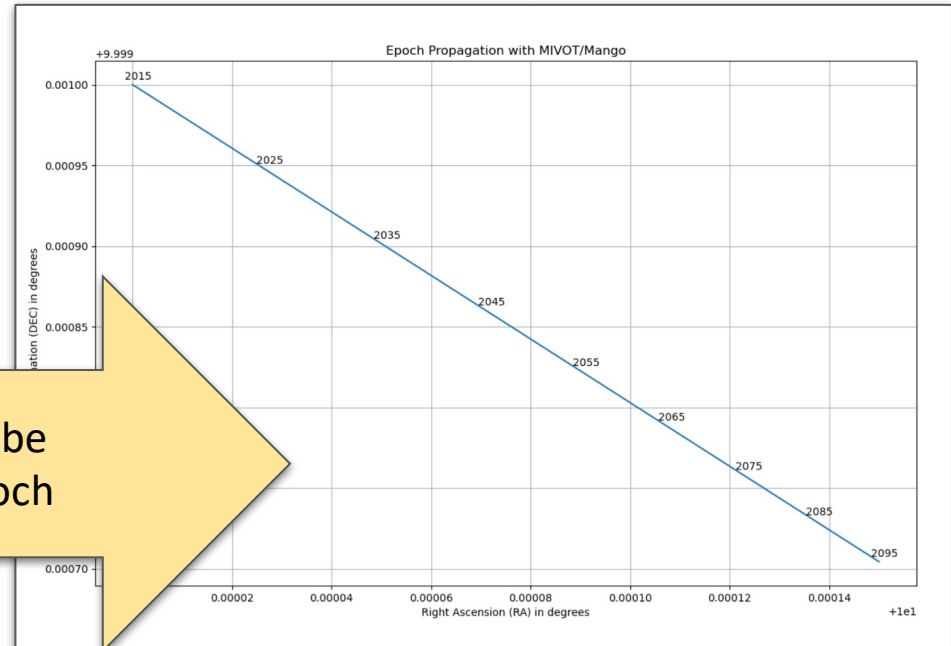**L. Michel**
**Intern:** **Somia Floret** (UTBM)

# The challenge:



Epoch Propagation with MIVOT/Mango

**In this VOTABLE, We have somewhere:**
- Position
- Proper motion
- Parallax
- Radial velocity
- Desired metadata

Regular VOtable: params, group, field, data

where my object will be located at a given epoch

The challenge: **find an appropriate way to do it**

# MIVOT: add a Mapping Block above the data Table

```xml
<VODML xmlns:dm-mapping="http://www.ivoa.net/xml/mivot">
    <REPORT status="OK">hand-made mapping</REPORT>
    <MODEL name="meas" url="https://www.ivoa.net/xml/Meas/20200908/Meas-v1.0.vo-dml.xml" />
    <MODEL name="coords" url="https://www.ivoa.net/xml/STC/20200908/Coords-v1.0.vo-dml.xml" />
    <MODEL name="mango" />
    <MODEL name="ivoa" url="https://www.ivoa.net/xml/VODML/IVOA-v1.vo-dml.xml" />
    <GLOBALS>
        <INSTANCE dmid="SpaceFrame_ICRS" dmtype="coords:SpaceSys">
            <INSTANCE dmrole="coords:PhysicalCoordSys.frame" dmtype="coords:SpaceFrame">
                <INSTANCE dmrole="coords:SpaceFrame.refPosition" dmtype="coords:StdRefLocation">
                    <ATTRIBUTE dmrole="coords:StdRefLocation.position" dmtype="ivoa:string" value="NoSet" />
                </INSTANCE>
                <ATTRIBUTE dmrole="coords:SpaceFrame.spaceRefFrame" dmtype="ivoa:string" value="ICRS" />
            </INSTANCE>
        </INSTANCE>
    </GLOBALS>
    <TEMPLATES>
        <INSTANCE dmtype="mango:EpochPosition">
            <REFERENCE dmrole="coords:Coordinate.coosys" dmref="SpaceFrame_ICRS"/>
            <ATTRIBUTE dmrole="mango:EpochPosition.longitude" dmtype="ivoa:RealQuantity" ref="pos_RA"/>
            <ATTRIBUTE dmrole="mango:EpochPosition.latitude" dmtype="ivoa:RealQuantity" ref="pos_DEC" />
            <ATTRIBUTE dmrole="mango:EpochPosition.pmLongitude" dmtype="ivoa:RealQuantity" ref="pm_RA" />
            <ATTRIBUTE dmrole="mango:EpochPosition.pmLatitude" dmtype="ivoa:RealQuantity" ref="pm_DEC"/>
            <ATTRIBUTE dmrole="mango:EpochPosition.pmCosDeltApplied" dmtype="ivoa:boolean" value="true"/>
            <ATTRIBUTE dmrole="mango:EpochPosition.radialVelocity" dmtype="ivoa:RealQuantity" ref="RV"/>
            <ATTRIBUTE dmrole="mango:EpochPosition.parallax" dmtype="ivoa:RealQuantity" ref="PARALLAX" />
            <ATTRIBUTE dmrole="mango:EpochPosition.epoch" dmtype="coords:Epoch" value="J2016.0" unit="year"/>
        </INSTANCE>
    </TEMPLATES>
</VODML>
```

● The space coordinate system is a **GLOBAL** object that can be referenced by any other MIVOT element

● Each table row can be interpreted as an instance of the class **EpochPosition** of the MANGO model

# MIVOT: PyVO implementation

## Astropy

MIVOT block extraction

Next data row readout

## PyVO

Build a generic model instance by resolving the internal MIVOT references

Build an XML serialization of the data row

Build a Py dictionary from XML view

Build/update a dynamic Python object from the dictionary

- No model specific code
  - The compliance with the model structure is not checked at this level
  - Can be done with the validator

Object denoting the moving position allowing to compute the epoch propagation

# ☐ Chatting with the services

- **A long-term work started** in 2017, Natural Language query translation to **understandable queries by** the **CDS services.**

- **Several presentations** at IVOA, ADASS, Astroinformatics.

- Following chatGPT advent in November 2022 experiments are on going in 2023 with OpenAI API.
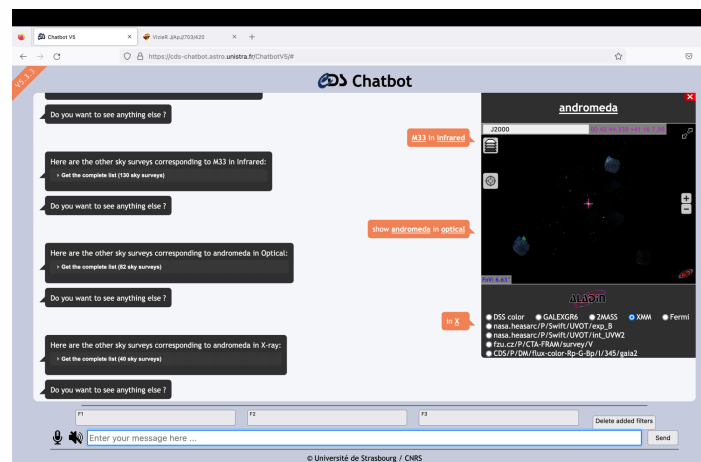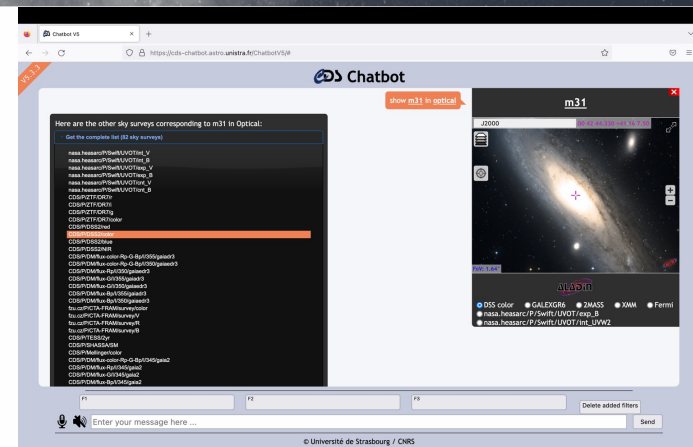




**A. Schaaff, T. Boch, S. Derriere**
**Interns: several students mainly from UTBM and cited in previous CDS scientific Councils**

# ☐ Chatting with the services

- OpenAI is making its API available to us so that we can use one of the Large Language Models (GPT-3.5 turbo in our case).

- It is possible to fine-tune the Model by making it learn our own data.

- In an OpenAI API call, you can specify functions and have the model intelligently choose to output a JSON object containing arguments to call one or many functions.
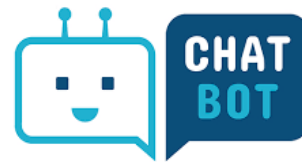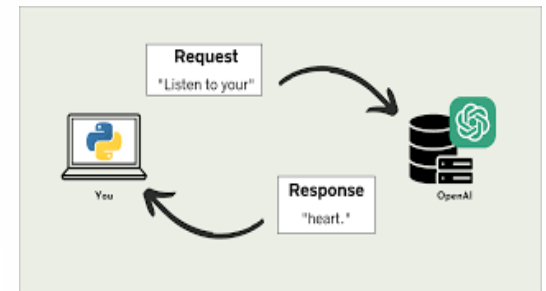


**A. Schaaff, T. Boch, S. Derriere**
**Intern: Robert habokyan (UTBM)**

# Example using Simbad and the criteria

**An example using Simbad and a criteria**

The user sends his request

```
You: show me 15 astronomical objects with a right ascension greater than 350
```

Here the bot understood via the function call. He knows the criteria from his apprenticeship (fine tune)

arguments='{\n"operator": ">",\n"criterion": "right_ascension",\n"filter_value": "350",\n"number": "15"\n}', name='get_criteria_information'),

Then, using an astroquery with its elements, we obtain this result:

The same thing can be done for the total number

of these objects ( a count query)



**A. Schaaff, T. Boch, S. Derriere**
**Intern: Robert habokyan (UTBM)**

# Aladin VR

- A Virtual Reality application based on Aladin Lite using HiPS to display 360 degrees panoramas.

- For whom ?
  - Everybody from young children to seniors
  - From home and regardless of the skill level.

- Why ?
  - Immerge the user in a real space environment
  - Introduce the users to scientific culture and data usage.

**S. Derriere**
**Intern: Flavian Theurel** (UTBM)

# ☐ Aladin VR (2)

- A set of Guided tours
  - Description and explanation about particular objects and phenomenon
  - List of celestial objects of interest.
- Tested with the Meta Quest.



**Derriere S.**
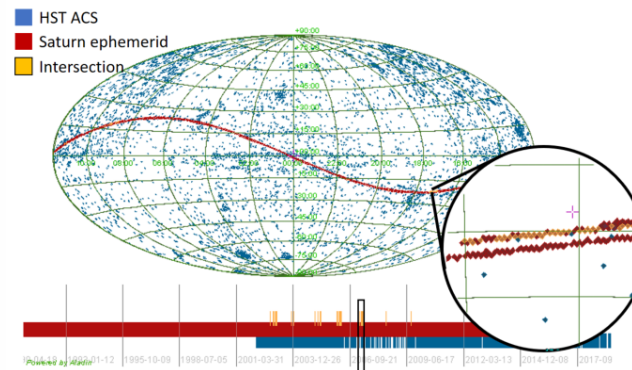**Intern: Flavian Theurel (UTBM)**

# ☐ F-MOC, towards a frequency MOC ?

- A MOC ? Quick reminder

  - Arbitrary coverage specification for Sky regions and/or time

  - very fast comparison mechanism provided

  - Based on discretization of space, time, dimensions

  - Based on a specific storage of the map coverage using predefined cell hierarchically

**See the IVOA MOC 2.0 document for details**

- **Standards**
  - MOC 1.0 => only spatial MOC
  - MOC 1.1 => + ASCII serialization
  - MOC 2.0 => Spatial + Temporal MOC

- **Data** from Oct 2021 to March 2023:
  - Spatial MOC: 23,832 -> 26,350
  - Temporal MOC: 1,212 -> 2,575
  - Spatio-temporal MOC: 1,045 -> 1,167

- **Tools** & **libraries**
  - MOCPy, MOC java
  - VO registry, MocServer, …
  - Aladin desktop, ESAsky, …





MOC: Multi-Order Coverage map
Version 2.0

IVOA Recommendation 2022-07-27

**P. Fernique, F.-X. Pineau, B. Cecconi (LESIA, Obs. Paris-Meudon), F. Bonnarel, D. Durand (CADC) et al Intern (2023) :**

# Space, Time… Energy ?

- The goal: reuse the same MOC principles to handle cove  on the electromagnetic axis

- Energy, Wavelength or frequency ? How to maps these values to a MOC ?

- Challenges around how to code/represent Energy (HEALPix indices for Space)

- Idea is to use frequencies

- Map values as a logarithmic expression, using the same principle as the coding of real numbers : mantissa and exponent

  - **52** bits for mantissa
  - **8** bits for exponent (not 11)
  - Save **4** bits for signature
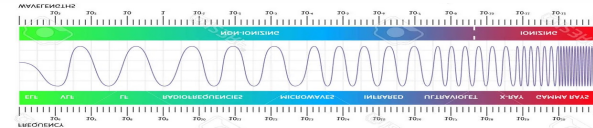
**Fast and positive evolution with Fast mapping, Tests, prototyping, Jupyter notebooks, etc.**
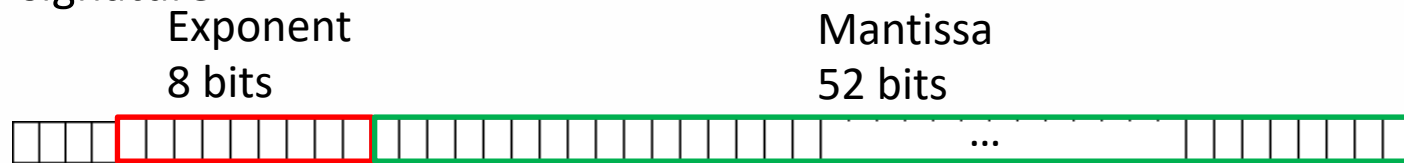
Exponent
8 bits

Mantissa
52 bits

…

P. Fernique, F.-X. Pineau, B. Cecconi **(LESIA, Obs. Paris-Meudon)**, F. Bonnarel, D. Durand **(CADC)** et al

# Astronomical tables annotations

- Model Instances in VOTables (MIVOT) defines a syntax to map VOTable data to any model serialized in VODML (VO Data Model Language). It's like a bridge between the data and the model.

- The internship objective was to generate examples of VOTable annotations in MIVOT

- And to create tools to make it easy.
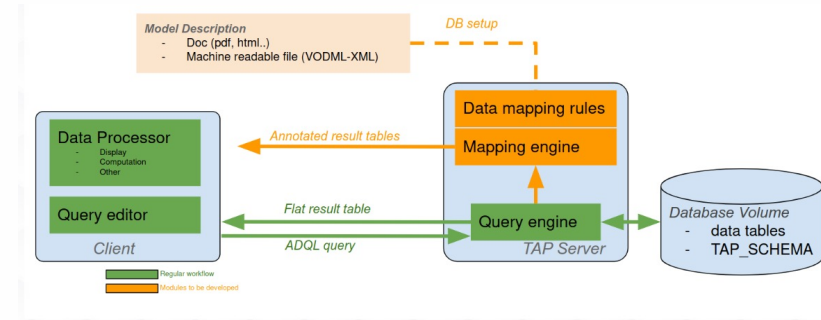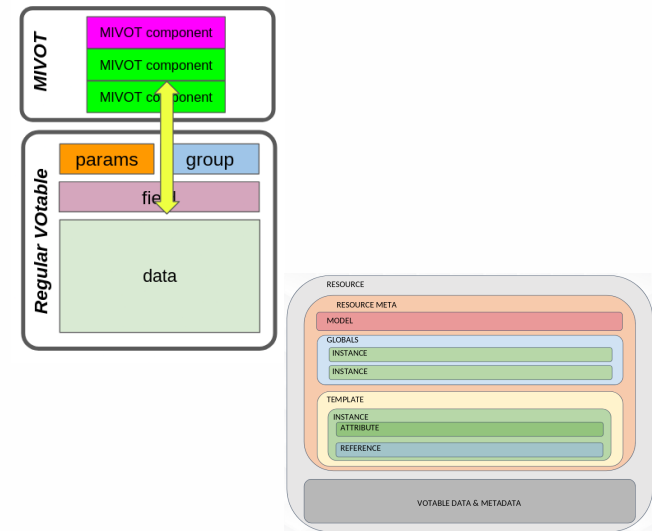


**M. Louys, L. Michel**
**Intern (2023) : Julien Abid (IUT Schuman Strasbourg)**

# ☐ Extension for Firefox

- **BCS** (Bibliography Center Supervisor) manages the papers import, storage and availability for the documentalistes.

- The CDS **XCDS** XML format is used for an efficient analysis with **DJIN**

- Problem: it is difficult for **Science** and **Nature**

- **Goal:** retrieve easily papers from Science and Nature

- 2 solutions (use the PDF version or load the HTML were not adapted.

- A third solution was to create and add an extension to the browser
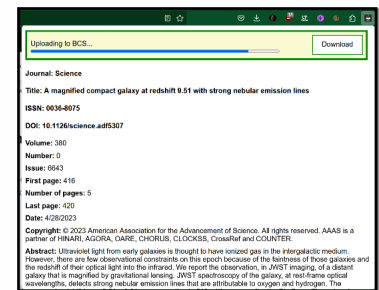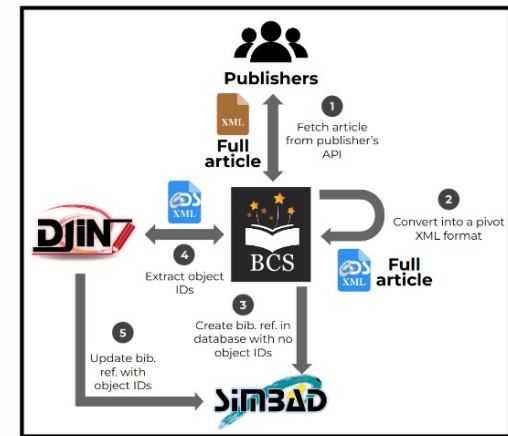




**Easy to use and time gain for the documentalistes**

**G. mantelet, A. Oberto**
**Intern (2023) : Théo Legros (IUT Schuman Strasbourg)**

# ☐ Examples of internships in VizieR

- Reuse of previous internship works
  - Foreign-Data-Wrapper in PostgreSQL, poster at ADASS 2023, (Adrien Vizier-Fontanesi)
  - UCD1+ resolver -> new version of UCD1+ in the VizieR renewal process (Louis Demange)
- 2023 Internship around the renewal of the VizieR photometry viewer based on Angular



**ANGULAR**



**G. Landais**
**Intern (2023) : Evguenia Sobine (IUT Schuman Strasbourg)**

# ☐ Future investigations

- A continuous R&D effort to provide an added value to the data access & presentation mechanisms.

- AI will be a major field of investigations.

- AstroInformatique (Ecole thématique CNRS) 2025 should be organised by the Observatory with the opportunity to train many local persons in AI (School main theme).

# Conclusion

- A coverage with various spin-offs:
  - improving the services, updating the staff skills
  - new (but not enough by itself !) technologies (e.g. AI).

- A team work since over 20 years at the Observatory level: contracts, workstations, presentation of the services and the professions to integrate quickly the newcomers, etc.

- Also a contribution to IT student Education and relations with the schools and their IT teachers and researchers.