European Summer Campus, 2011

Introduction to the Virtual Observatory

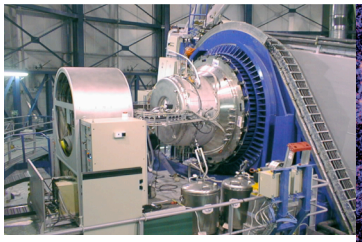
C. Bot & P. Fernique

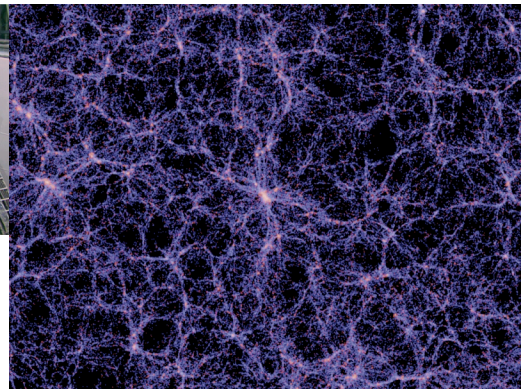
Centre de Données de Strasbourg

Observatoire
Astronomique
de Strasbourg

Research in astronomy today is based on:

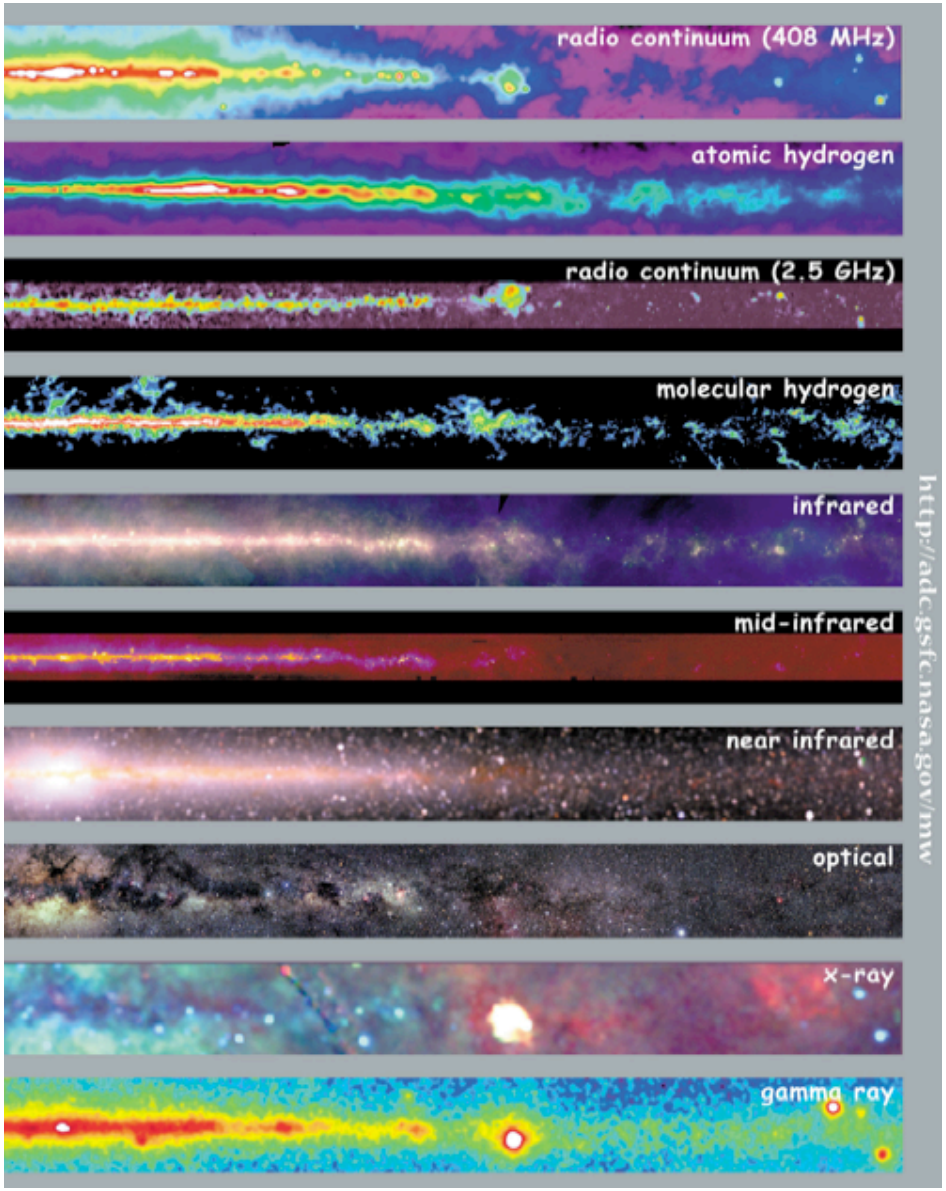
- Telescopes and their instrument
  - spatial or ground based
  - covering the whole range of the electromagnetic spectrum
- Data reduction or data analysis softwares
  - depending on the instrument
- Data archives
  - Preserve, maintain and add value to the data
- Theory, models and numerical simulations, laboratory experiments
- Journals -> publications





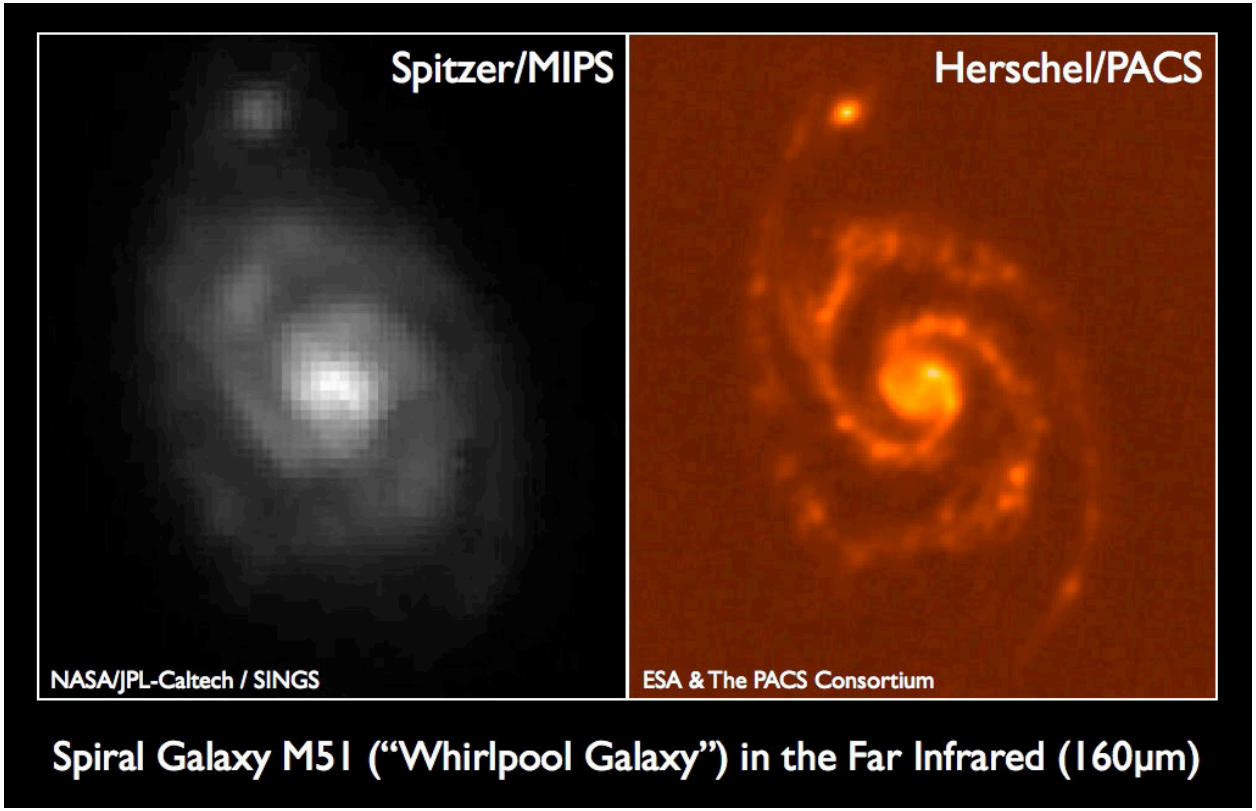
Evolution of astronomy

- Multi-λ astronomy
  - Combine data with different types, different formats…
  - Different data archives, different data access
  - Different data analysis softwares, different techniques



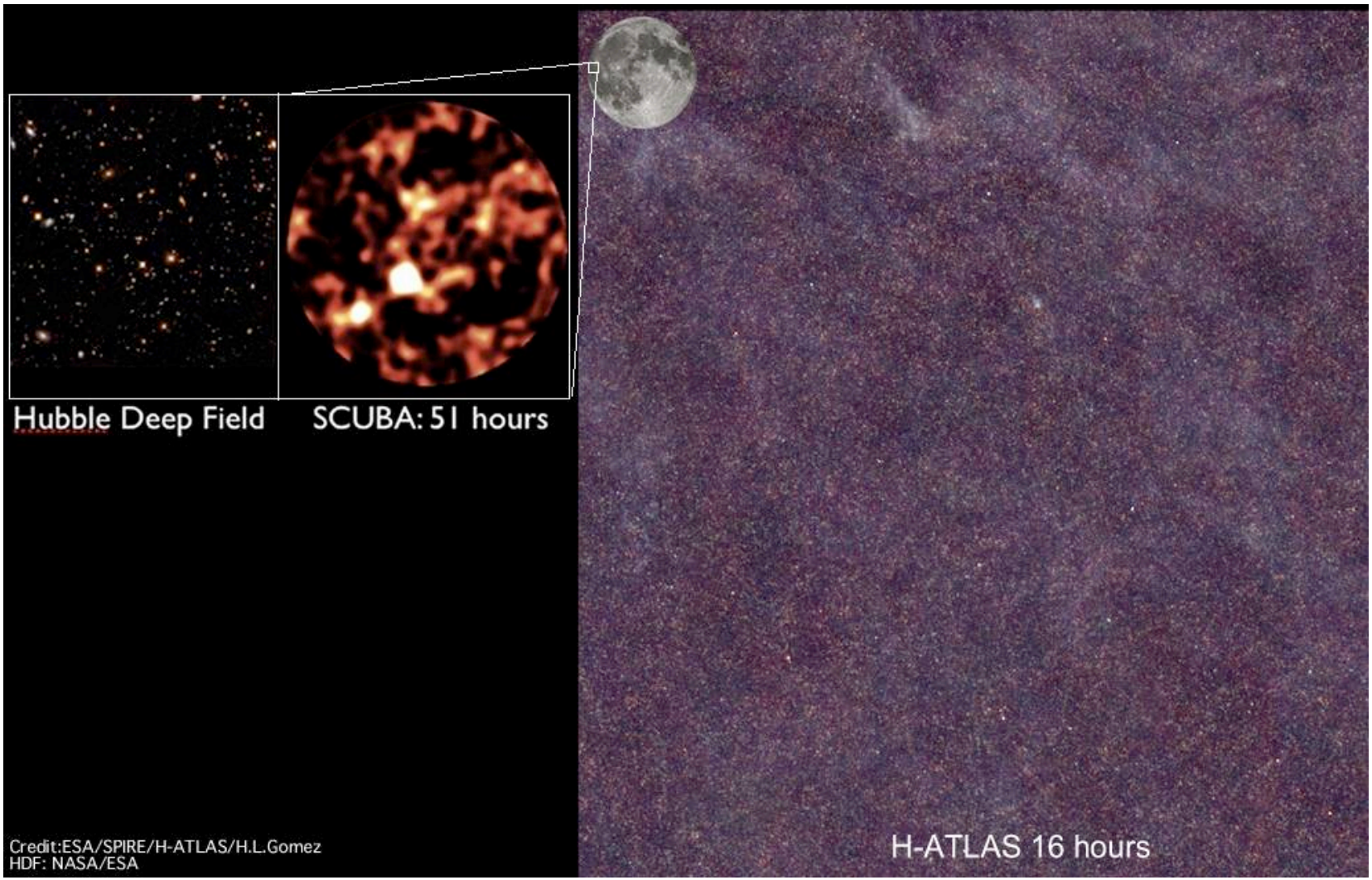
Evolution of astronomy

- More and more resolution



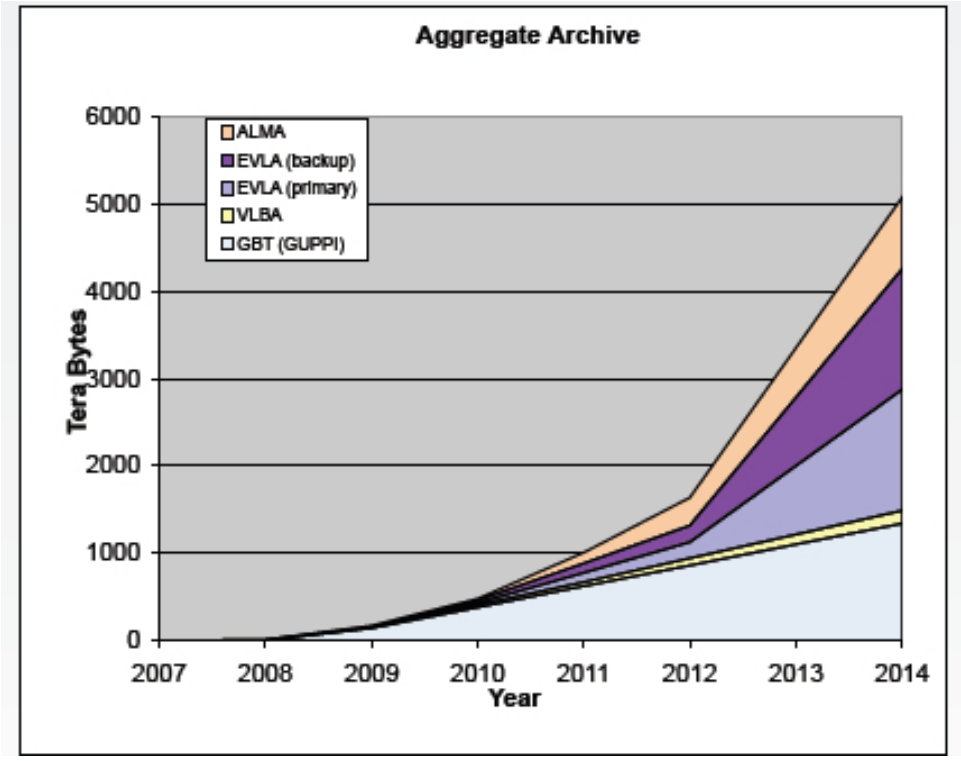
Evolution of astronomy

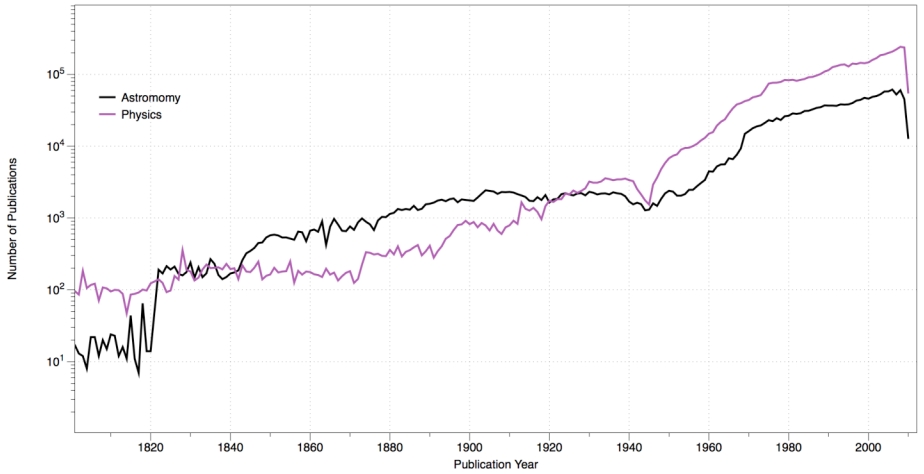
- Large surveys, deep fields



Evolution of astronomy

- Avalanche of data quantity!





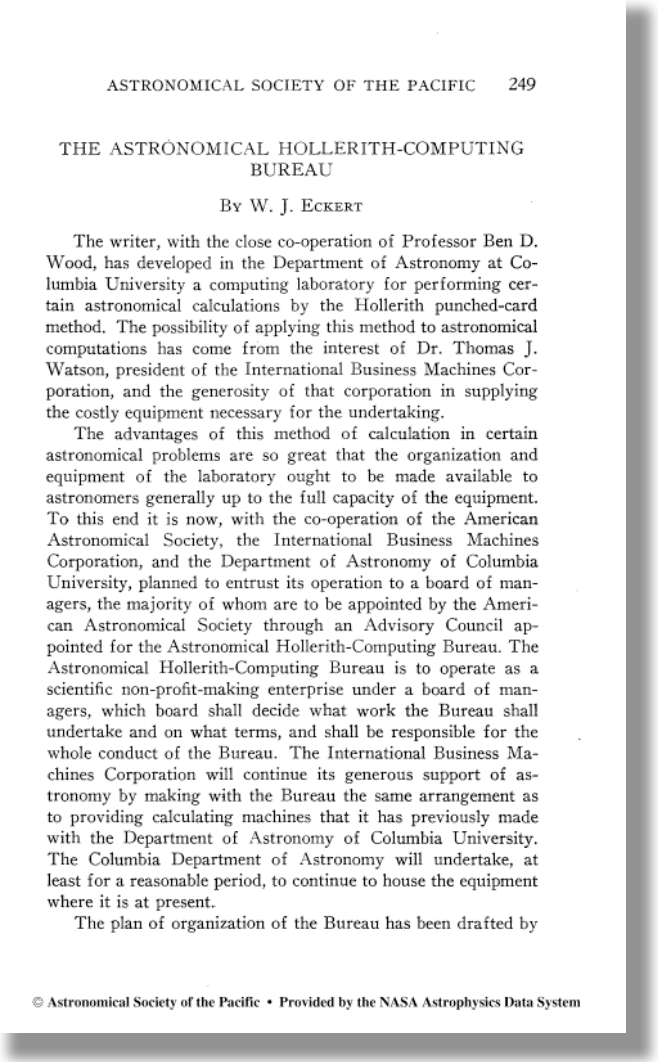
Data avalanche

- *Very large surveys:*100 million of sources, less than 3000 sources per night ⇒ > 100 nights to identify them

- *Very large data collections*: dowloading and analysing the data on personnal computers becomes problematic, or even impossible. For example: to download the data from the Sloan Digital Sky Survey (SDSS) DR6 :
  - ✓ images (10 Terabytes) ⇒ ~ 3 months at 10 Mbps
  - ✓ catalogues (2 Terabytes) ⇒ ~ 3 weeks
  - ✓ On DVDs ⇒ you will need ~ 2100 of them
- How do you do *data analysis*?? (similar sizes for MACHO, 2MASS …)

astronomy & computer science

- Long term relationship
  - 1937: collaboration IBM & Columbia university to compute 40 000 stellar locations (YALE catalog)
  - 1988: first TCP/IP transatlantic connection for the 20th UAI congress to access Simbad (CDS)
- No commercial value of the data
  - Free, open access
  - Simplifies collaborations, data exchange…
- Data volume always at the limit of the technology



THE ASTRONOMICAL HOLLERITH-COMPUTING BUREAU

By W. J. Eckert

The writer, with the close co-operation of Professor Ben D. Wood, has developed in the Department of Astronomy at Columbia University a computing laboratory for performing certain astronomical calculations by the Hollerith punched-card method. The possibility of applying this method to astronomical computations has come from the interest of Dr. Thomas J. Watson, president of the International Business Machines Corporation, and the generosity of that corporation in supplying the costly equipment necessary for the undertaking.

The advantages of this method of calculation in certain astronomical problems are so great that the organization and equipment of the laboratory ought to be made available to astronomers generally up to the full capacity of the equipment. To this end it is now, with the co-operation of the American Astronomical Society, the International Business Machines Corporation, and the Department of Astronomy of Columbia University, planned to entrust its operation to a board of managers, the majority of whom are to be appointed by the American Astronomical Society through an Advisory Council appointed for the Astronomical Hollerith-Computing Bureau. The Astronomical Hollerith-Computing Bureau is to operate as a scientific non-profit-making enterprise under a board of managers, which board shall decide what work the Bureau shall undertake and on what terms, and shall be responsible for the whole conduct of the Bureau. The International Business Machines Corporation will continue its generous support of astronomy by making with the Bureau the same arrangement as to providing calculating machines that it has previously made with the Department of Astronomy of Columbia University. The Columbia Department of Astronomy will undertake, at least for a reasonable period, to continue to house the equipment where it is at present.

The plan of organization of the Bureau has been drafted by

Publications of the Astronomical Society of the Pacific,
Vol. 49, No. 291, p.249

Courtesy of P. Quinn

So, what is the « Virtual Observatory »?

The Virtual Observatory

- Web: all documents on a personnal computer
- VO: *all astronomical data bases on a personnal computer*
- VO ⇒ democratisation of astronomy!

- What the VO is not:
  - A centralized data base that contains all astronomical data
  - A monolithic software
  - A peer-to-peer system

- All this requires that the different parties speak the same language ⇒ VO standards VO and protocols are defined and adopted by the IVOA (International Virtual Observatory Alliance) that includes 16 projects across the world

- But the OV layer remains transparent to the users

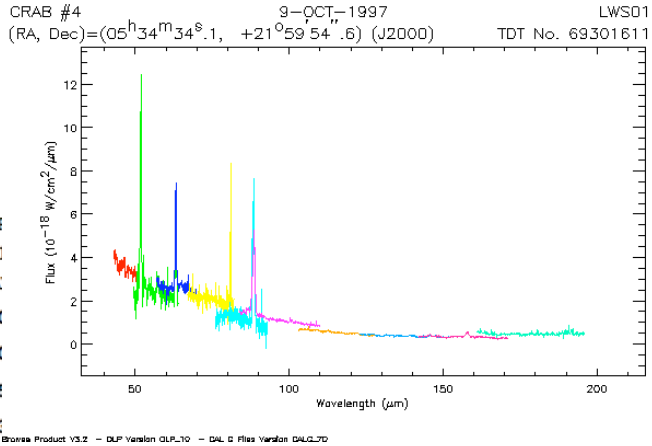

Member Organizations
Europe
Chine
Australie
Inde
Canada
UK
Russie
Espagne
USA
Italie
Corée
Arménie
Japon
Hongrie
Allemagne
France
IVOA
EURO VO
Aus-VO
Astro Grid
China-VO
vo-i
CVO
SVO
NVO
NATIONAL VIRTUAL OBSERVATORY
VObs.it
ArVO
JVO
JAPANESE VIRTUAL OBSERVATORY
HVO
Hungarian Virtual Observatory
GERMAN ASTROPHYSICAL
GAVO
VIRTUAL OBSERVATORY
France

The Virtual Observatory

- The Virtual Observatory (VO) is an inovating and evolving system, that enables astronomers to:
  - Benefit from the ever increasing quantity of astronomical data
  - Query multiple data centers in a transparent way and to take the best out of astronomical data
  - Remotely compute and perform data analysis
  - Triger *new science*

The Virtual Observatory

- Astronomers questions the VO can answer:
  - Are there data that…
  - Where can I find data/tools that…
  - What is the format of…
  - What is the content of…

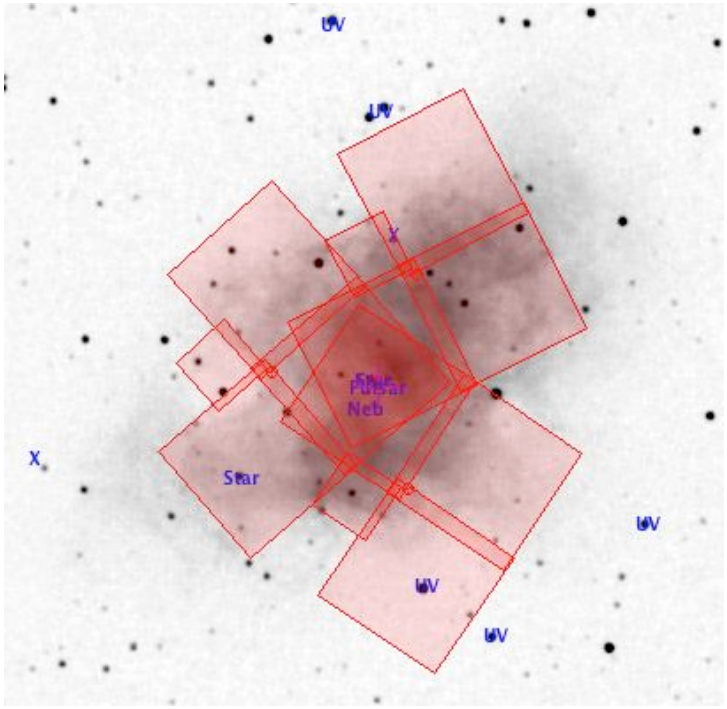


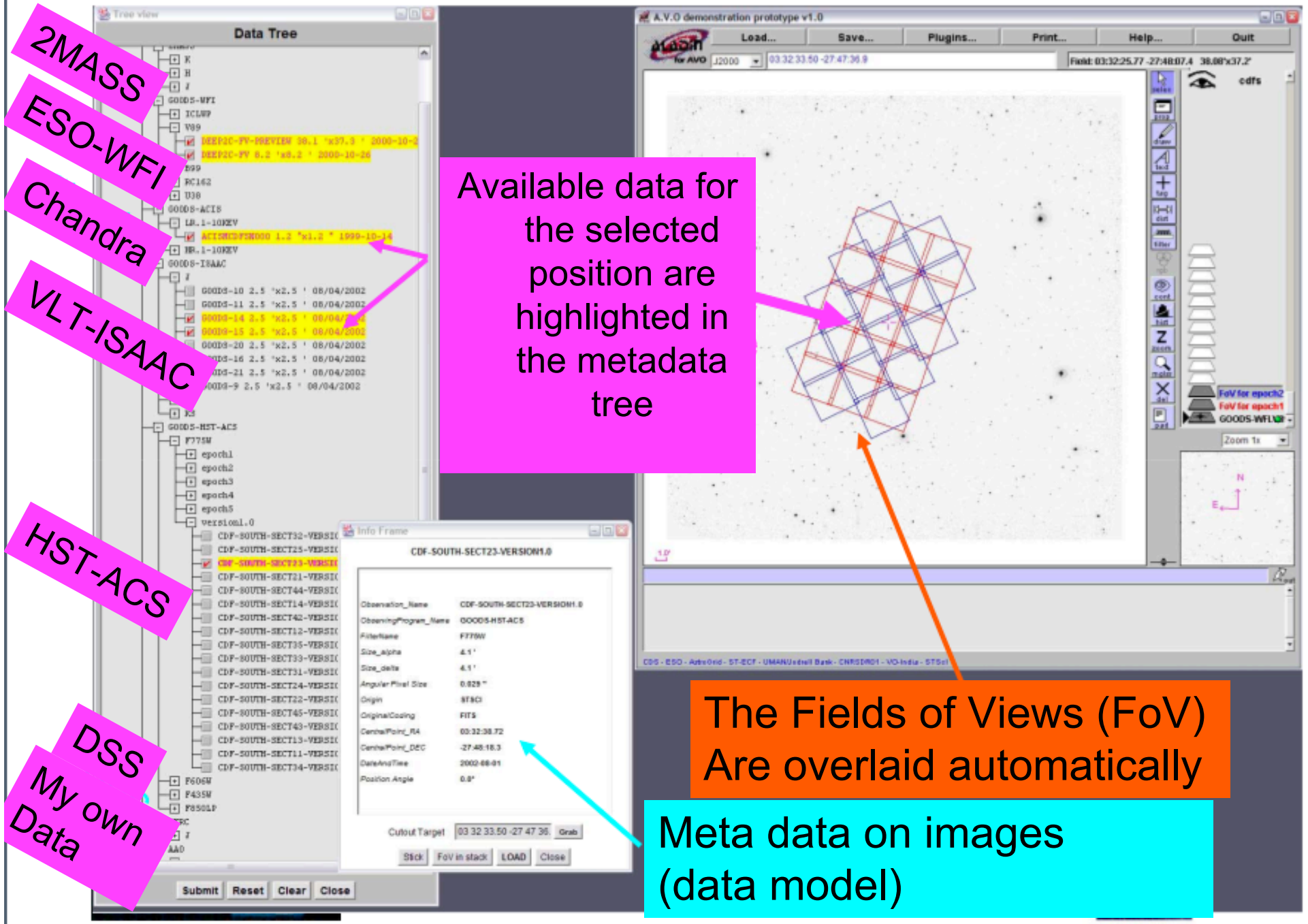
Basic data :

V* CM Tau -- Pulsar

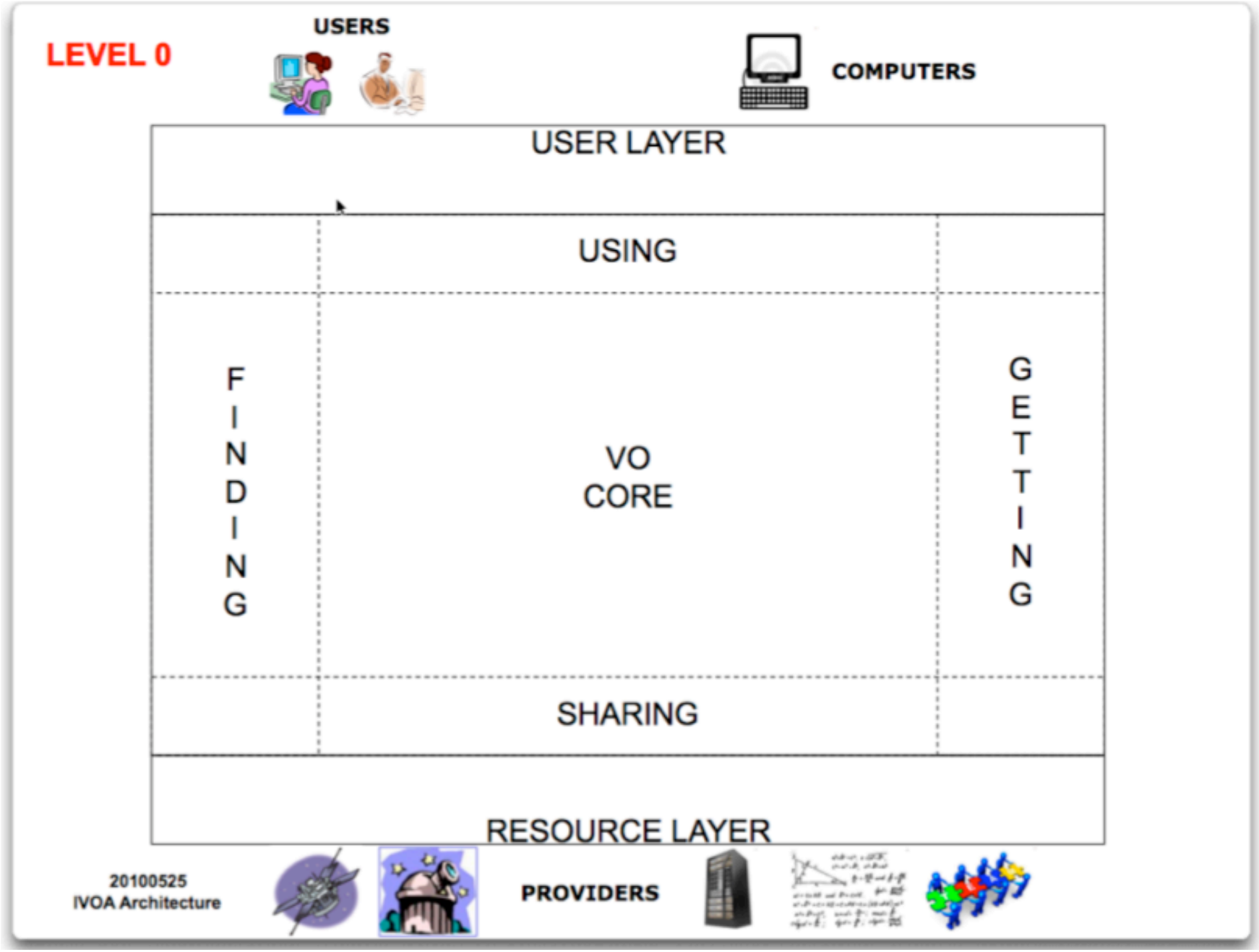
Other object types:	**Rad** (2C,3C,4C,3C… (3A,2E,1ES,1H,H,… (LBN,SH) ,**SNR** (A…
ICRS coord. (ep=2000) :	**05 34 31.97 +22** (
FK5 coord. (ep=2000 eq=2000) :	**05 34 31.97 +22** (
FK4 coord. (ep=1950 eq=1950) :	**05 31 31.43 +21** !
Gal coordinates :	**184.5575 -05.784**…
Parallaxes *mas*:	**7 [10] D** 1952GCTP..C.....0J
Spectral type:	**F D ~**

essential notes: • PSR 0531+21 is the central source of SNR 184.6-5.8 [*05-Jul-2004*].




2MASS
ESO-WFI
Chandra
VLT-ISAAC
HST-ACS
DSS
My own Data
Data Tree
Available data for the selected position are highlighted in the metadata tree
A.V.O demonstration prototype v1.0
Load...
Save...
Plugins...
Print...
Help...
Quit
The Fields of Views (FoV) Are overlaid automatically
Meta data on images (data model)
Info Frame
CDF-SOUTH-SECT23-VERSION1.0
Submit
Reset
Clear
Close

Ok, but what is it really?



Virtual Observatory

Data Centers

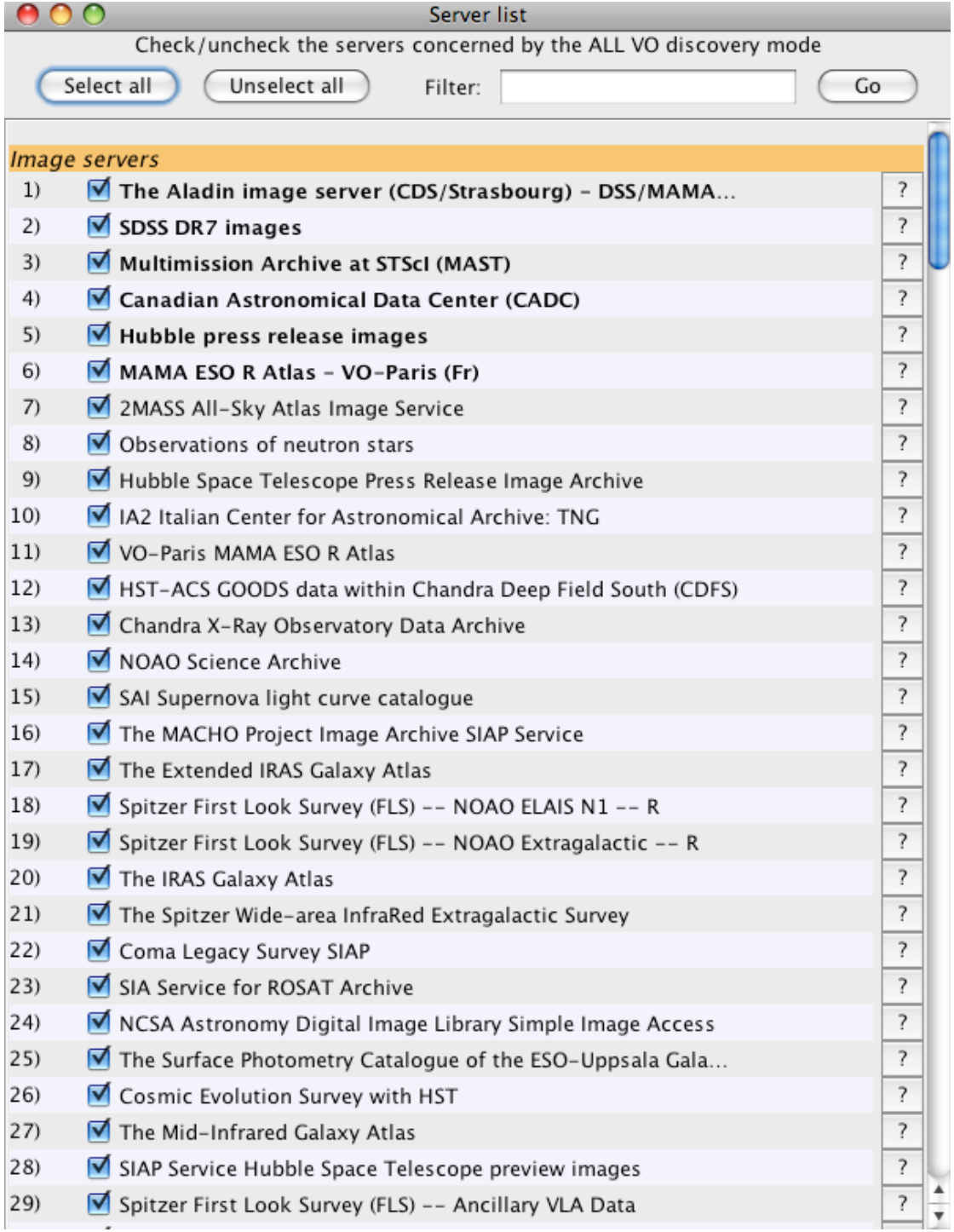
- The essential bricks of the VO
- Data centers in the VO can be very different, including very small teams
- Archives data, services with added value, development of softwares,…
- Keywords: user services, quality, durability
- In Strasbourg:

VO core: Standards!

- Protocols to:
  - Access images, access spectra, access tables
  - Exchange messages between applications
- Standard data formats
  - VOTable, FITS
  - UCDs (Unified Content Descriptors)
- Data models & meta-data
  - Describe the data
    - Instrument used, wavelength, epoch, …
  - What makes this dataset different from others?

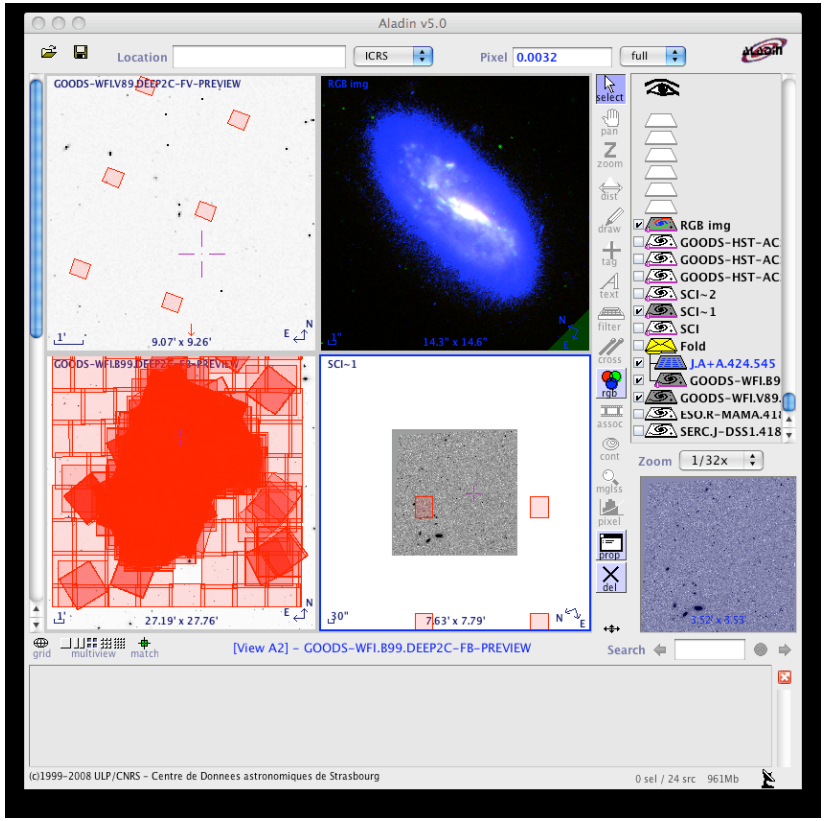
Services

- Registry: The yellow pages of the VO
  - Who is hosting data right now?
  - Are they available?
  - Which URL should I use?



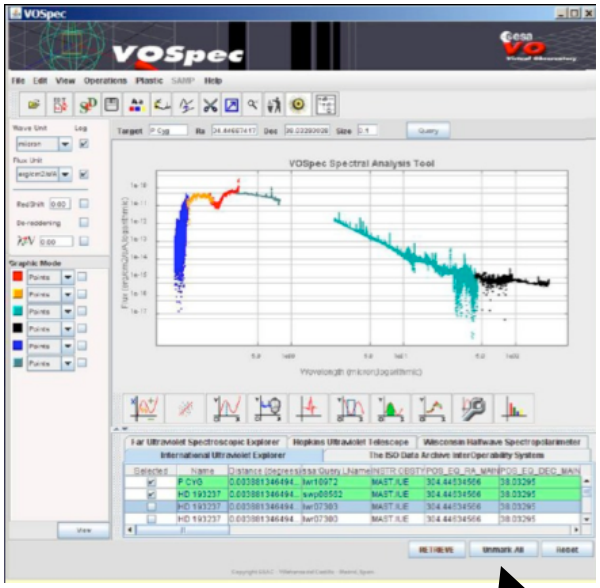
Tools

- Query, explore, visualize data
- Selection tools
- Estimate how usefull the data is for your science with simple analysis functionalities
- Aladin: one of the VO portals

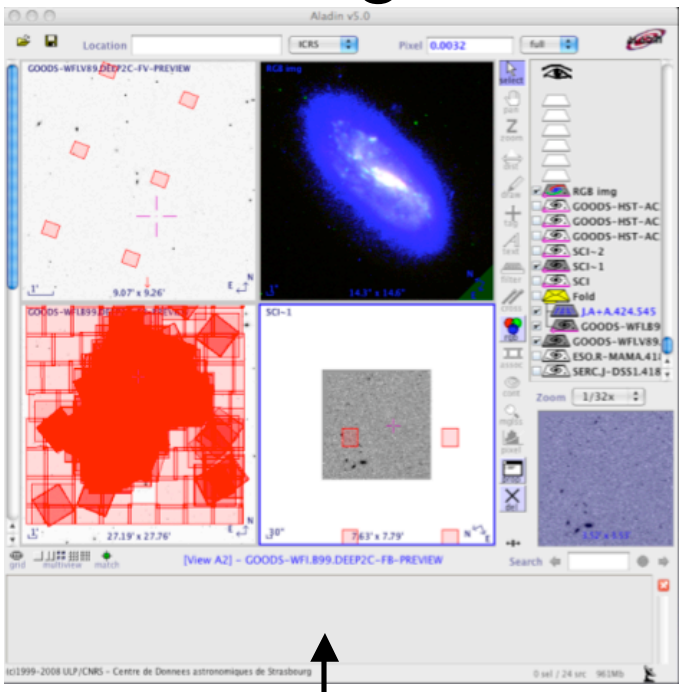


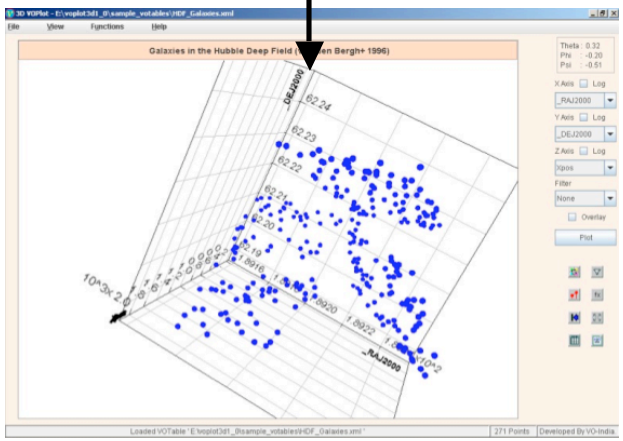
Standards and interoperability

- Interaction between tools as if one single tool



SAMP





Basic data :

V* CM Tau -- Pulsar

query a

Other object types:	**Rad** (2C,3C,4C,3CR,CTA,CTB,Cul,DA,DB,GRS,Mills,NRAO,NRL,PKS,SIM,VRO,W, (3A,2E,1ES,1H,H,1M,2U,3U,4U,X,[BM83],[KRL2007b]) ,**gam** (1CG,2CG,2EG,3E (LBN,SH) ,**SNR** (AJG,SNR) ,**Psr** (PSR,PULS) ,* (CSI,PLX) ,**IR** (IRAS,2MASS)
ICRS coord. (ep=2000) :	**05 34 31.97 +22 00 52.1 (~) [139 100 90] D** 1999A&AS..136..571H
FK5 coord. (ep=2000 eq=2000) :	**05 34 31.97 +22 00 52.1 (~) [139 100 90] D** 1999A&AS..136..571H
FK4 coord. (ep=1950 eq=1950) :	**05 31 31.43 +21 58 54.4 (~) [139 100 90] D** 1999A&AS..136..571H
Gal coordinates :	**184.5575 -05.7843 (~) [139 100 90] D** 1999A&AS..136..571H
Parallaxes *mas*:	**7 [10] D** 1952GCTP..C.....0J
Spectral type:	**F D ~**

essential notes: • PSR 0531+21 is the central source of SNR 184.6-5.8 [*05-Jul-2004*].

And now, let us show you how it works!