# CSP Status Report

**Bruno Merín**
IVOA Committee on Science Priorities (CSP)
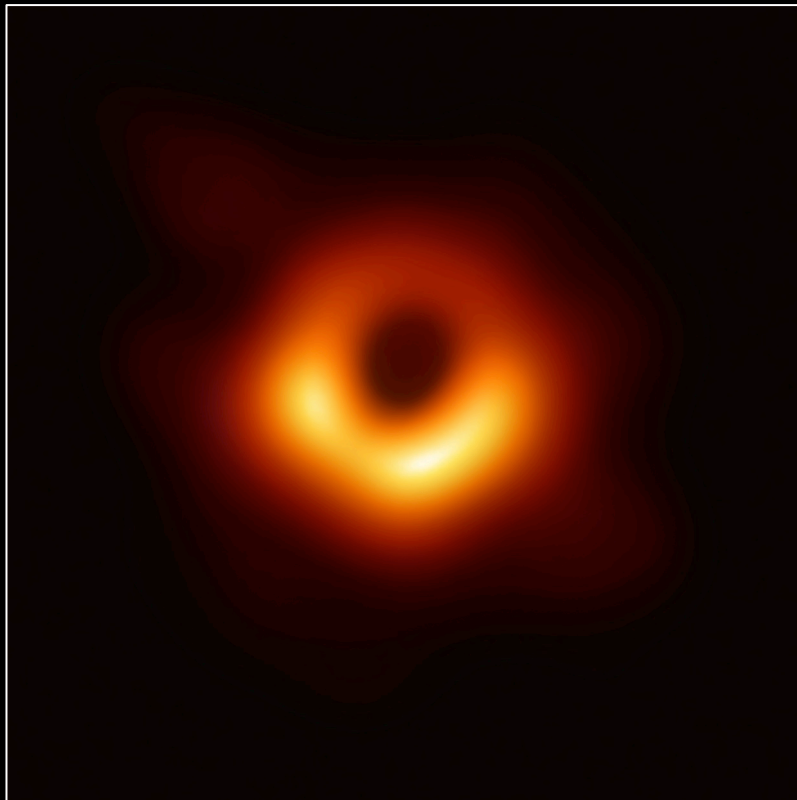http://wiki.ivoa.net/twiki/bin/view/IVOA/IvoaSciencePriorities

ESAC Science Data Centre (ESA), Madrid, Spain

IVOA Paris Interop, 13/05/2019

# Outline

1. Motivation

2. Scientific priorities

   1. Currently identified

   2. Upcoming

3. Final recommendations

European Space Agency

- 5 PBs of data, collected by plane
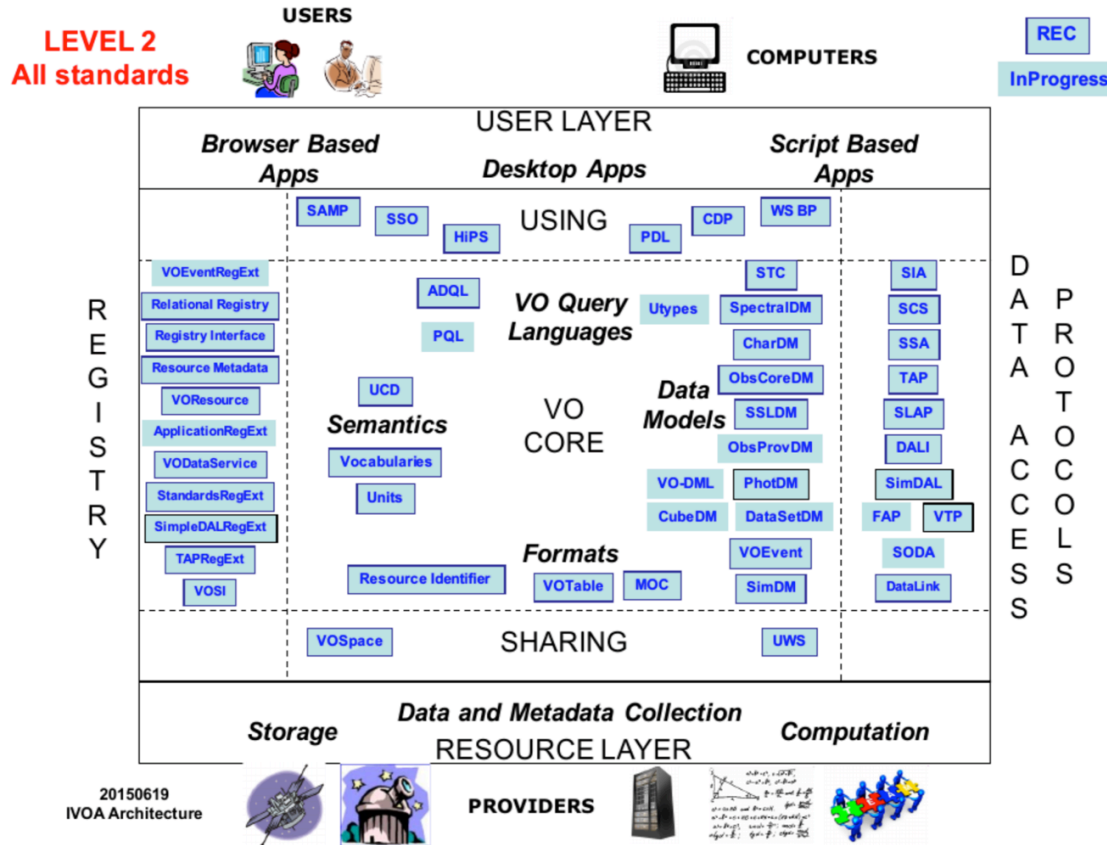- Months of data processing

# So what do scientists need?

1. *Visualization tools*

2. *Simple / easy* access to *reliable*, *relevant* and *big* data

3. *Fast* computation on new data

4. *Easy* comparison tools between data and models/theory

5. Data *completeness* and *consistency*

6. *Reproducible* data representation

European Space Agency

# What does the IVOA provide?



Technology

P. Dowler
TCG Report

European Space Agency

# What does the IVOA provide?

1. *Visualization tools* -> SAMP, HiPS, (T-)MOC

2. *Simple / easy* access to data
-> registry, ObsCore, SAMP, TAP, SODA, SIA/SSA, HiPS, (T-)MOC

3. *reliable* data
-> DataModels, Semantics

4. *relevant* data
-> ??? (missing, links to papers?, data ratings?)

# What does the IVOA provide?

*5.  Fast* computation on new data
-> GWS, Computing resources close to the data, VOSpace interface for distributed storage -> Focus Session on Big Data

*6.  Easy* comparison tools between data and models/theory
-> SimDAL, but models usually created by users..

7.  Data *completeness* and *consistency*
-> Registry complete and consistent? Glots?

*8.  Reproducible* data representation
-> Scripting interfaces, python? ADQL, TOPCAT -> Astropy Hackathon

European Space Agency

The best way to make progress is via a constant dialogue:

*science* ⟷ *technology*

European Space Agency

- **Time-domain astronomy**: light-curves -> see new proposed TIMESYS and T-MOC

- **Multi-dimensional data**: spectral or time cubes (sky + wavelength/frequency or sky + time)

  - **Python reference implementations** prioritized for major services?
  - Ways for accessing **large amounts of data** from future surveys?
  - Other growing areas/priorities?

User requirements defined in 2013:

- Data Discovery (Query) as a function of
  - RA, Dec
  - Frequency/wavelnegth
  - Polarization states
  - Spatial scale
  - Angular resolution
  - Integration time
  - Time of observations
  - Data Access
    - Download complete science data
    - Download simple cutouts

Simple cut-outs

Spatial : a circle (a coordinate + radius)

Energy : one interval (energy1 – energy2)

Time : one interval (time1 – time2)

Polarization : a list

Additional requirements for cut-outs:

Sum along any one or more axes

Re-bin in one or more axes

Multiply by a function

Other action on the data

# Example case for the Multi-Dimensional use-case

- October 2012: "Time-Domain Astronomy and Multi-Dimensional data were AGREED as priority areas for the IVOA" FM46S

- May 2013: Focus session on multi-dimensional data in Heidelberg

- May 2014: Focus session on multi-dimensional data at ESAC/ Madrid

- June 2015: DataLink 1.0 IVOA Recommendation

- December 2015: SIA 2.0 IVOA Recommendation 23 Dec. 2015

- May 2016: Focus Session on "Interoperability of data from major astronomical projects"

- May 2017:
  - SODA 1.0 IVOA Recommendation 17 May 2017
  - Obscore 1.1 IVOA Recommendation 9 May 2017
  - Announced the milestone – IVOA multi-d standards: first set of standards to address Discovery, Access and Simple cut-out of multi-D data

- July 2018: IVOA note on Feedback on the DAL protocols relevant to Multi-D standards

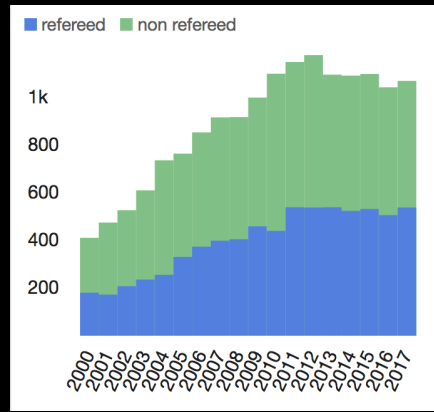- Feb. 2019: Reference implementation on ASTERICS DADI Technology Forum
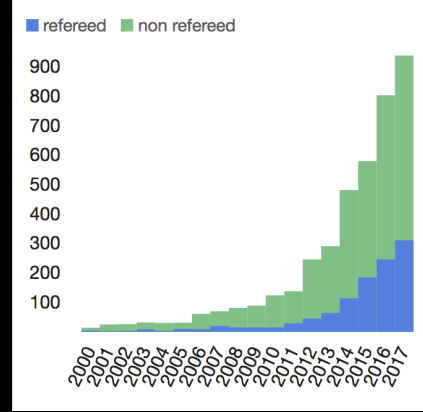
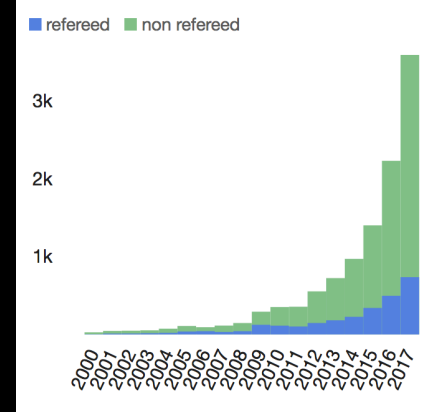# Upcoming scientific priorities at IVOA



Gravitational waves
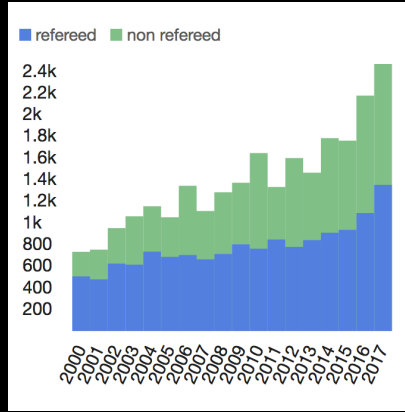
Multi-wavelength

Python

Machine learning

ADS-listed articles containing those key-words as a function of time

# IVOA – Astropy hackathon

- Several core developers of the Astropy and Astroquery groups will attend part of the IVOA interop and will run a two-days hackathon/sprint with a few Astropy contributors from IVOA in a parallel session (Tue and Wed).

- Goals:

  - Explore how best to build cooperation between the IVOA and Astropy

  - Discuss role of PyVO

  - Deliver new VO features !

# Focus session on Big Data Challenges in Astronomy

**Tuesday, May 14, 16:00--17:30, Salle Le Verrier**

| Speaker | Title | Presentation | Time |
|---|---|---|---|
| Bruno Merín | Motivation for the Focus Session | | 2 |
| Gregory Dubois-Felsmann | LSST data exploitation plans | | 5+5 |
| Tom Donaldson | Pan-STARRS, WFIRST and TESS data exploitation plans | | 5+5 |
| Juan González | Gaia data exploitation plans | | 5+5 |
| Jesús Salgado | Euclid data exploitation plans | | 5+5 |
| Séverin Gaudet | SKA RC data exploitation plans | | 5+5 |
| | Open discussion on challenges and opportunities | | 28 |

Questions for large surveys :

- Describe the data volumes and types of data expected from the mission/survey.
- Describe your data dissemination/exploitation plan for users.
- Are you looking at sending data to users or looking at a code to the data approaches?
- How would you cross-correlate data with different surveys?
- How and where does the IVOA fit into your plans?

# The IVOA needs you

- We need active and enthusiastic scientists at the Committee of Science Priorities!!

- Talk to us if you are interested!!

European Space Agency

# Final recommendations from the CSP

- Always ask the question: how is the user going to use this?

- Always follow the user workflow to the paper and keep the big picture (is provenance clear? Can I explain/make a plot of this?)

- Connect to the future generation of users where they are: e.g. python, github, open source projects, social media, online open fora, connected to new big astronomy projects, using mobile devices and expecting quick answers

European Space Agency

# Thanks!

**Committee on Science Priorities : csp@ivoa.net**

European Space Agency