The VO: A Powerful Tool for Global Astronomy

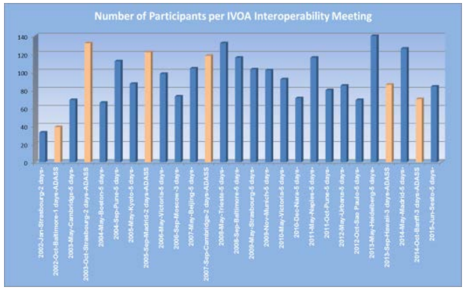
Christophe Arviset,
Head of ESAC Science Data Centre, IVOA Chair

26/10/2015

Issue/Revision: 1.0
Reference: VO: A powerful tool for astronomy
Status: Draft

IVOA Today

1. Created in 2002, IVOA today has 20 diverse member projects

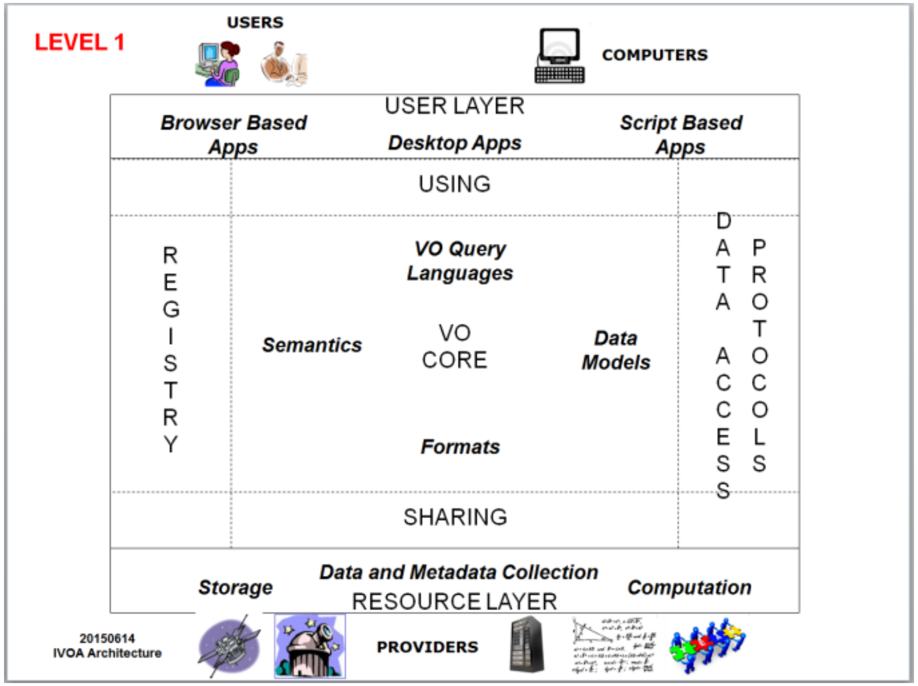


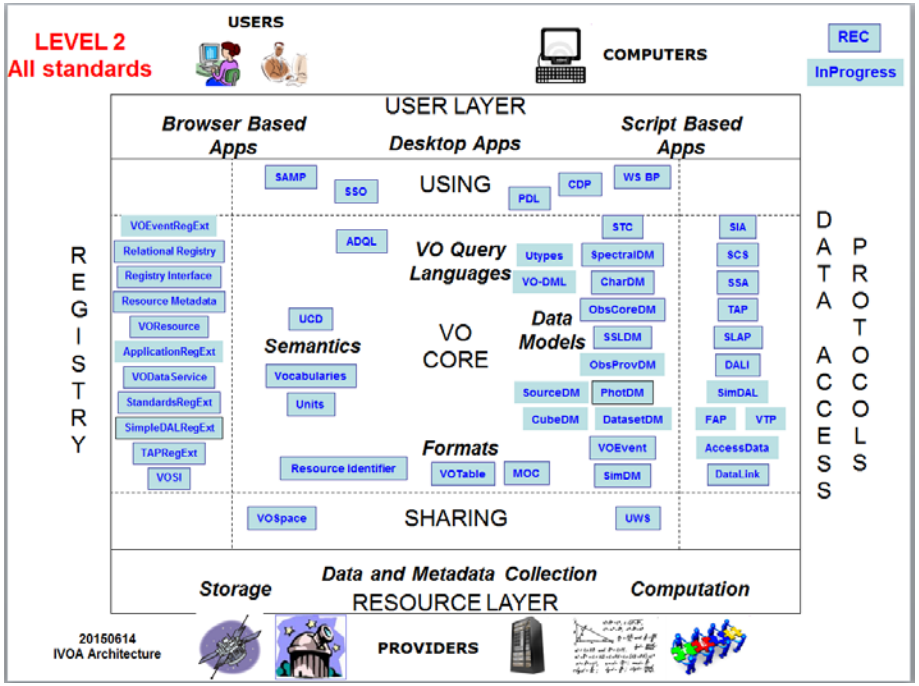
2. 2 well attended interoperability meetings per year
3. Technical Coordination Group
   - 6 active Working Groups (Applications, Data Access Layer, Data Model, Grid and Web Services, Registry and Semantics)
   - 7 Interest Groups (Data Curation and Preservation, Education, Theory, Time Domain, Operations, Knowledge Discovery in Databases)
4. Members have come and gone, some projects have persisted and some others have stopped
5. Hard lessons learnt on management and developing standards
6. IVOA has survived because VO is a good idea

IVOA Successes - 1

1. International collaboration of many VO projects worldwide

2. Stable IVOA Architecture, with well established interoperability standards
   - ➢ tables, images, spectra, registries

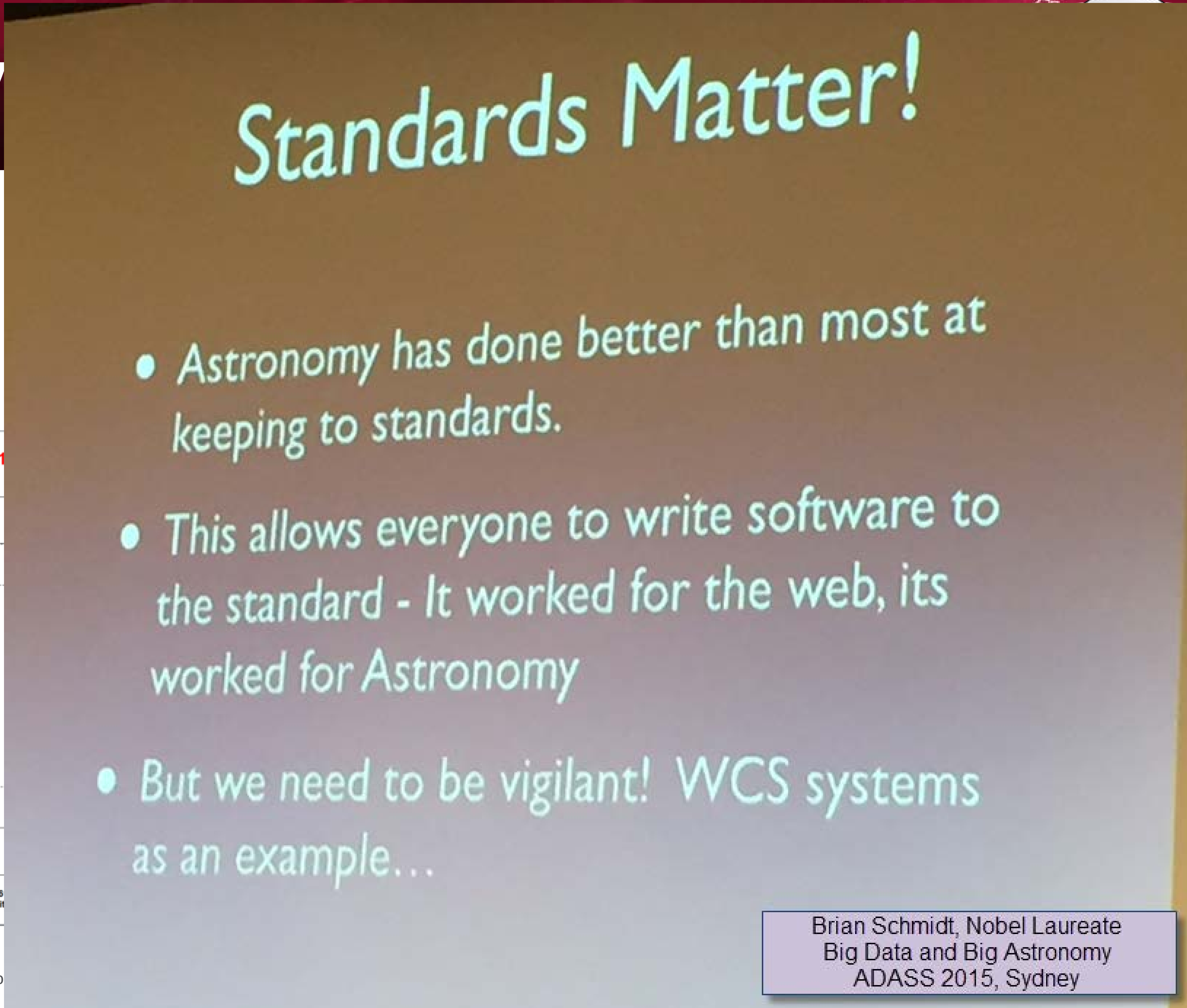




IV

1.

2.



Brian Schmidt, Nobel Laureate
Big Data and Big Astronomy
ADASS 2015, Sydney

LEVEL 1

REC

InProgress

DATA ACCESS

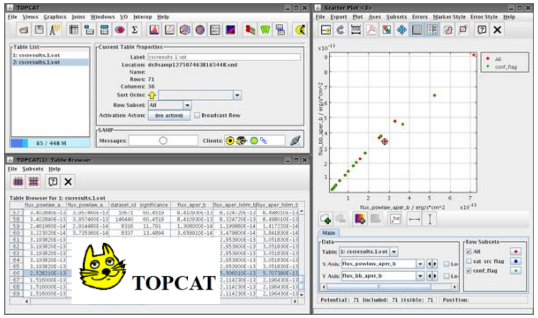
PROTOCOLS

201506
IVOA Archit

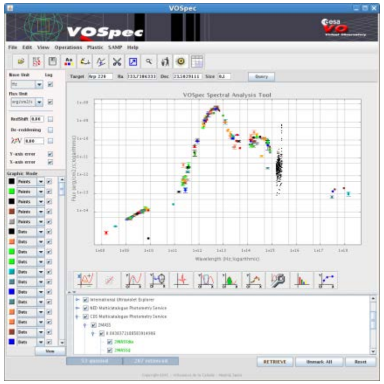
IVOA Successes - 1

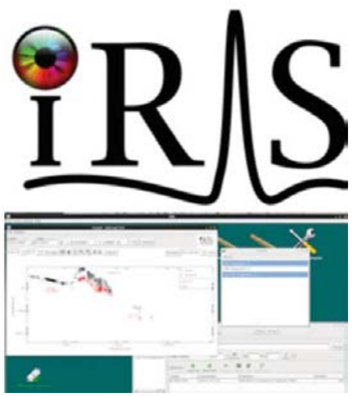
1. International collaboration of many VO projects worldwide

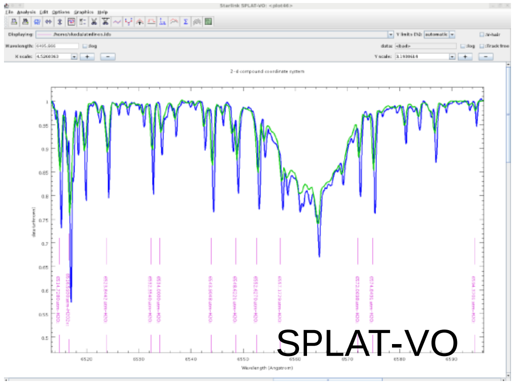
2. Stable IVOA Architecture, with well established interoperability standards
   - tables, images, spectra registries

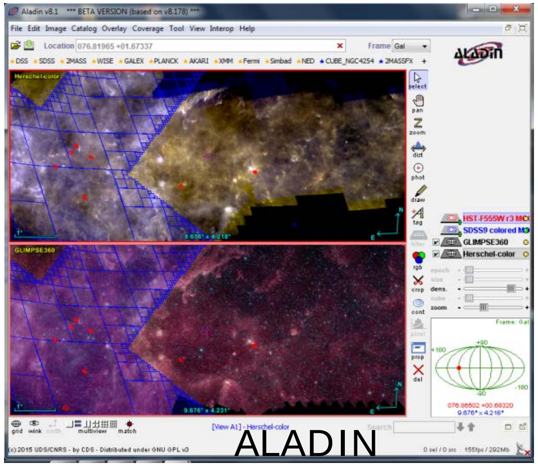


3. VO *interoperable* applications
   - Topcat, Aladin, VOSpec, SPLAT-VO, Iris, ...





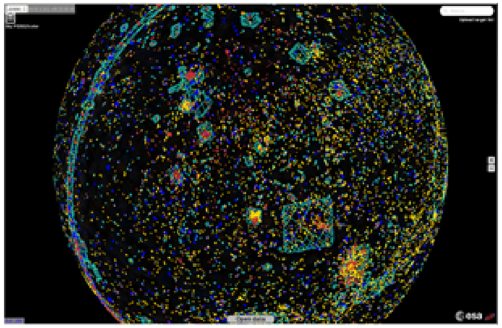




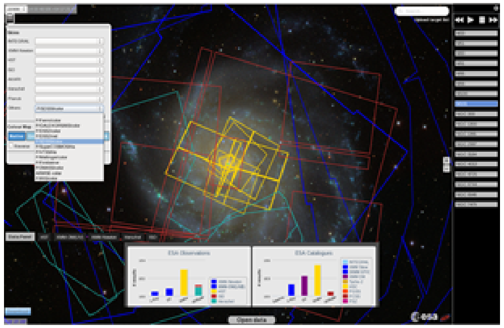
IVOA Successes - 1

ESA Sky

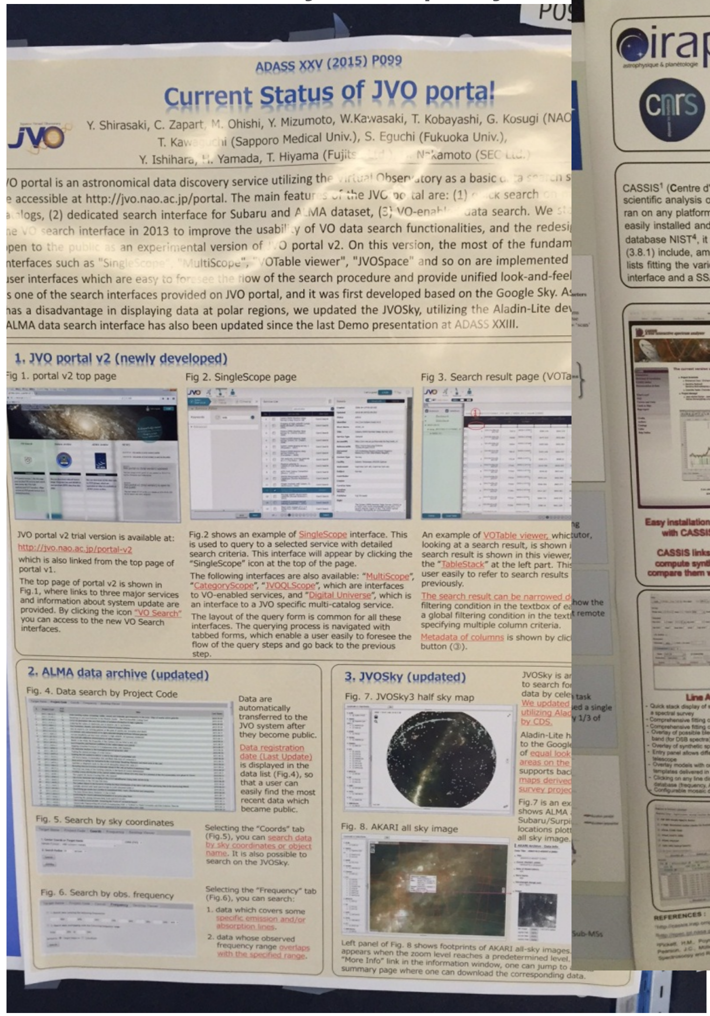
- science driven discovery portal for all ESA Astronomy missions
- facilitates data discovery and archival science for ALL users
- all-sky multi-resolution HiPS maps of full mission datasets

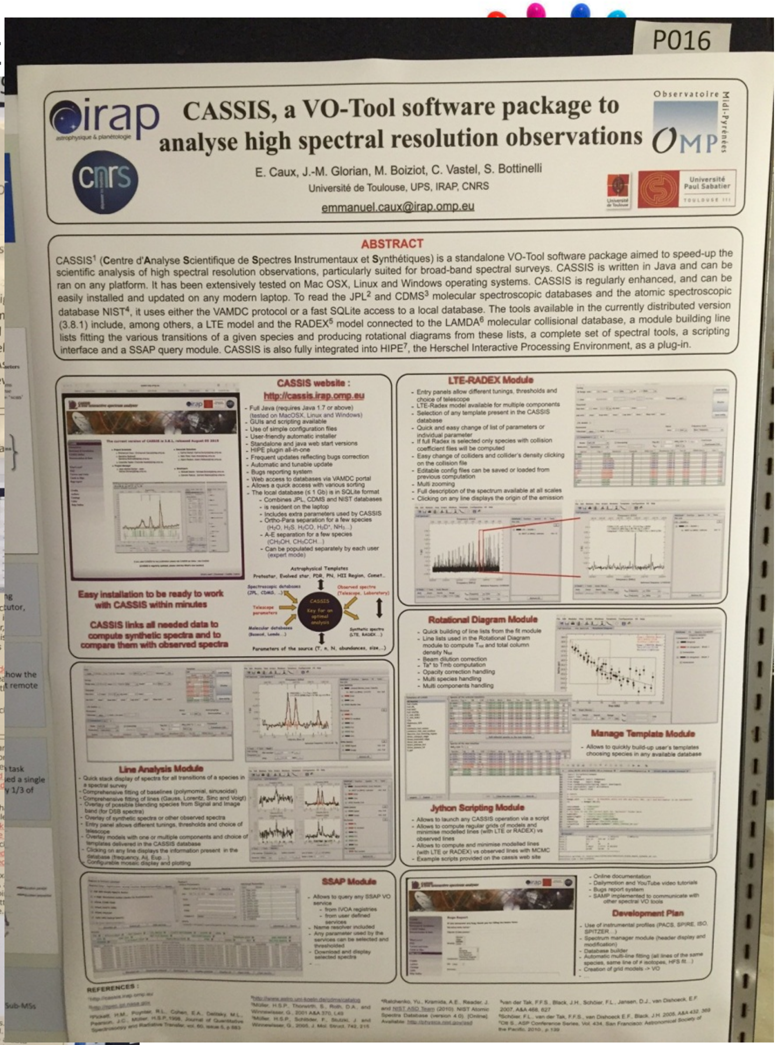


- detailed geometrical footprints to connect the all-sky mosaics to individual observations
- direct access to the underlying mission-specific science archives
- seamless access to mission source catalogues



ration of many VO project





IVOA Successes - 1

1. International collaboration of many VO projects worldwide

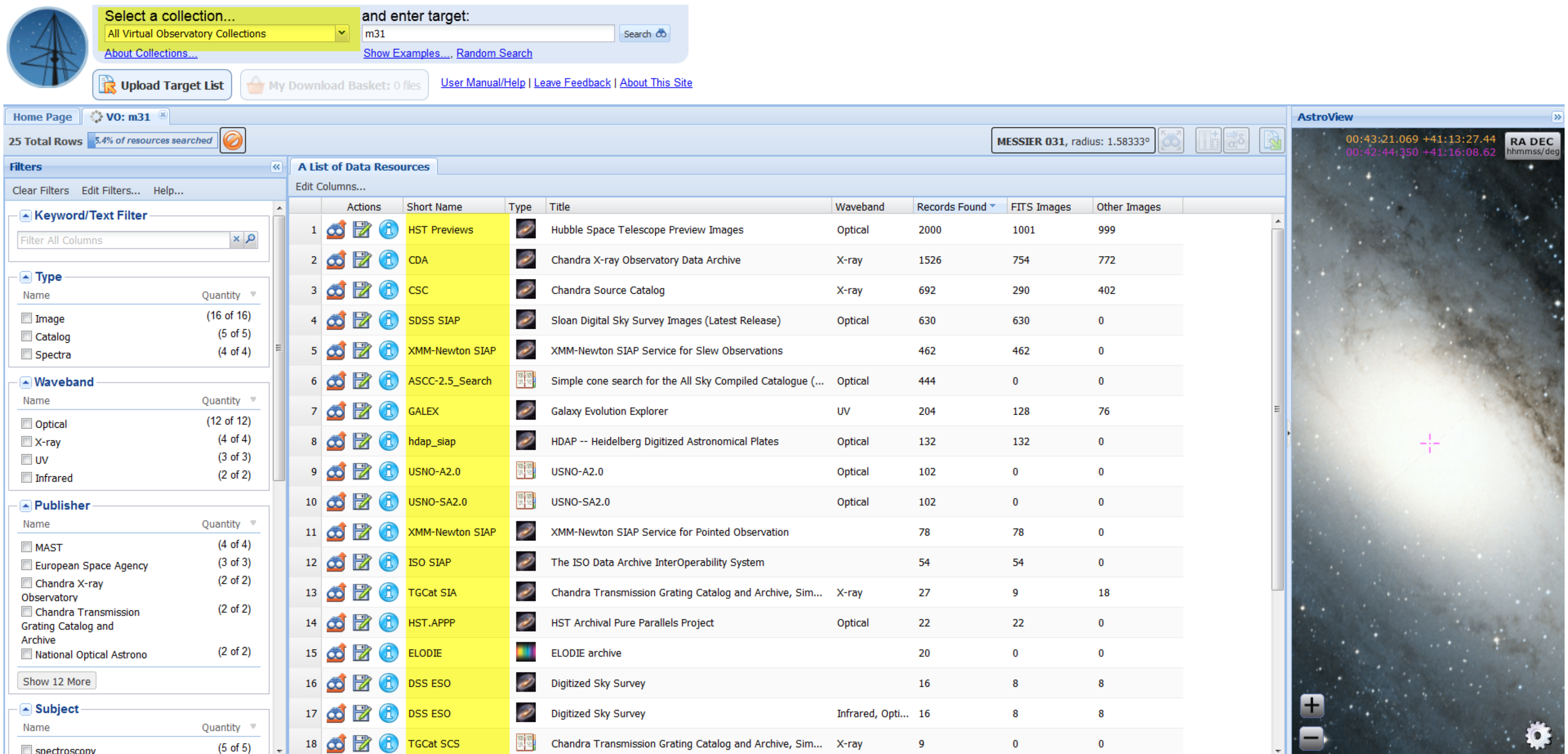
2. Stable IVOA Architecture, with well established interoperability standards
   - tables, images, spectra registries

3. VO applications
   - Topcat, Aladin, VOSpec, SPLAT-VO, Iris, ...
   - ESASky, JVO Portal, Cassis, ...

4. VO recognized and supported as an e-Infrastructure
   - E.g. ASTRONET European Infrastructure Roadmap
   - NASA support to US-VOA
   - Open and shared software and infrastructure components
   - Registries, TAP libraries, VOTable parsers, Data publishing SW
   - Embeddable interface components coming (e.g. Aladin Lite)

IVOA Successes - 2

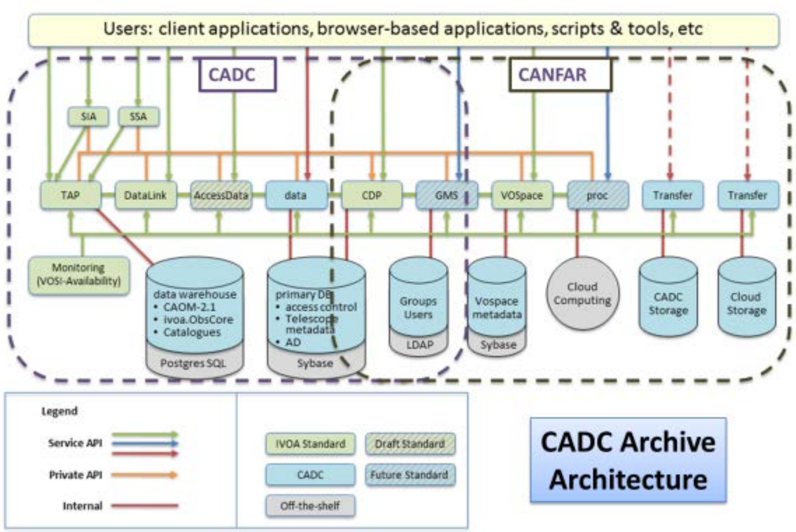
1. Major astronomical data collections accessible through the VO
    - CDS, CADC, MAST, ESA, Chandra, …

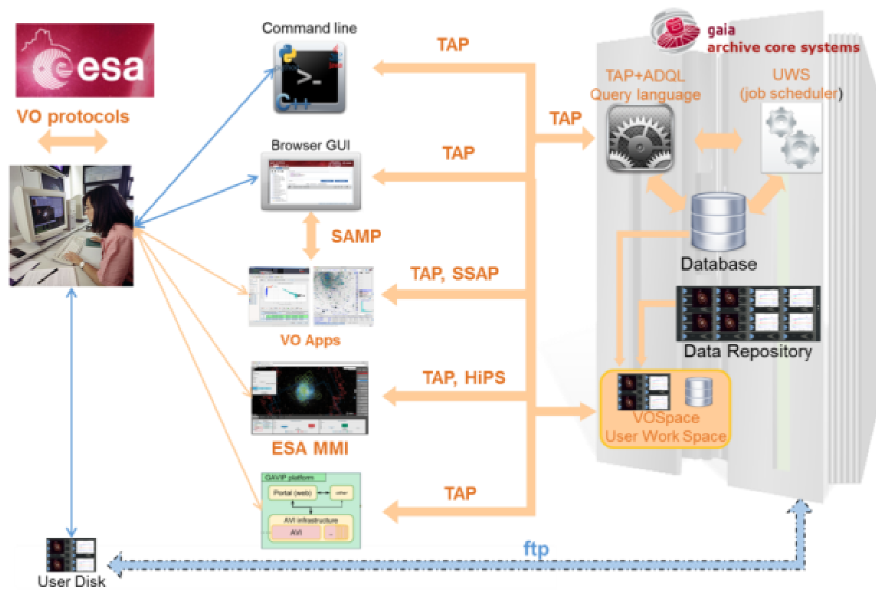


IVOA Successes - 2

1. Major astronomical data collections accessible through the VO
   - CDS, CADC, MAST, ESA, …

2. VO being used to build new astronomical data infrastructure
   - CADC, Gaia, Euclid, …





IV

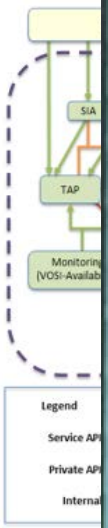
1.

2.

SkyMapper Database

- SDSS-depth 5-year survey 20,000 sq deg
- Volume
  - 300,000 images
  - 5 Giga-rows of measurements × 100 features
  - Tables with images, measurements, distilled objects

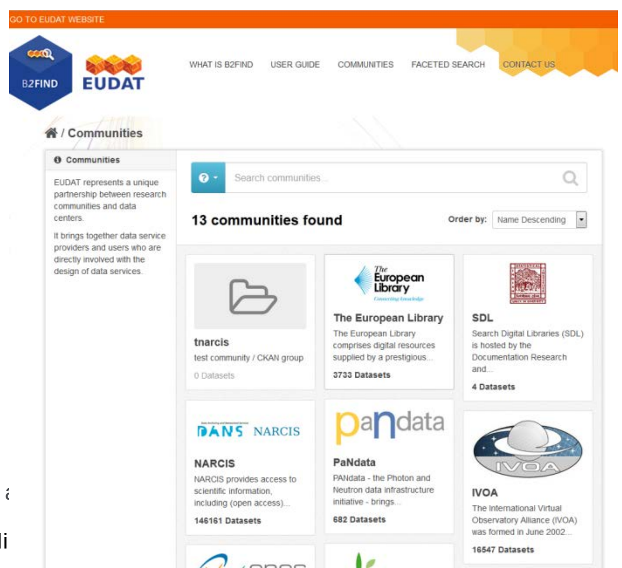
- Query functions & IVOA standards
  - ADQL catalogue query – Table Access Protocol TAP
  - Cone search – Cone search
  - Image cutouts – Simple Image Access Protocol SIAP
- Service discoverable in TopCat
- Web access wrapped around backend functions

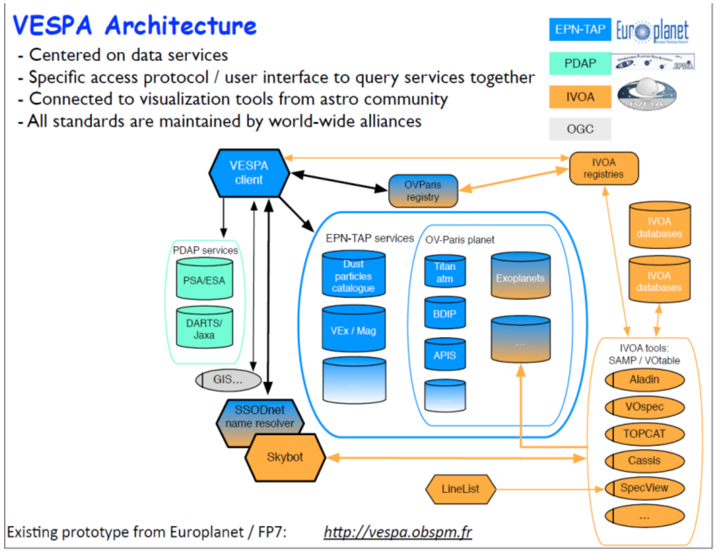


All Sky VO, Australia
ADASS 2015, Sydney

IVOA Successes - 2

1. Major astronomical data collections accessible through the VO
   - CDS, CADC, MAST, ESA, …
2. VO being used to build new astronomical data infrastructure
   - CADC, Gaia, Euclid, …
3. VO re-used by neighbour disciplines
   - Planetary Science,
   - VAMDC,
   - EUDAT





IVOA Problem Areas

1. Unrealistic initial expectations
   - Big ideas were needed to get off the ground
   - Some aspects were "over-sold"

2. Sometimes VO perceived as a closed-shop
   - VO standards seen as too complex
   - Difficulty to implement rich VO services
3. Long time to deliver VO standards

4. Wrong perception, many people initially thought VO was a killer application, but
   - IVOA defines VO "ecosystem" and interoperability standards
   - Astronomy projects and data services build VO services and VO applications

VO is a data management interoperability infrastructure

rather than an astronomy "app"

VO Challenges – take up

1. Engagement of big projects and how the VO can help them ?

   - How to capture their requirements into IVOA standards development process ?
   - How to adapt to their development timelines and constraints ?
   - How the VO can help with “Big Data” and bring the code to the data ?

2. Facilitate and improve VO take-up by Data Centres
   - Two different models of VO implementations (VO layer or VO built-in)
   - VO publishing tools for small data centres with little IT expertise
   - VO software libraries for bigger data centres with more IT expertise
   - Simple services (S*AP) vs advanced services (TAP, DM)

Need to make big projects and data centres

“participants” not “customers”

VO Challenges – Operations

1. VO is in use, but how to see it ?
2. How to define VO success (or failure) metrics ?
   - Counting scientific publications ?
     - But doesn´t capture VO use...
   - VO services access statistics ?
     - But not uniformly collectable and comparable...

3. VO Services Operations
   - Reliability of existing VO Ecosystem
   - Hundreds of VO services in the IVOA Registries...
     - Compliance of VO services ?, VO Services validators
   - IVOA to become more active in this

Is the VO success to be “invisible” ?

VO for Science

1. VO very good for data discovery and quick exploration of data (e.g. ObsCoreDM, HiPS, MOC)
   - From various data sources and data centres
   - Then push to VO applications for display and analysis (SAMP)
   - Well defined for table and images, spectra not yet reached full exploitation potential

2. Getting science use cases from scientists
   - IVOA Committee on Science Priorities
     - ObsCore DM, SED, Multi Dimensional data, TimeDomain
   - IVOA to define standards and infrastructure to enable implementation of science use cases

The goal of the VO is to make access to and use of data easier

to enable science !

Conclusions

1. *"The VO is **not** a magic solution to all astronomy data management challenges but it offers a **powerful tool** and **useful solutions** to some of them"*
   - Interoperability amongst datasets
     - and VO science applications will be the key
   - IVOA Standards to help building archive data centre infrastructure
   - Need to address "bring and run code next to the data"

2. *"If one wants to **take something out** of the IVOA, one needs to **bring something in**"*
   a. Need to convince big projects to participate in standards development so they can better fit their needs
   b. And IVOA needs to go faster so we can meet projects deadlines

Special Thanks to my co-authors

1. Allen, Mark, CDS
2. Aloisi, Alessandra, STScI / MAST
3. Berriman, Bruce, IPAC
4. Boisson, Catherine, Observatoire de Paris / CTA
5. Cecconi, Baptiste, Observatoire de Paris / Europlanet/VESPA
6. Ciardi, David, NASA Exoplanet Science Institute, IPAC
7. Evans, Janet, SAO / CXC
8. Fabbiano, Giuseppina, SAO / CXC
9. Genova, Francoise, CDS
10. Groom, Steve, IRSA / IPAC
11. Jenness, Tim, LSST
12. Mann, Bob, Wide-Field Astronomy Unit, University of Edinburgh
13. McGlynn, Tom, NASA / HEASARC
14. O'Mullane, Wil, ESA-ESAC / Gaia
15. Schade, David, CADC
16. Stoehr, Felix, ESO / ALMA
17. Zacchei, Andrea, INAF-OATs / Euclid

We want the VO because…

(CADC) “*it makes implementation easier by leveraging the collective inputs of a large community through the use of the VO standards.”*

(CDS) *“it helps astronomers do science with our services and it enables the services to interoperate with others”*

(CXC) *“VO cone search is our most popular interface to the Chandra Source Catalog”*

(ESA) *“it helps us built up new generation of archives (e.g. Gaia, Euclid) and makes our data reachable via other interfaces”*

(ESO / ALMA) *"VO technology (ADQL, ObsCoreDM) is at the core of our archive interfaces."*

(LSST) *“we want our alert stream to be usable by the broadest user community through VOEvent”*

(STScI / MAST) *“it helps our users to tap into new scientific opportunities by cross-correlating data collections spread throughout the world!.* The era of "Global Astronomy" has become a reality!“

(WFAU) *“it enables our users to do multi-wavelength astronomy using our data”*