Astronomical Virtual Observatory and its Scientific Challenges

Petr Škoda

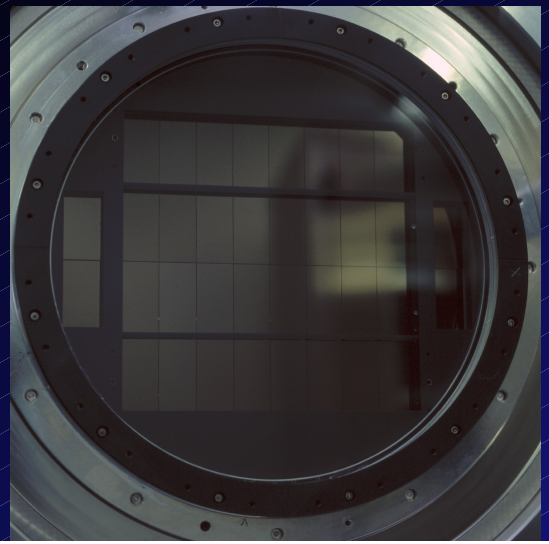
Astronomical Institute Academy of Sciences
Ondřejov
Czech Republic

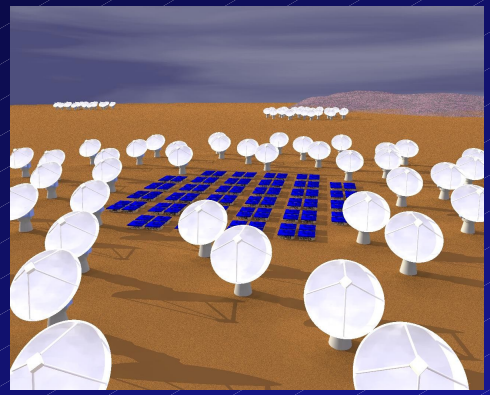
Institute of Astronomy of BAS, Sofia, Bulgaria, 1-st December 2009

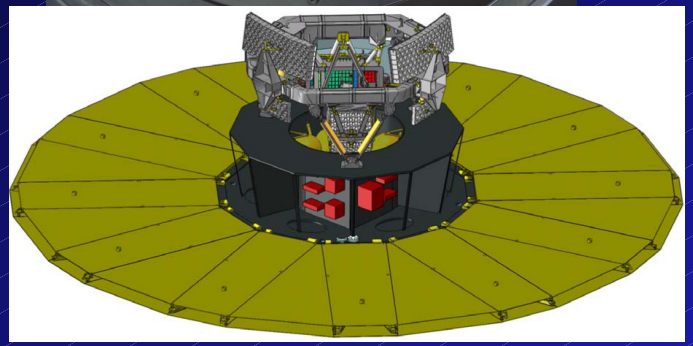
Data Avalanche

- *Huge surveys*:100 million sources at < 3000 sources per night ⇒> 100 years to identify them
- *Huge data collections*: download and data analysis on desktop problematic/impossible.
- Example: downloading Sloan Digital Sky Survey (SDSS) DR6 data:
  - ✓ images (10 Terabytes) ⇒~ 3 months at 10 Mbps
  - ✓ catalogues (2 Terabytes) ⇒~ 3 weeks
  - ✓ on DVDs ⇒~ 2,100 of them
- And *data analysis*?? (similar size for MACHO, 2MASS et

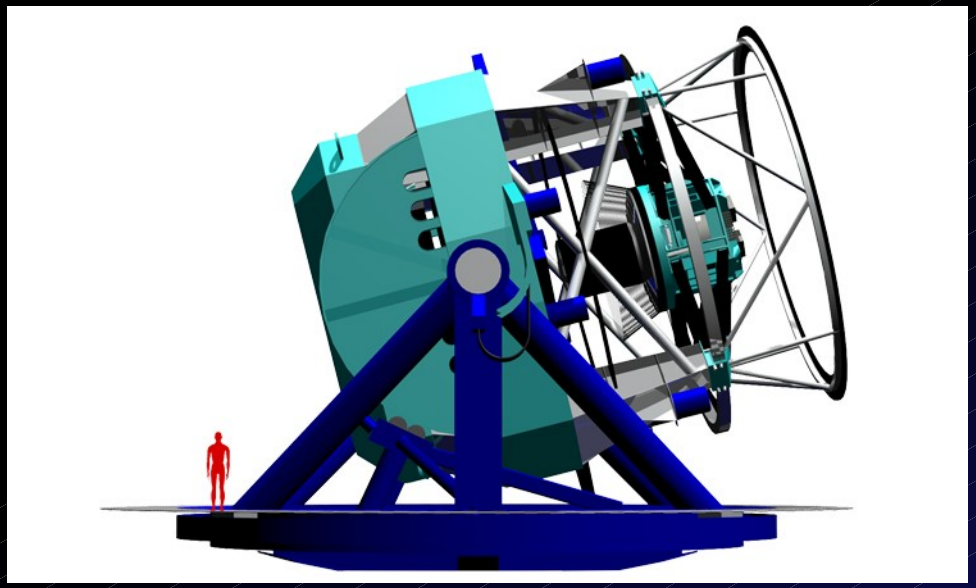
Data Avalanche

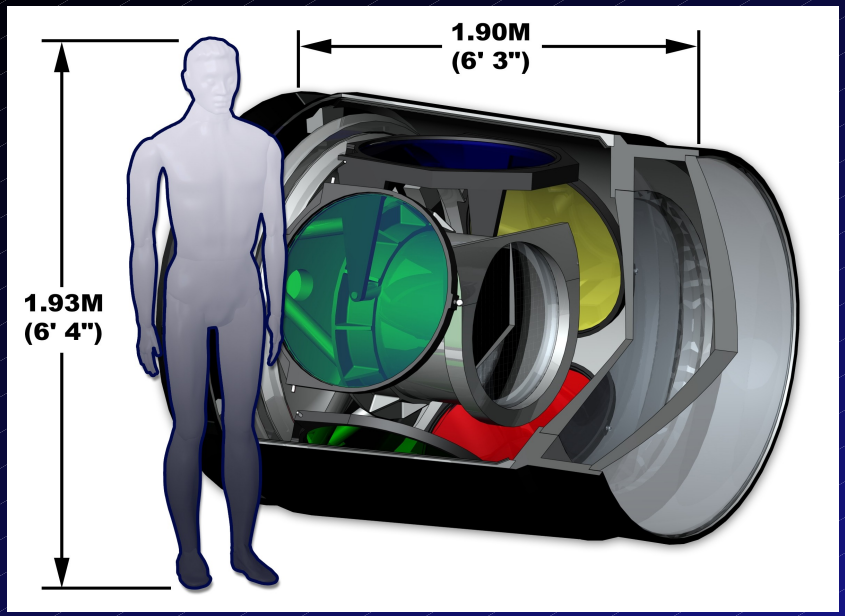


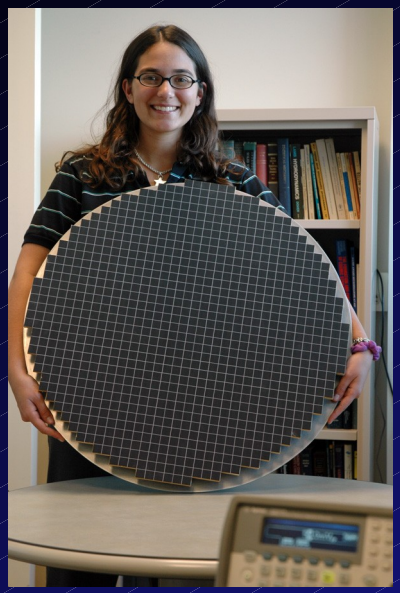




LSST (8.4m)





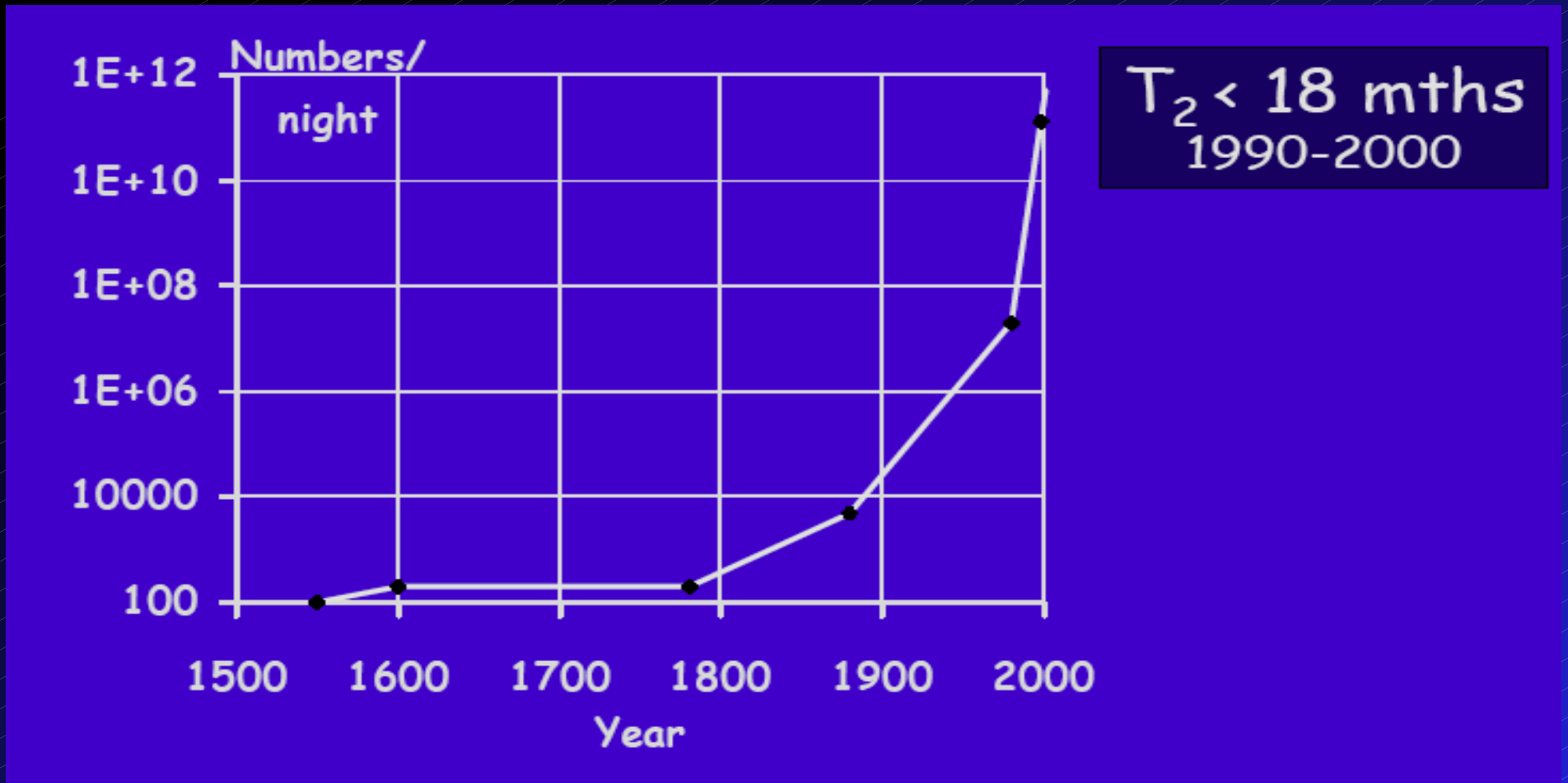


200 CCD 4kx4k,
32 channels (6400)
3.2 Gpix every 20 sec
64cm diameter
3.5 deg FOV
30 TB/night
2 TFLOPS
detection of changes
within 60sec

Data Avalanche

Moore law for chips –doubling 1.5 year (1000/10 yr)

Data in astronomy – doubling < 1 yr !



History of VO

idea VO end 2000

Federation of archives (MAST)

unified IF, data format for transport

much data – distributed processing

GRID - started in HEP

Multispectral research : radio---gama

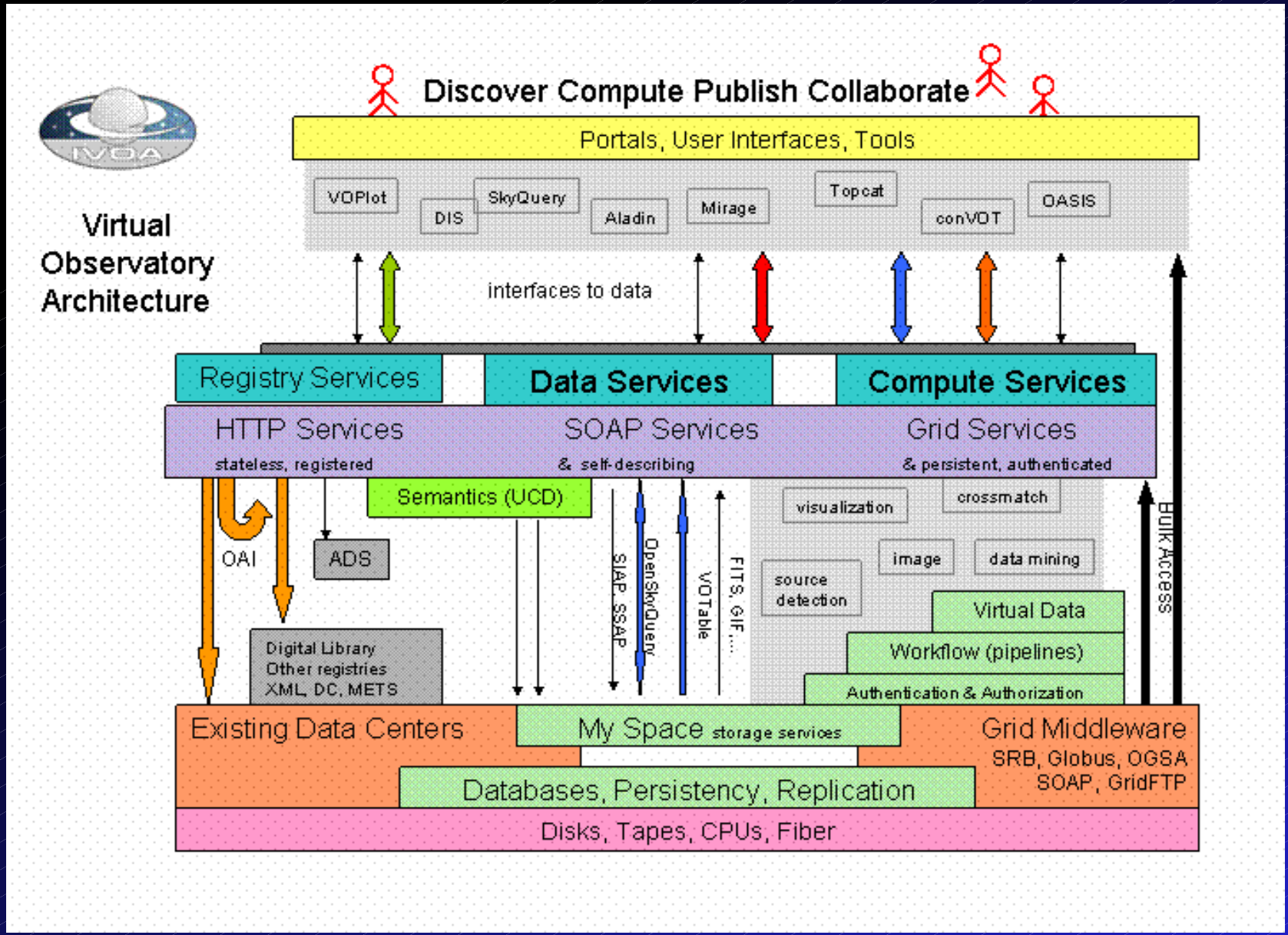
Output of simulations in NCSA

Data for SDSS – key research

VO Paradigma

- METADATA (name of column), ontologies (name)
- Unique format (VOTable – e.g Vizier)
- Transparent search, download, conversion
- Query for data – processing done on servers
- Federation of astronomical archives (protocols)
- Unified presentation – automatic units conversion (A,MeV,MHz->nm), Wm^-2s^-1 → Jy)
- Background computing on GRIDS
- Multiwavelength approach (SED)

Architecture of VO



Technology of VO

Unified data format– VOTable, UCD

Transparent transport (SOAP , REST)

Web services (WS)

VOregistry (like DNS)

VOSpace (network home disc)

protocols (CGI services)

- ConeSearch (searching in circle on sky)
- SIAP (Simple Image Access Protocol)
- SSAP(Simple Spectral Access Protocol)
- SLAP(Simple Line Access Protocol)
- TAP (Table Access Protocol)

Technology of VO

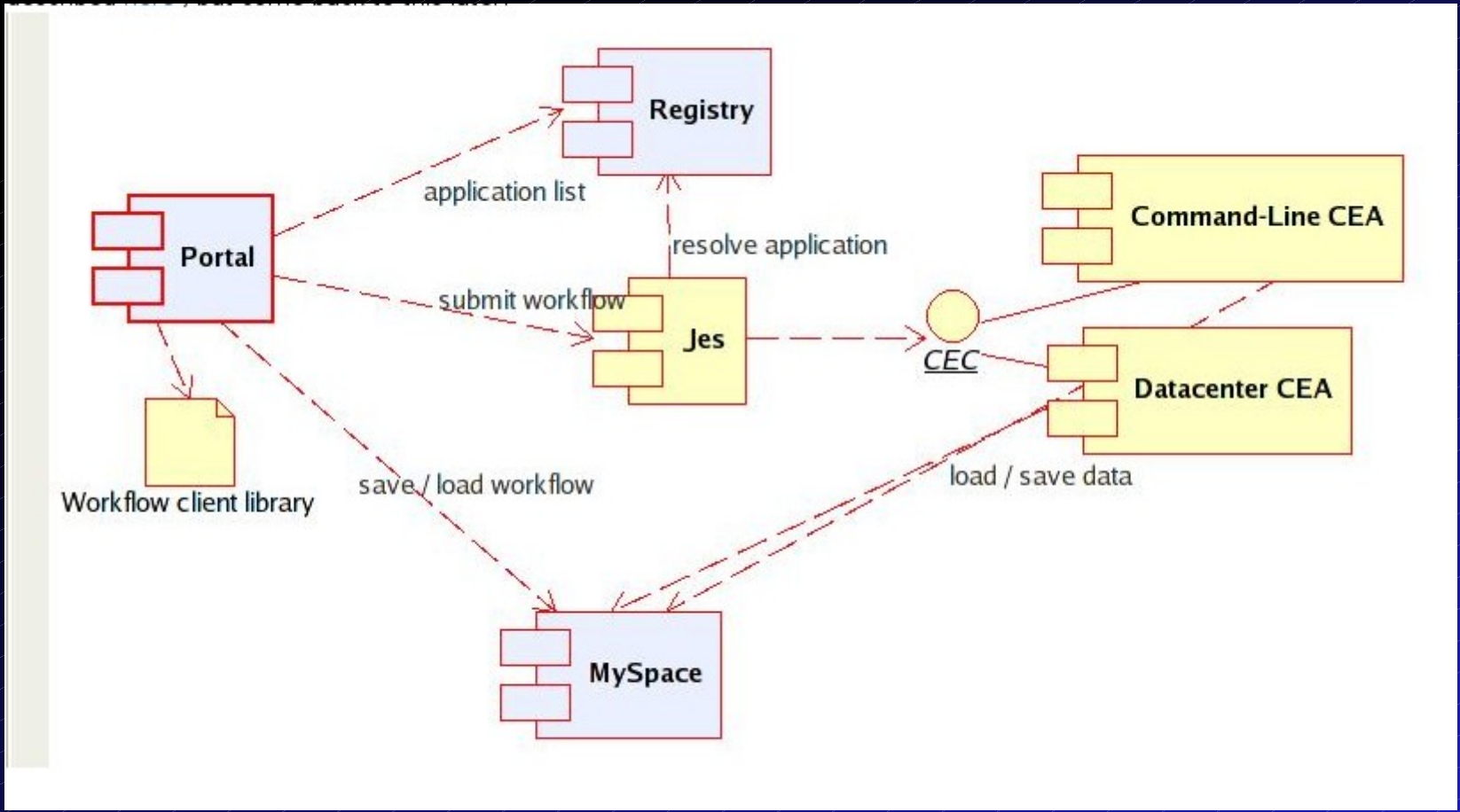
ADQL

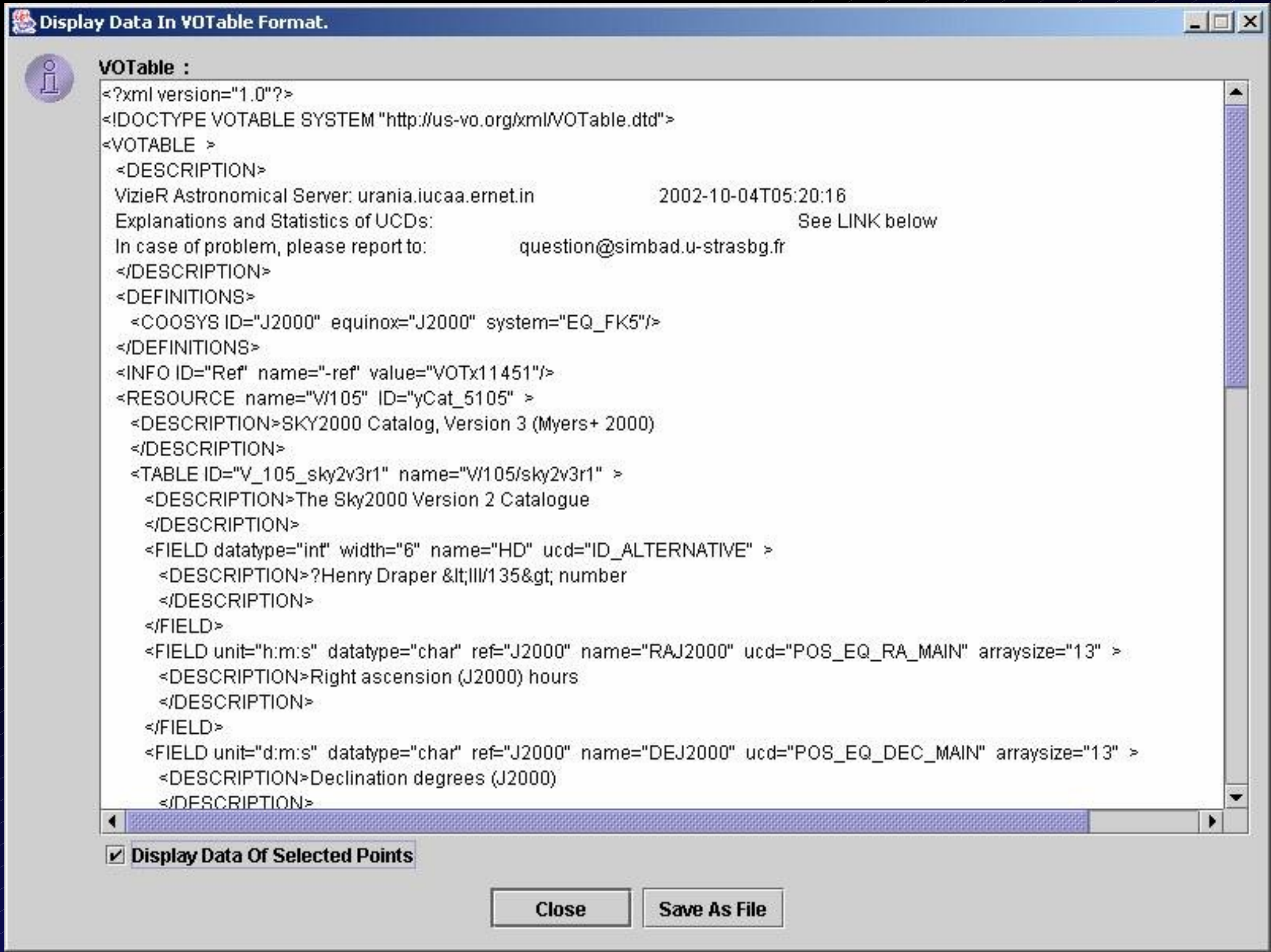
XMATCH, REGION

VOSpace – cache(WS, database,http put get)

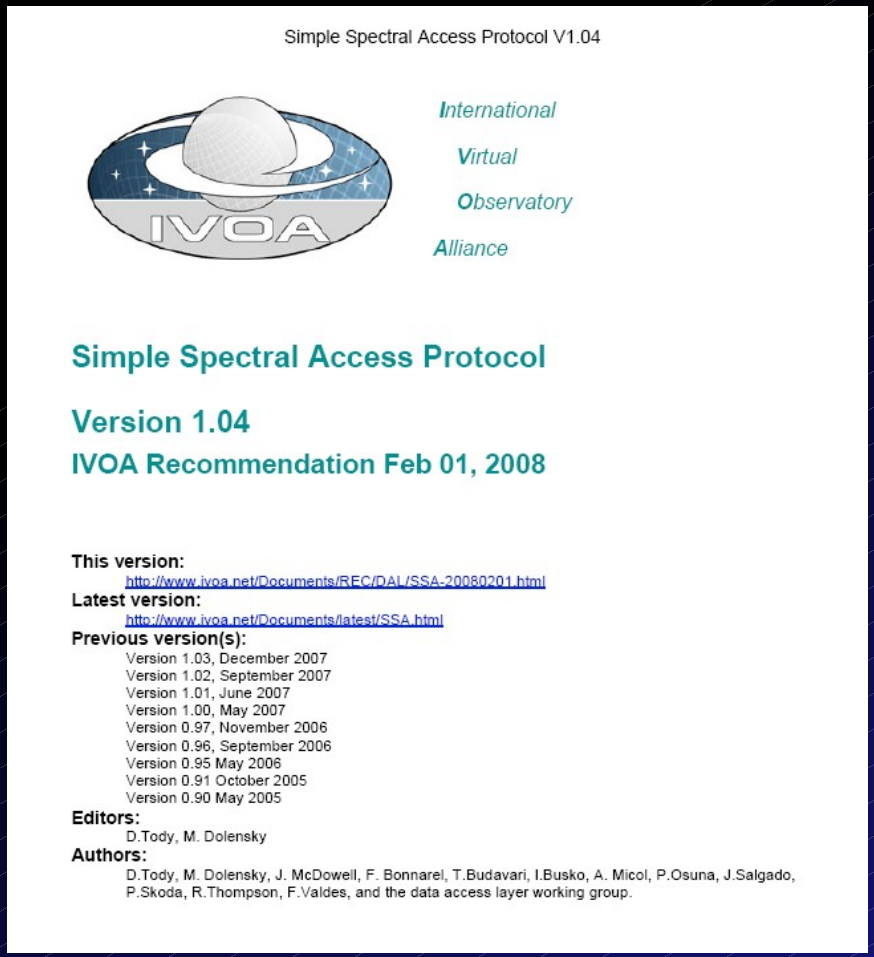
Application interop – PLASTIC, SAMP

Workflows - Astrogrid




Display Data In VOTable Format.
VOTable :
<?xml version="1.0"?>
<!DOCTYPE VOTABLE SYSTEM "http://us-vo.org/xml/VOTable.dtd">
<VOTABLE >
<DESCRIPTION>
VizieR Astronomical Server: urania.iucaa.ernet.in 2002-10-04T05:20:16
Explanations and Statistics of UCDs: See LINK below
In case of problem, please report to: question@simbad.u-strasbg.fr
</DESCRIPTION>
<DEFINITIONS>
<COOSYS ID="J2000" equinox="J2000" system="EQ_FK5"/>
</DEFINITIONS>
<INFO ID="Ref" name="-ref" value="VOTx11451"/>
<RESOURCE name="V/105" ID="yCat_5105" >
<DESCRIPTION>SKY2000 Catalog, Version 3 (Myers+ 2000)
</DESCRIPTION>
<TABLE ID="V_105_sky2v3r1" name="V/105/sky2v3r1" >
<DESCRIPTION>The Sky2000 Version 2 Catalogue
</DESCRIPTION>
<FIELD datatype="int" width="6" name="HD" ucd="ID_ALTERNATIVE" >
<DESCRIPTION>?Henry Draper <III/135> number
</DESCRIPTION>
</FIELD>
<FIELD unit="h:m:s" datatype="char" ref="J2000" name="RAJ2000" ucd="POS_EQ_RA_MAIN" arraysize="13" >
<DESCRIPTION>Right ascension (J2000) hours
</DESCRIPTION>
</FIELD>
<FIELD unit="d:m:s" datatype="char" ref="J2000" name="DEJ2000" ucd="POS_EQ_DEC_MAIN" arraysize="13" >
<DESCRIPTION>Declination degrees (J2000)
</DESCRIPTION>
Display Data Of Selected Points
Close
Save As File

Simple Spectra Access Protocol
Spectral Data Model



Simple Spectral Access Protocol V1.04

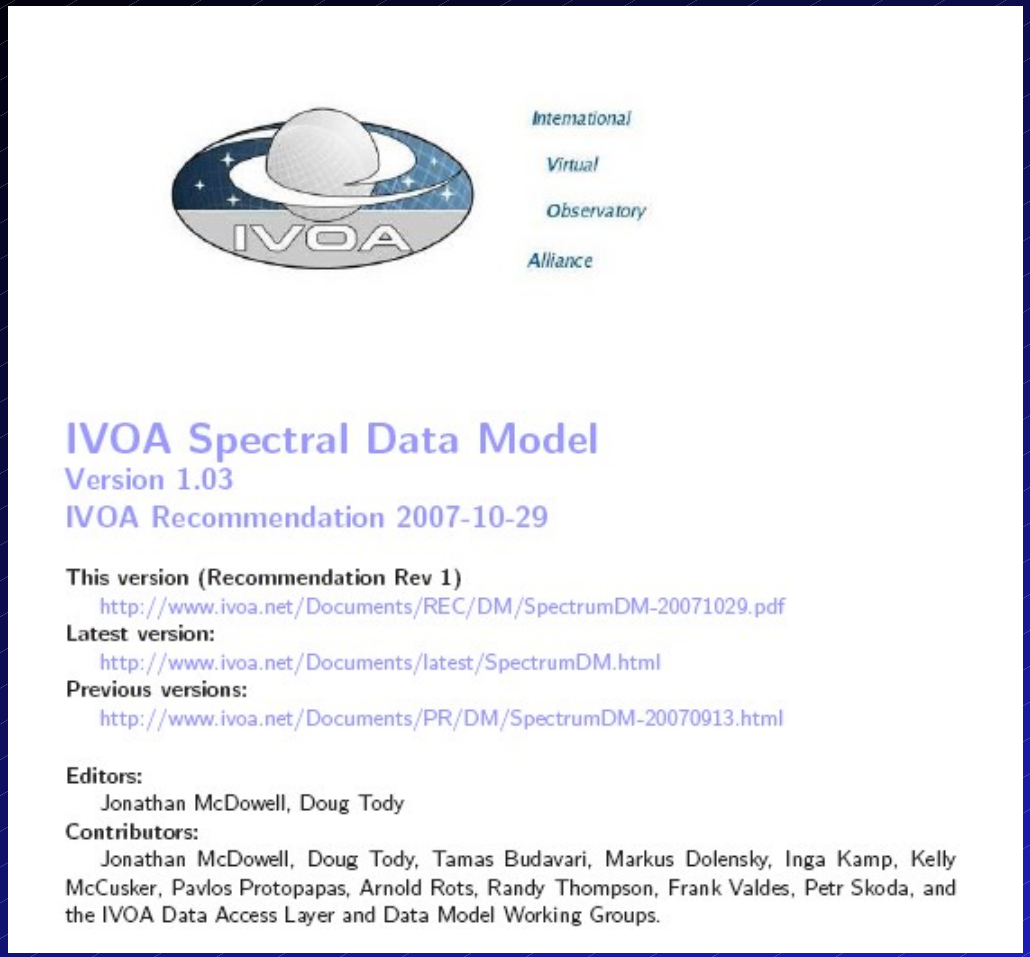
International Virtual Observatory Alliance

Simple Spectral Access Protocol

Version 1.04

IVOA Recommendation Feb 01, 2008

This version:
http://www.ivoa.net/Documents/REC/DAL/SSA-20080201.html
Latest version:
http://www.ivoa.net/Documents/latest/SSA.html
Previous version(s):
Version 1.03, December 2007
Version 1.02, September 2007
Version 1.01, June 2007
Version 1.00, May 2007
Version 0.97, November 2006
Version 0.96, September 2006
Version 0.95 May 2006
Version 0.91 October 2005
Version 0.90 May 2005
Editors:
D.Tody, M. Dolensky
Authors:
D.Tody, M. Dolensky, J. McDowell, F. Bonnarel, T.Budavari, I.Busko, A. Micol, P.Osuna, J.Salgado, P.Skoda, R.Thompson, F.Valdes, and the data access layer working group.



International Virtual Observatory Alliance

IVOA Spectral Data Model

Version 1.03

IVOA Recommendation 2007-10-29

This version (Recommendation Rev 1)
http://www.ivoa.net/Documents/REC/DM/SpectrumDM-20071029.pdf
Latest version:
http://www.ivoa.net/Documents/latest/SpectrumDM.html
Previous versions:
http://www.ivoa.net/Documents/PR/DM/SpectrumDM-20070913.html

Editors:
Jonathan McDowell, Doug Tody
Contributors:
Jonathan McDowell, Doug Tody, Tamas Budavari, Markus Dolensky, Inga Kamp, Kelly McCusker, Pavlos Protopapas, Arnold Rots, Randy Thompson, Frank Valdes, Petr Skoda, and the IVOA Data Access Layer and Data Model Working Groups.

SSAP Parameters

4.1.1 Mandatory Query Parameters

The following parameters **must** be implemented by a compliant service:

Parameter	*Sample value*	*Physical unit*	*Datatype*
POS	52,-27.8	degrees; defaults to ICRS	string
SIZE	0.05	degrees	double
BAND	2.7E-7/0.13	meters	string
TIME	1998-05-21/1999	ISO 8601 UTC	string
FORMAT	votable	-	string

4.1.2 Recommended and Optional Query Parameters

Parameter	Sample value	Unit	Req	Datatype
APERTURE	0.00028 (=1°)	degrees	OPT	double
SPECRP	2000	λ/dλ	REC	double
SPATRES	0.05	degrees	REC	double
TIMERES	31536000 (=1yr)	seconds	OPT	double
SNR	5.0	dimensionless	OPT	double
REDSHIFT	1.3/3.0	dimensionless	OPT	string
VARAMPL	0.77	dimensionless	OPT	string
TARGETNAME	mars		OPT	string
TARGETCLASS	star		OPT	string
FLUXCALIB	relative		OPT	string
WAVECALIB	absolute		OPT	string
PUBDID	ADS/col#R5983		REC	string
CREATORDID	ivo://auth/col#R1234		REC	string
COLLECTION	SDSS-DR5		REC	string
TOP	20	dimensionless	REC	int
MAXREC	5000		REC	string
MTIME	2005-01-01/2006-01-01	ISO 8601	REC	string
COMPRESS	true		REC	boolean
RUNID			REC	string

The spatial, spectral, and time resolution of the data may all be used as query constraints

IVOA


Aus-VO
EURO VO
China-VO
VO-i
Astro Grid
CVO
RVO
SVO
NVO
IVOA
ArVO
JVO
HVO
GAVO
France

VO-enabled tools

Aladin

VOPlot

TOPCAT

VOSpec

SpecView

SPLAT

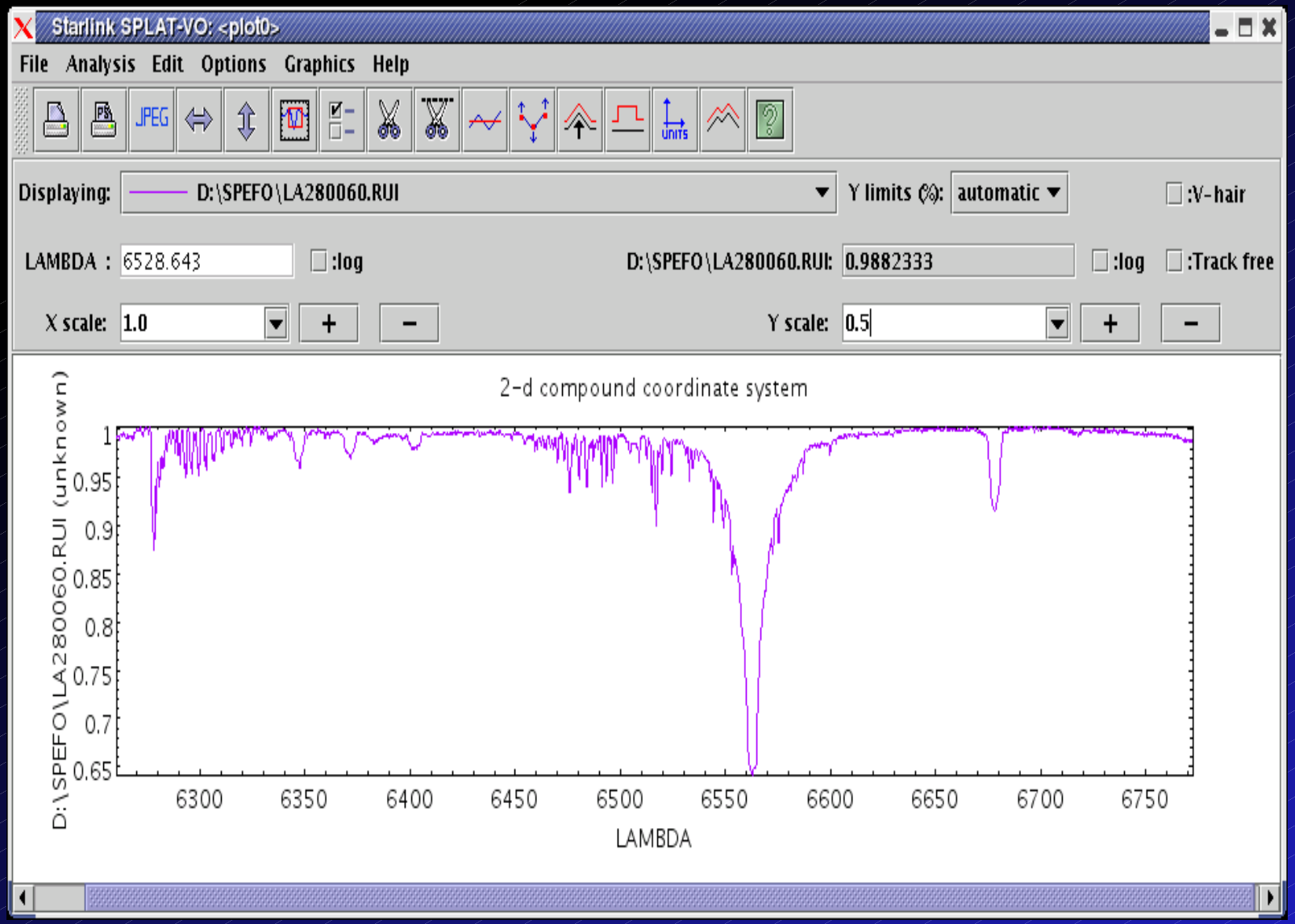
ViSiVio

VOSED

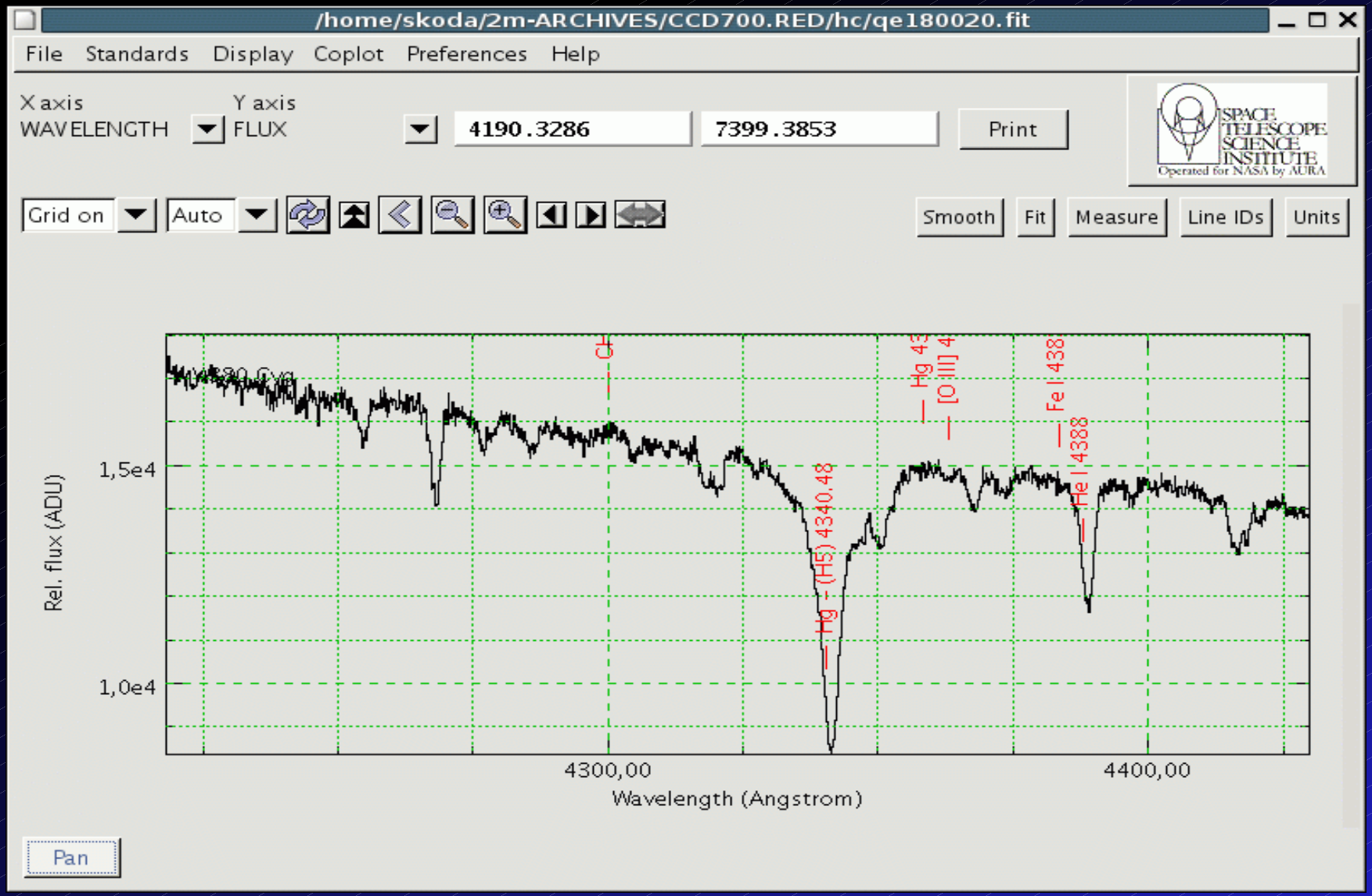
Sextractor – WESIX (Web Enabled Source Identification with Cross Matching)

PLASTIC HUB, Treeview + Period04 (since 18.9.08)

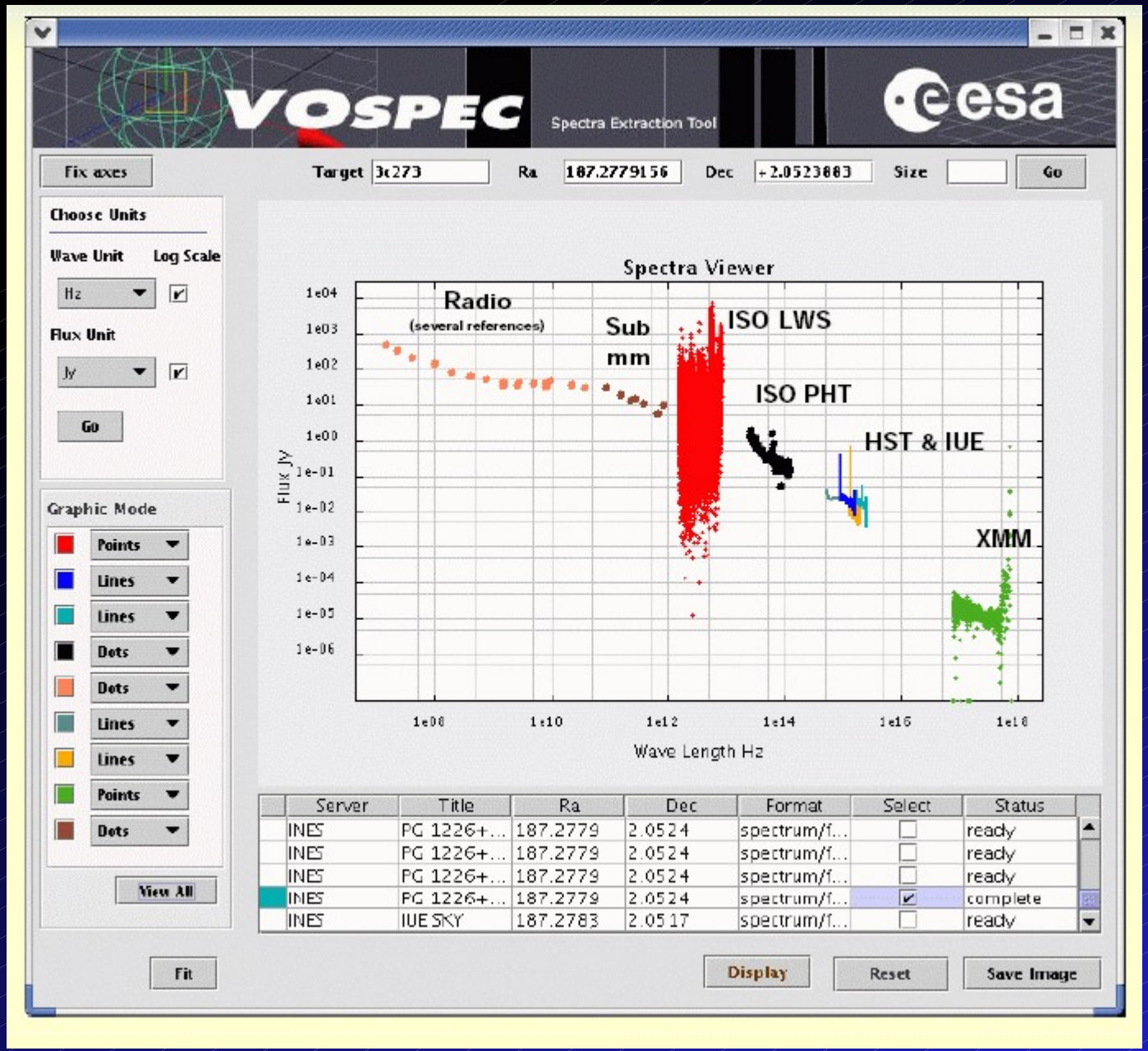
SPLAT-VO (Starlink, JAC)



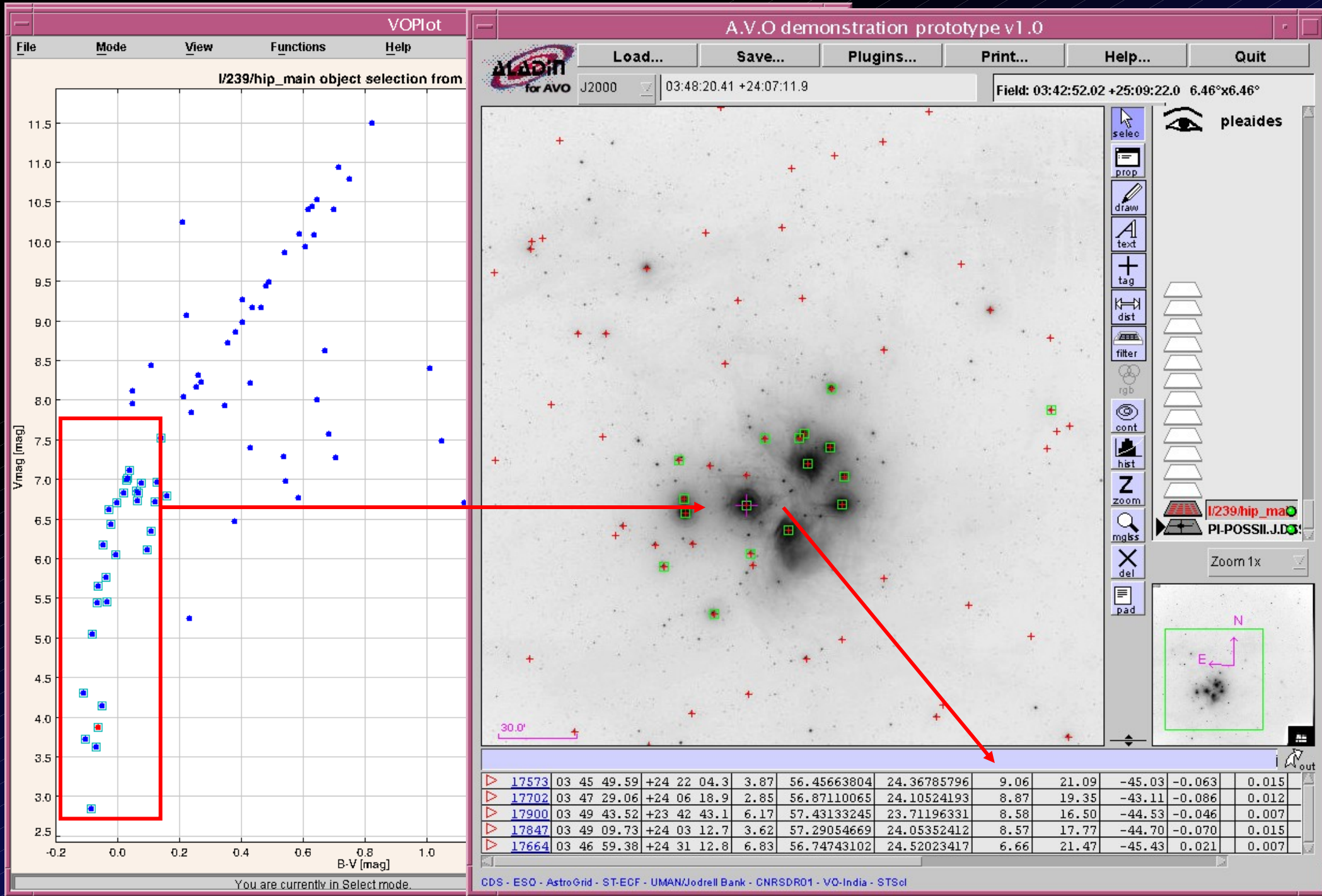
SpecView (STScI)

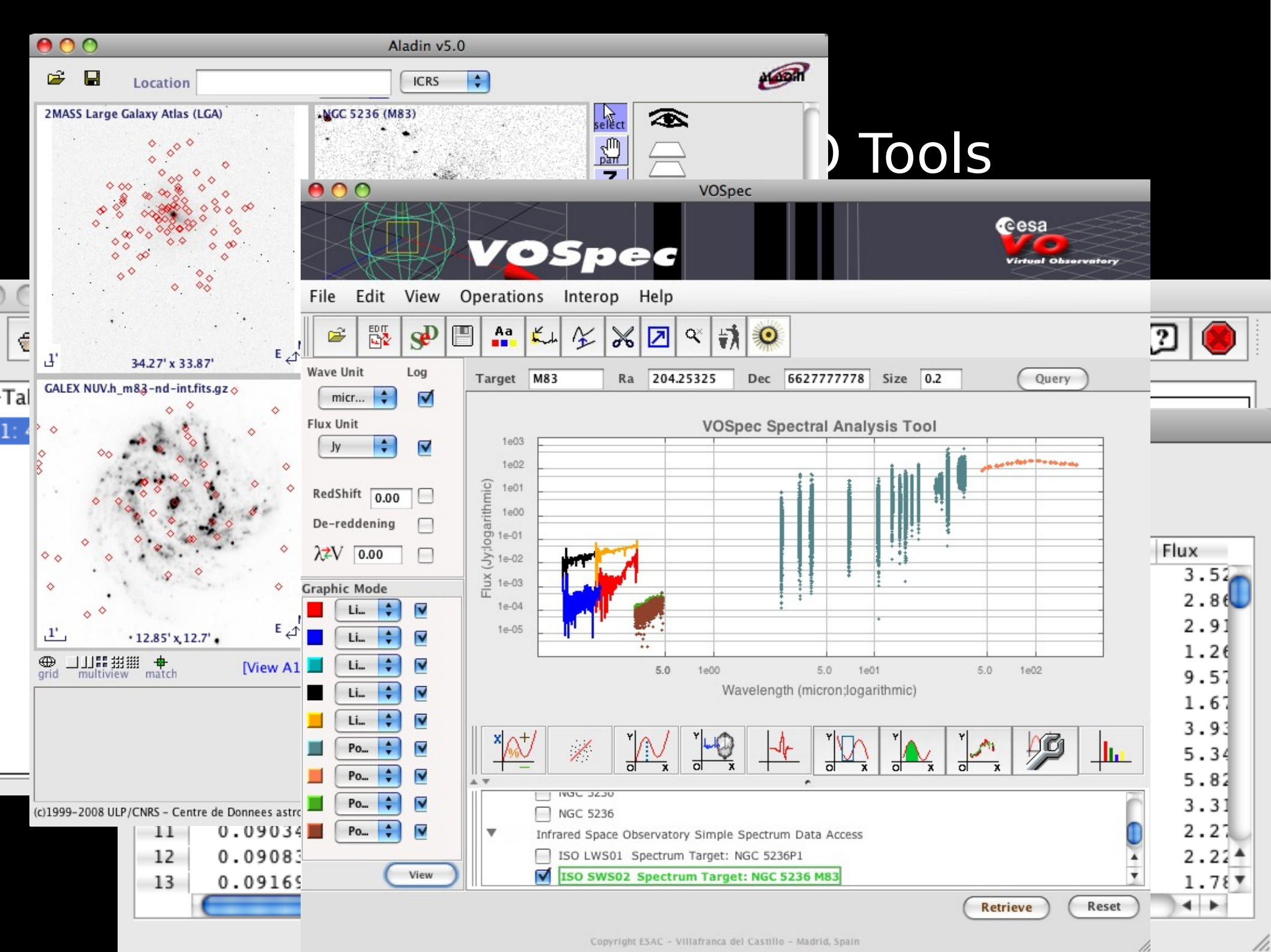


VOspec (ESAC)

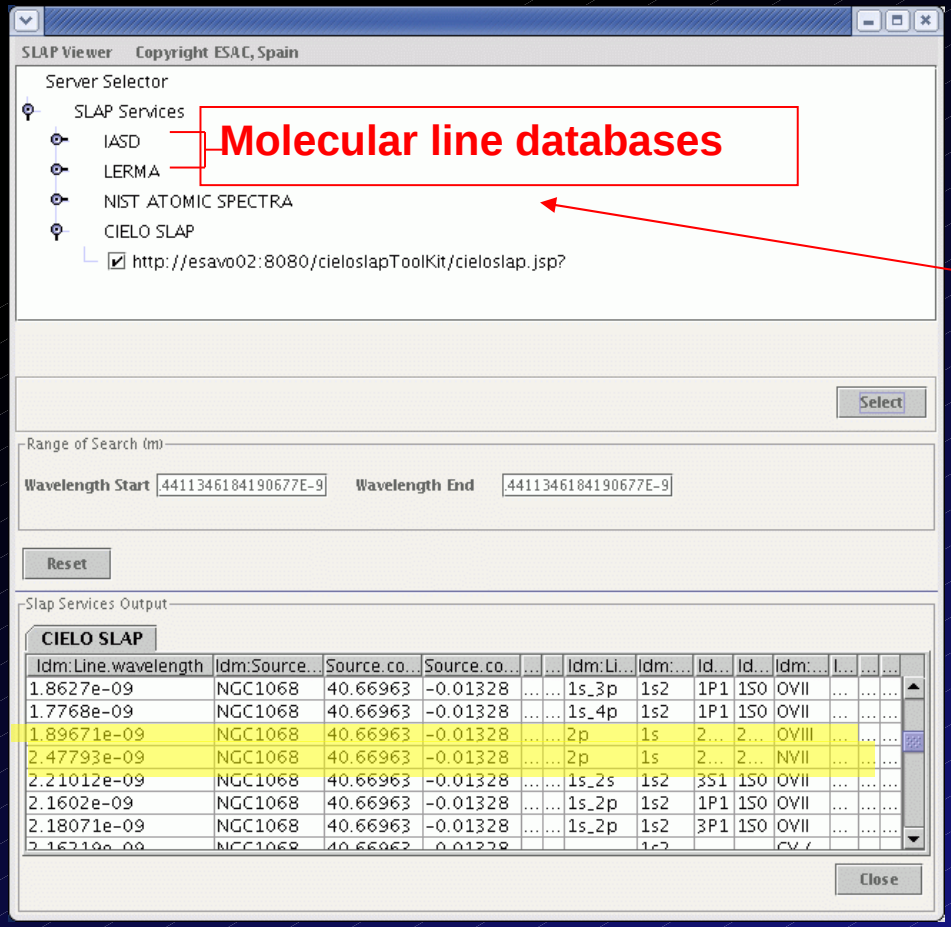

VOSPEC
Spectra Extraction Tool
esa
Fix axes
Target 3c273
Ra 187.2779156
Dec +2.0523883
Size
Go
Choose Units
Wave Unit
Log Scale
Hz
Flux Unit
Jy
Go
Graphic Mode
Points
Lines
Lines
Dots
Dots
Lines
Lines
Points
Dots
View All
Spectra Viewer
Radio
(several references)
Sub
mm
ISO LWS
ISO PHT
HST & IUE
XMM
Flux Jy
1e04
1e03
1e02
1e01
1e00
1e-01
1e-02
1e-03
1e-04
1e-05
1e-06
1e08
1e10
1e12
1e14
1e16
1e18
Wave Length Hz
Server
Title
Ra
Dec
Format
Select
Status
INES
PG 1226+...
187.2779
2.0524
spectrum/f...
ready
INES
PG 1226+...
187.2779
2.0524
spectrum/f...
ready
INES
PG 1226+...
187.2779
2.0524
spectrum/f...
ready
INES
PG 1226+...
187.2779
2.0524
spectrum/f...
complete
INES
IUE SKY
187.2783
2.0517
spectrum/f...
ready
Fit
Display
Reset
Save Image

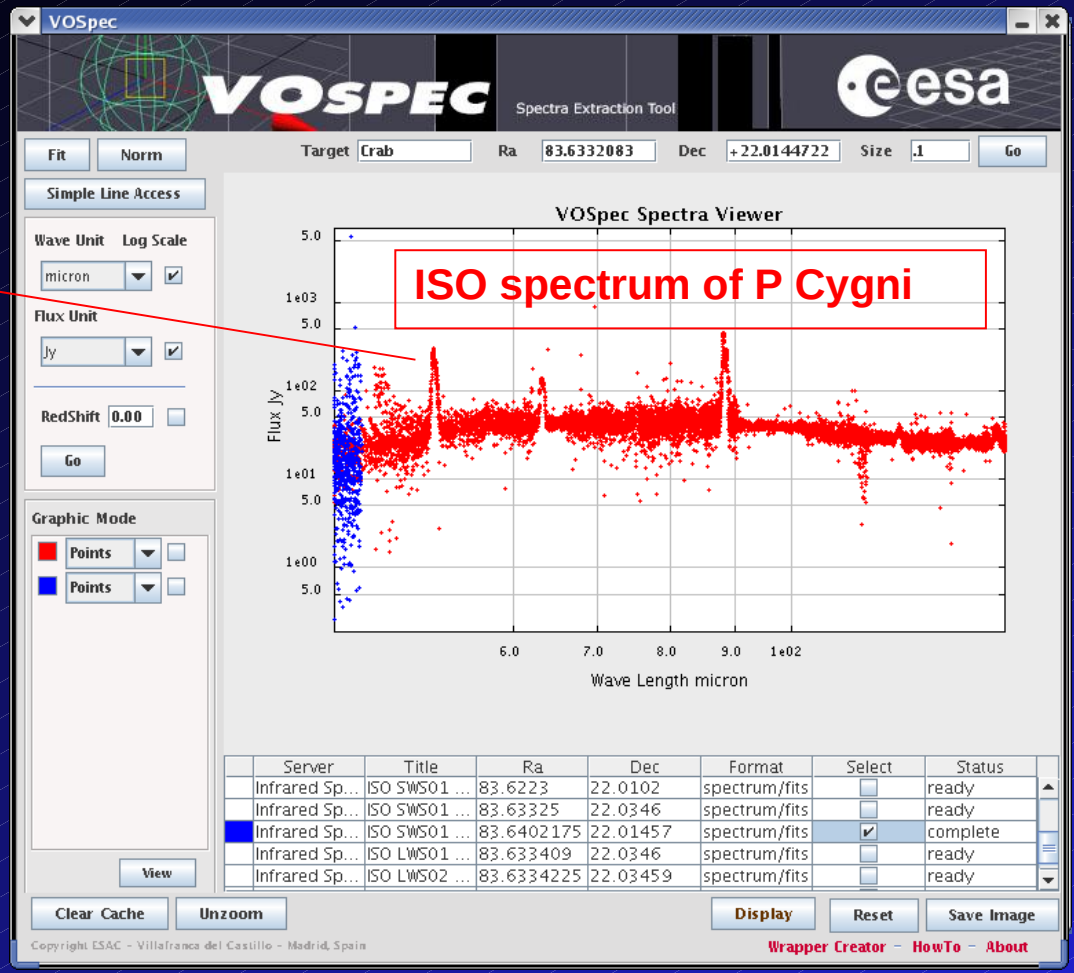
Colour-magnitude diagram




Aladin v5.0
Location
ICRS
2MASS Large Galaxy Atlas (LGA)
NGC 5236 (M83)
select
pan
34.27' x 33.87'
GALEX NUV.h_m83-nd-int.fits.gz
12.85' x 12.7'
grid
multiview
match
[View A1
(c)1999-2008 ULP/CNRS - Centre de Donnees astro
11 0.09034
12 0.09083
13 0.09169
Tools
VOSpec
VOSpec
esa
Virtual Observatory
File Edit View Operations Interop Help
Wave Unit
micr...
Log
Flux Unit
Jy
RedShift 0.00
De-reddening
0.00
Graphic Mode
Li...
Po...
View
Target M83
Ra 204.25325
Dec 6627777778
Size 0.2
Query
VOSpec Spectral Analysis Tool
Flux (Jy;logarithmic)
1e03
1e02
1e01
1e00
1e-01
1e-02
1e-03
1e-04
1e-05
5.0
1e00
5.0
1e01
5.0
1e02
Wavelength (micron;logarithmic)
NGC 5236
Infrared Space Observatory Simple Spectrum Data Access
ISO LWS01 Spectrum Target: NGC 5236P1
ISO SWS02 Spectrum Target: NGC 5236 M83
Retrieve
Reset
Copyright ESAC - Villafranca del Castillo - Madrid, Spain
Flux
3.52
2.86
2.91
1.26
9.57
1.67
3.93
5.34
5.82
3.31
2.27
2.22
1.78

CIELO VO – line catalogue SLAP





(IVOA Line Data Model: Dubernet, Osuna et al., in preparation)
(Simple Line Access Protocol: Salgado et al., in preparation)

VO for Atomic and Molecular Data

VAMDC (06/2009-12/2012 FP7)

13 organizations

Virtual Atomic and Molecular Data Centre

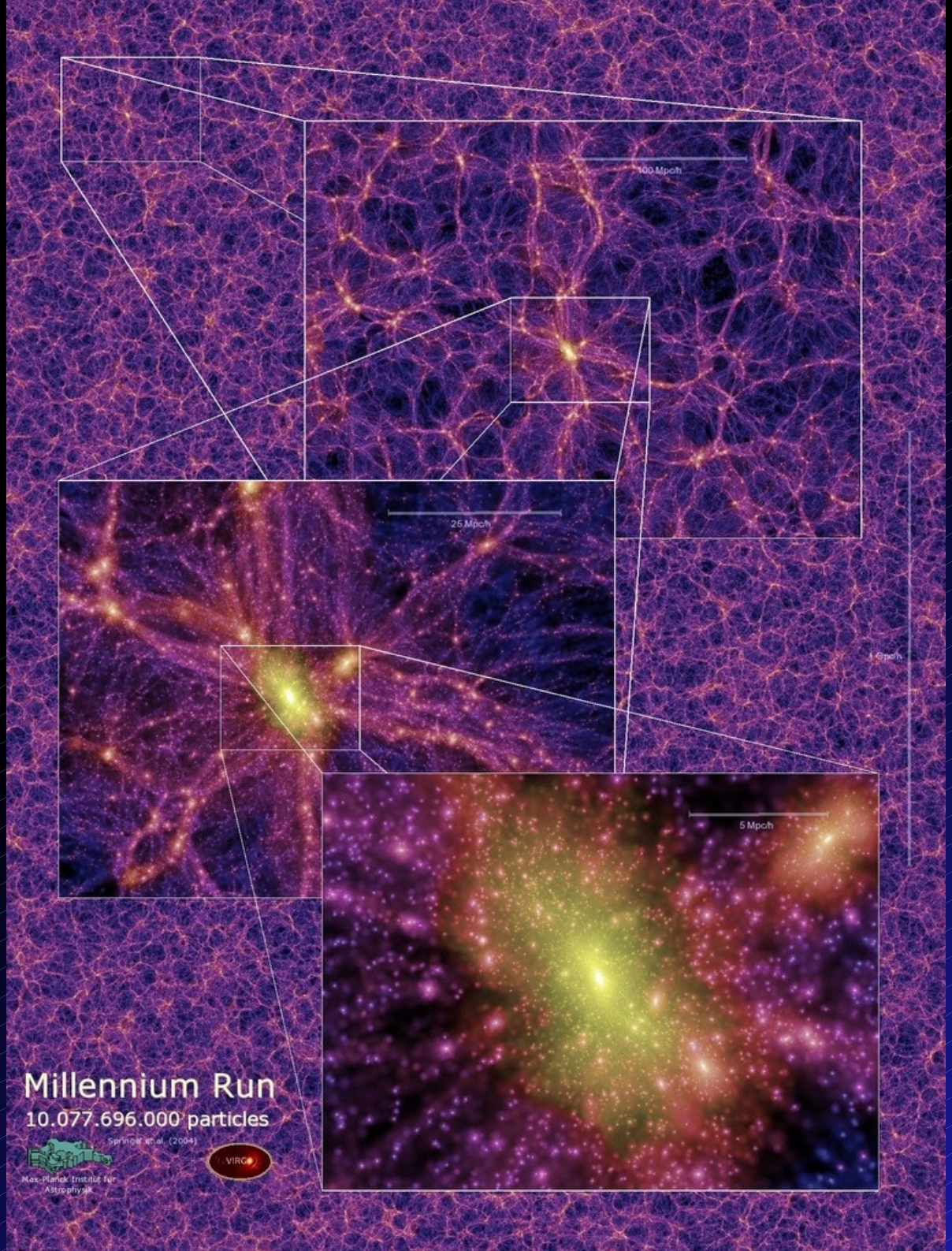
VO principles (web services, integration, registry, SAMP, VODesktop, TOPCAT, VOSpec)

(includes VALD extractor, NIST)

extended citation system (all providers acknowledged)

Theory VO (TVO)

- Methods of VO (parameters in DB, SQL...) for study of results of simulations , catalogues of simulated objects like SDSS...(PCA)
- Browsing of simulation space along different axes – parameters, regions...
- Evolutionary tracks, Photo Dissociation Regions
- Formation of artificial galaxies, clusters – N body models (MR - 10 billions, 25TB)
- Theoretical Spectra



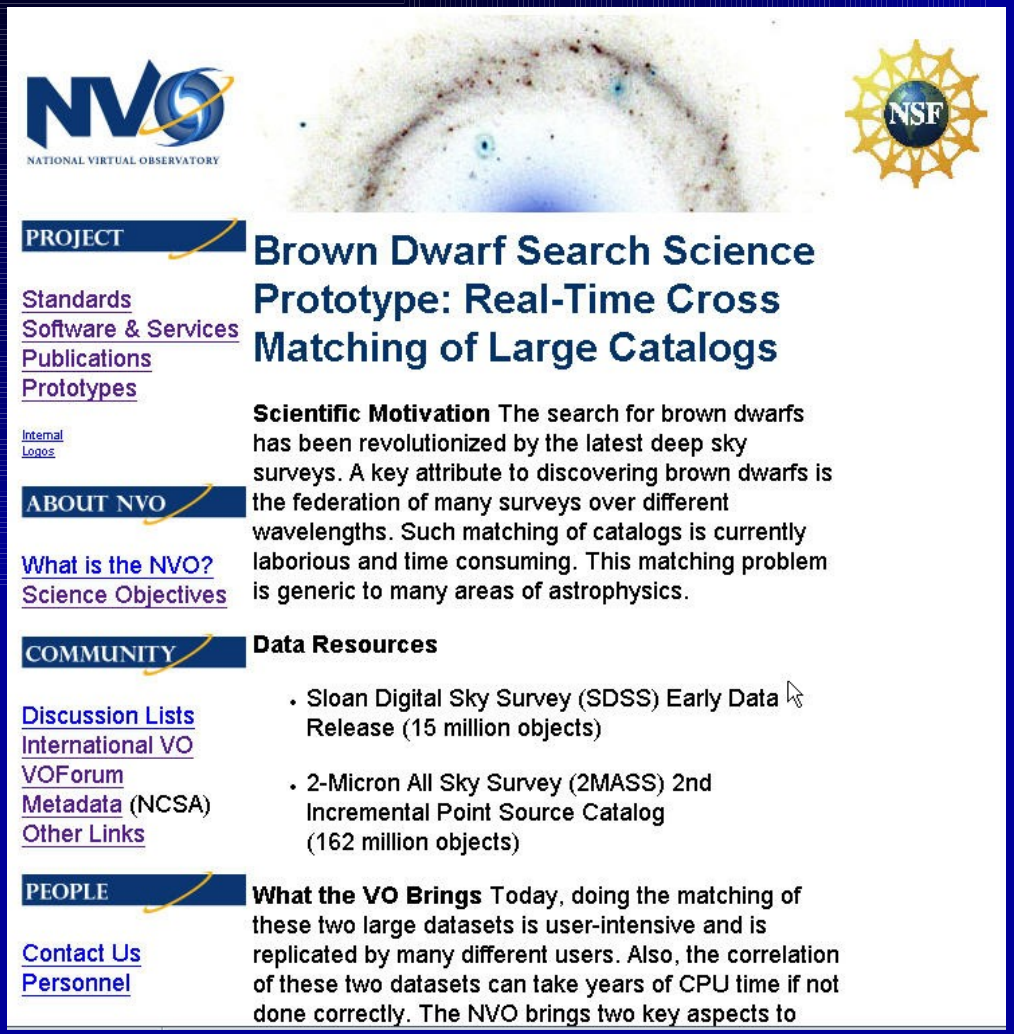
Millenium Run

- 10^10 particles
- Several Gpc to
- 10 kpc
- Cube 2 billion ly
- One month MPSSC
- 25 TB
- Evolution of 20 mil galaxies
- Evolution merger tree

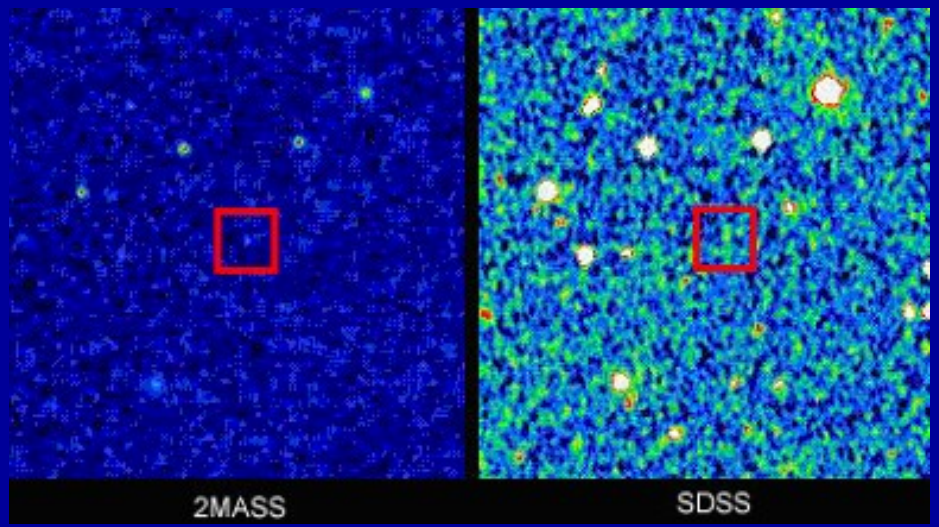
Examples of VO Science

- Discovery of obscured QSO (Padovani)
- Brown dwarfs (about 20 candidates)
- Brightest (WD?) (Caballero et al.) Albus-1
- Widest CPM binaries
- AGB to PN (100 new to 200 known using VO)
- SED (Spectrum Energy Distribution)
- Bolometric magnitude
- VOEvent – robotic telescopes (GRB, transits,)
- Outreach , Education (MS WWT, GoogleSky)

BDs discovered using VO



- **Filtering criteria:** z & J-only detections with *z- J > 2.75*
- *SDSS: 15M obj.*
- *2MASS: 160M obj.*
- *300000 objects in common.*

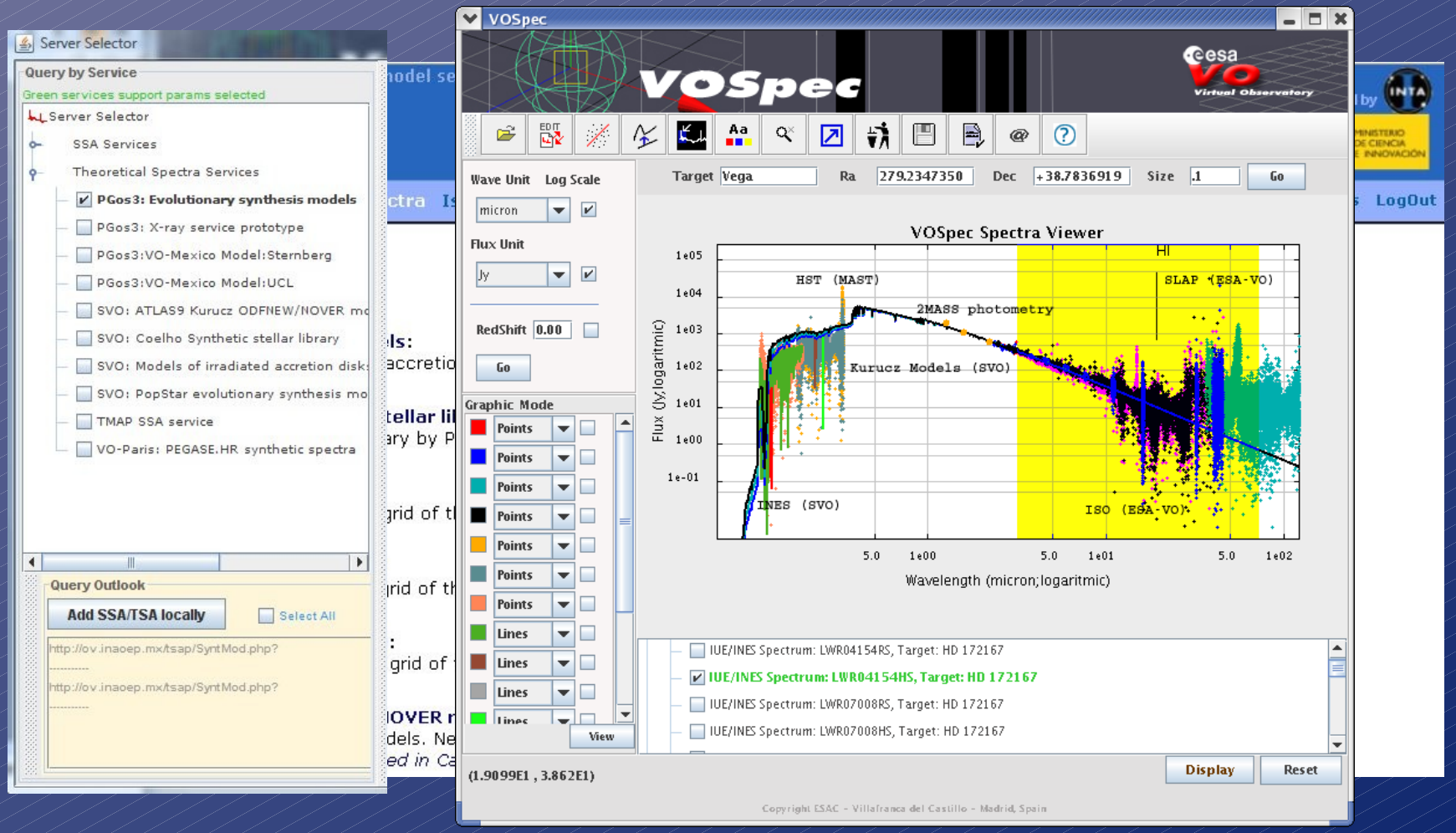


✓ *However, systematic searches using a VO methodology have not been performed so far.*

Archives, Theory, VO-Science, DataMining, E&O

Simple Spectral Access Protocol V1.04

Appendix A: Theoretical Spectral Access Use Case



Archives, Theory, VO-Science, DataMining, E&O

- Easy access to multi-λ info.
- Easy access to models (S3).

Astronomy & Astrophysics manuscript no. Synth˙VO˙PR1˙ref˙format
August 2, 2008

VOSA: Virtual Observatory SED Analyzer.

An application to the Collinder 69 open cluster

A. Bayo[1,2], C. Rodrigo[1,2], D. Barrado y Navascués[1,2], E. Solano[1,2], R. Gutiérrez[1,2], M. Morales–Calderón[1], and F. Allard[3]

Astronomy & Astrophysics manuscript no. 09595
April 14, 2008
© ESO 2008

Young stars and brown dwarfs surrounding Alnilam (ϵ Ori) and Mintaka (δ Ori)

J. A. Caballero[1,2] and E. Solano[3]

[1] Dpto. de Astrofísica y Ciencias de la Atmósfera, Facultad de Física, Universidad Complutense de Madrid, E-28040 Madrid, Spain
[2] Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany
[3] Laboratorio de Astrofísica Espacial y Física Fundamental, INSA, Apdo. 78, E-28691 Villanueva de la Cañada, Madrid, Spain

Received 17 February 2008; accepted 9 April 2008

Archives, Theory, VO-Science, DataMining, E&O

- Discovery of exotic objects. Access to large databases
  - 2MASS, SDSS, UKIDSS, IPHAS,…

Astronomy & Astrophysics manuscript no. aaVLM-iphas-v1
October 22, 2008
© ESO 2008

An IPHAS-based search for accreting very low-mass objects using VO tools

L. Valdivielso[1], E. L. Martín[1,2], H. Bouy[1,3], E. Solano[4,5], J. E. Drew[6,7], R. and Vink, J. S.[11]

Albus 1: A very bright white dwarf candidate

José Antonio Caballero[1,2]

Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany

caballero@mpia.de

Draft version November 6, 2009
Preprint typeset using LaTeX style emulateapj v. 05/04/06

THE FIRST L SUBDWARF IDENTIFIED IN THE UKIDSS INFRARED DEEP SKY SURVEY

N. Lodieu[1,2], M. R. Zapatero Osorio[3], E. L. Martín[3], E. Solano [3], M. Aberasturi[3]

Draft version November 6, 2009

Archives, Theory, VO-Science, DataMining, E&O

Theoretical model services

SVO

VOSA: VO Sed Analizer

VO SED Analyzer

Services: VOSA Filters TSAP S3if

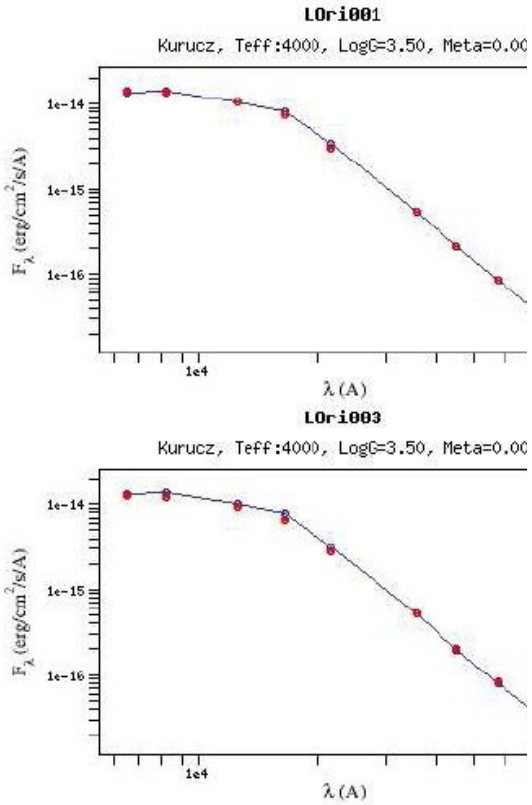
Astronomy & Astrophysics manuscript no. Synth˙VO˙PR1˙ref˙format
August 2, 2008

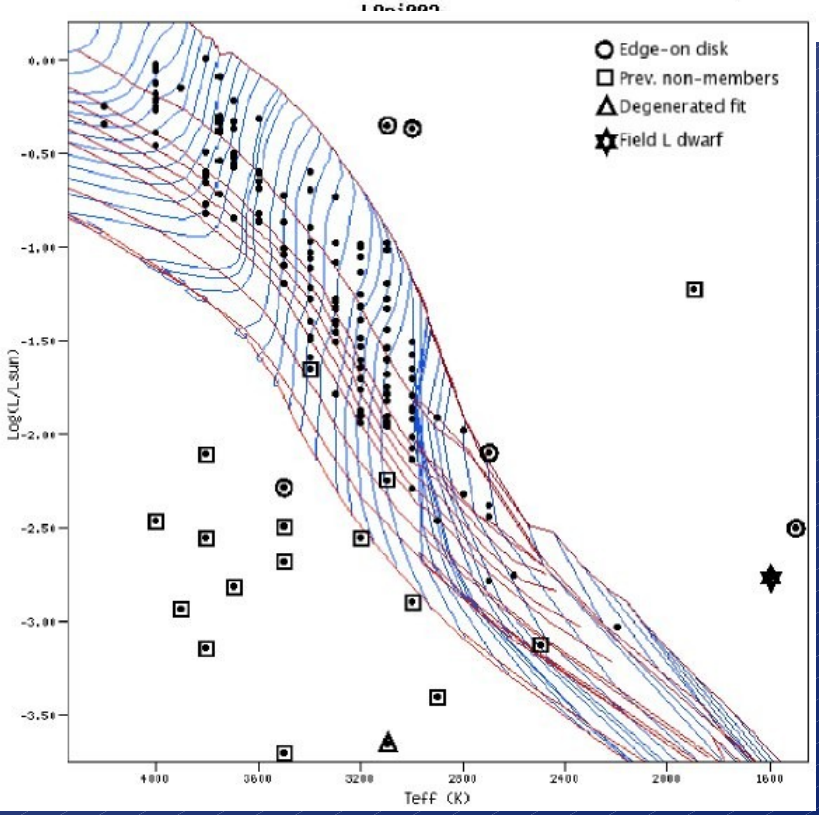
VOSA: Virtual Observatory SED Analyzer.

An application to the Collinder 69 open cluster

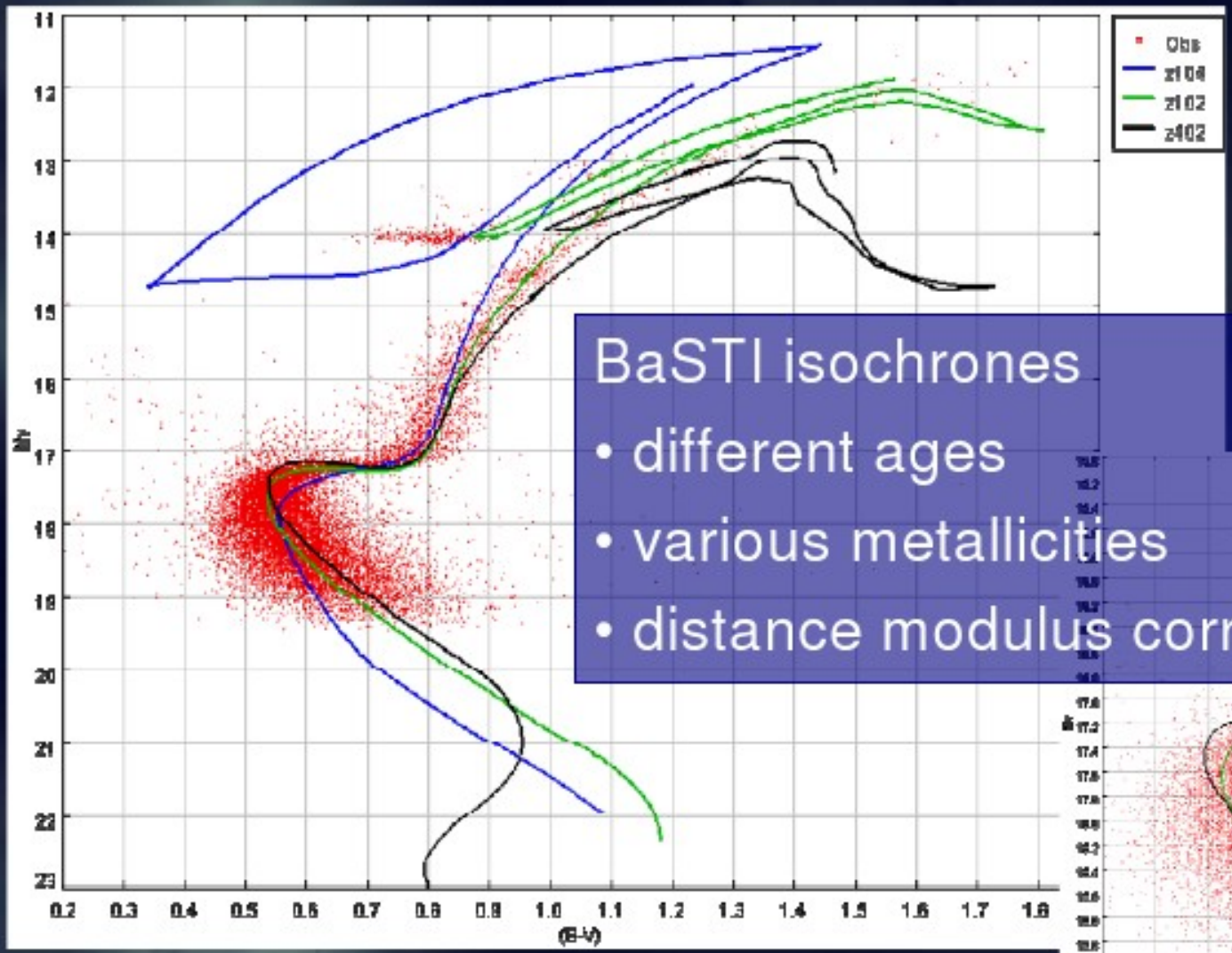
A. Bayo[1,2], C. Rodrigo[1,2], D. Barrado y Navascués[1,2], E. Solano[1,2], R. Gutiérrez[1,2], M.

Object	Model	T_{eff}	LogG	Metallicity	χ^2
LOri001	Kurucz	4000	3.50	0.00	1.71e
LOri002	NextGen	3800	4.5	0	1.44e
LOri003	Kurucz	4000	3.50	0.00	5.05e
LOri004	Kurucz	3750	4.00	0.00	2.82e
LOri005	NextGen	4000	4.0	0	9.20e
LOri006	Kurucz	4000	3.50	0.00	3.36e
LOri007	Kurucz	4000	4.50	0.00	2.49e
LOri008	Kurucz	4000	3.50	0.00	4.43e
LOri009	NextGen	4000	3.5	0	6.22e
LOri010	NextGen	4200	4.0	0	1.07e
LOri011	NextGen	3900	4.5	0	1.20e
LOri012	NextGen	4000	4.5	0	6.56e
LOri013	NextGen	3700	4.5	0	3.50e
LOri014	Kurucz	4000	4.50	0.00	4.70e
LOri015	Kurucz	4000	3.50	0.00	8.73e
LOri016	Kurucz	3750	4.50	0.00	5.63e
LOri017	NextGen	4200	4.0	0	7.56e
LOri018	Kurucz	3750	3.50	0.00	4.31e
LOri019	Kurucz	3750	3.50	0.00	1.90e
LOri020	NextGen	3800	4.5	0	2.98e
LOri021	Kurucz	4000	3.50	0.00	3.08e
LOri022	Kurucz	3750	4.00	0.00	1.78e
LOri023	NextGen	4000	4.5	0	2.35e
LOri024	Kurucz	3750	3.50	0.00	2.22e
LOri025	NextGen	3700	4.5	0	1.37e
LOri026	NextGen	3700	4.5	0	4.01e
LOri027	NextGen	4000	4.5	0	2.35e
LOri028	Kurucz	3750	4.00	0.00	1.26e
LOri029	NextGen	3100	4.5	0	7.28e
LOri030	NextGen	3700	4.5	0	2.15e
LOri031	NextGen	3800	4.5	0	2.77e
LOri032	NextGen	3700	4.5	0	1.70e
LOri033	NextGen	3700	4.5	0	7.12e
LOri034	NextGen	3000	4.0	0	1.77e
LOri035	NextGen	3700	4.5	0	1.61e



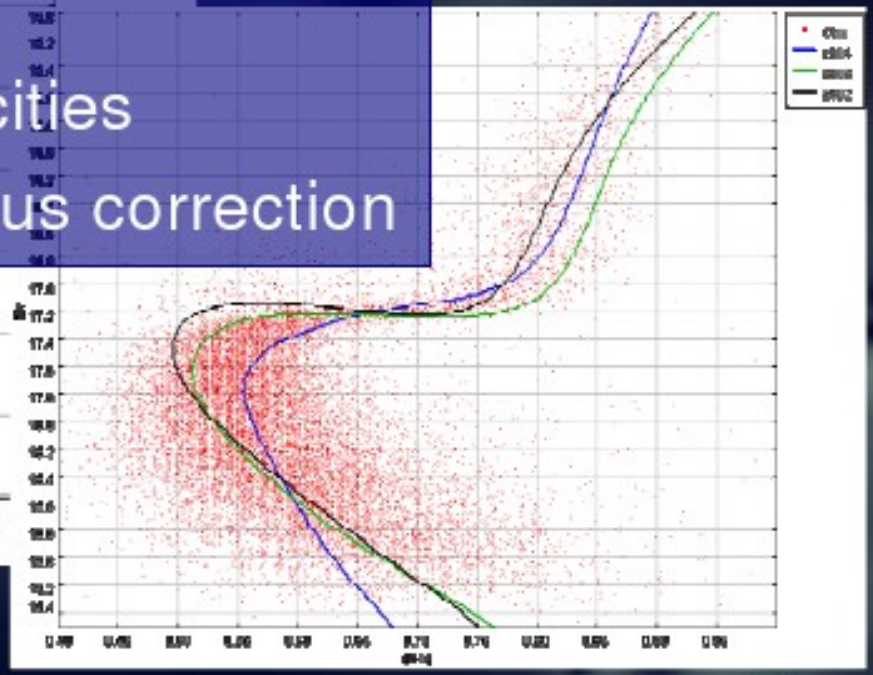


BaSTI Isochrones



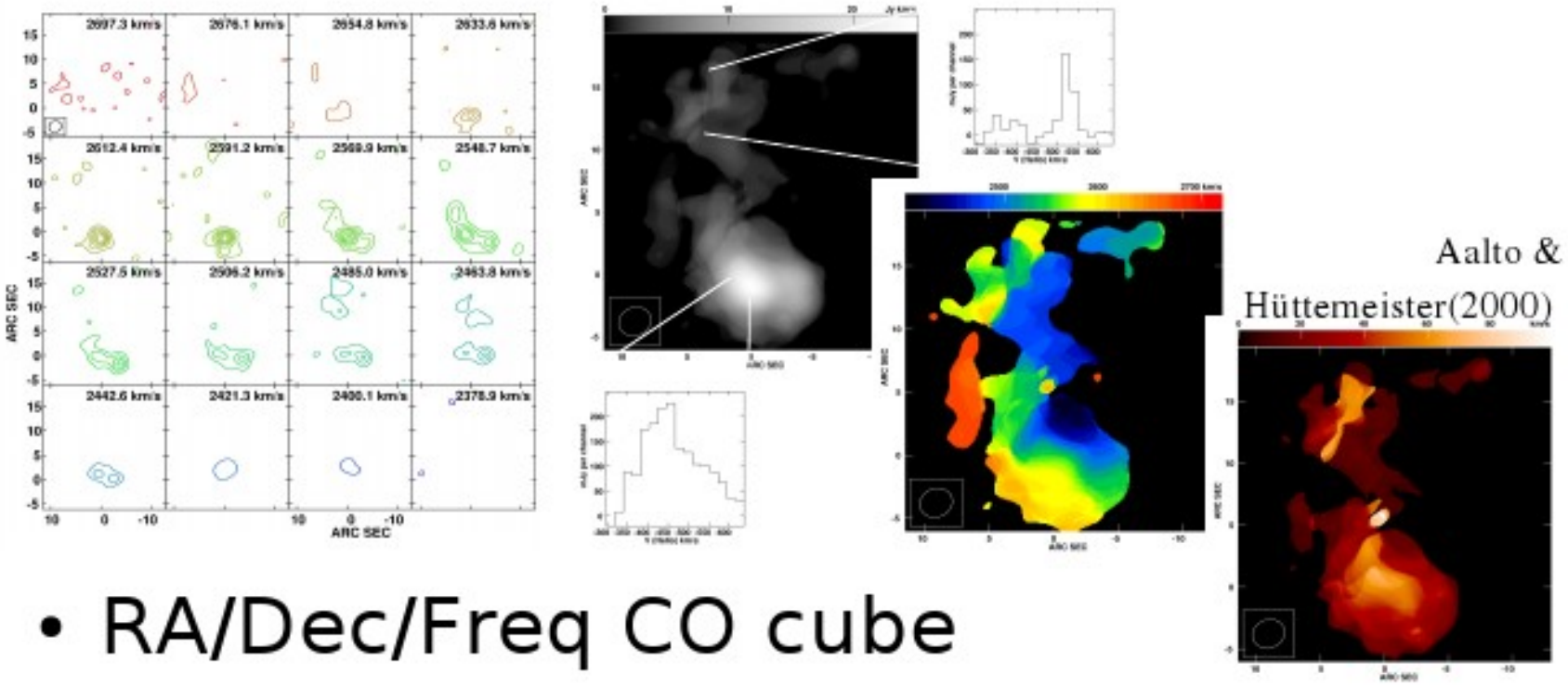
step 1
metallicity

BaSTI isochrones
- different ages
- various metallicities
- distance modulus correction



z = 0.01 (α-enh) ; 0.008 (scaled solar)

ALMA/IRAM use case



- RA/Dec/Freq CO cube
  - Convert to velocity (LSR, radio convention)
  - Cutouts, simple squashes - VO tools?
  - Smoothed spectra, moments with noise cut-off
    - Specialised server-side pipeline controlled via UWS

Objections to VO

Data quality – garbage in - garbage out

How and whom to give credit ? (button)

embedded ivo:// data in ApJ

VO for dissemination only

- technology for OPTICON, nextgen
- Virtual science – VO technology

VO only for public data ! Proprietary ?

(data jealousy)

- local archive - available data marked

Other VOs

Virtual Solar Observatory

Virtual Solar-Terrestrial Observatory

Virtual Magnetospheric Observatory

Virtual Space Physics Observatory

Virtual Meteor Observatory

Interest of climatology, meteorology

New branch of Science = e-Science

Democratization of Science

- Digital Divide (data access free, journals ?)
- International Council for Science (ICS UNO) CODATA (Committee on Data for Science and Technology)
- OECD, UNESCO
- CASPAR

Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval

Digital curation centers

ADS and VO (links to ivo://, metadata, ontologies – understanding, semantic web)

Archive importance: 5x IUE , 3x HST results from archives than PI articles

Effectivity – 50% of published data appears in Journals, links to data automatic ?

The Astronomer's Data Manifesto

at 26 IAU GA Prague SPS3

(a) All significant tables, images, and spectra published in journals should appear inastronomical data centres.

(b) All data obtained with publicly-funded observatories should, after appropriate pro-prietary periods, be placed in the public domain.

(c) In any new major astronomical construction project, the data processing, storage,migration, and management requirements should be built in at an early stage of the project plan, and costed along with other parts of the project.

(d) Astronomers in all countries should have the same access to astronomical data and information.

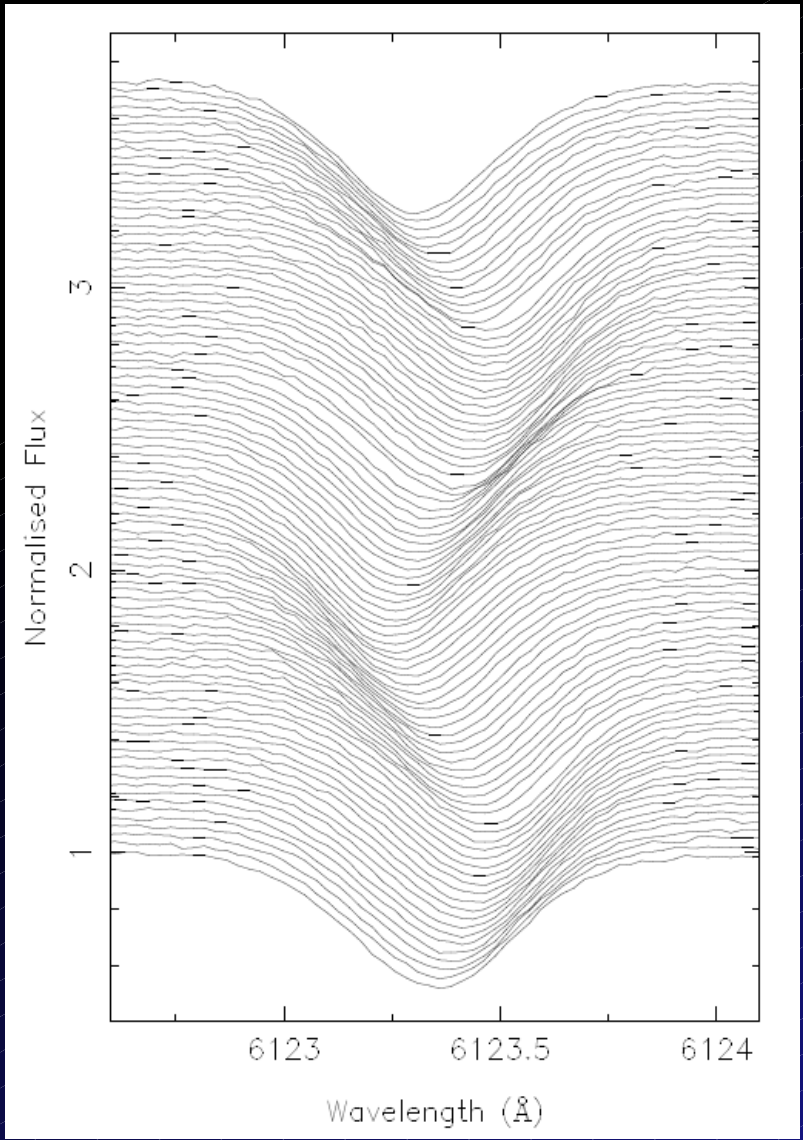
(e) Legacy astronomical data can be valuable, and high-priority legacy data should be preserved and stored in digital form in the data centres.

(f) The IAU should work with other international organisations to achieve our common goals and learn from our colleagues in other fields. ”

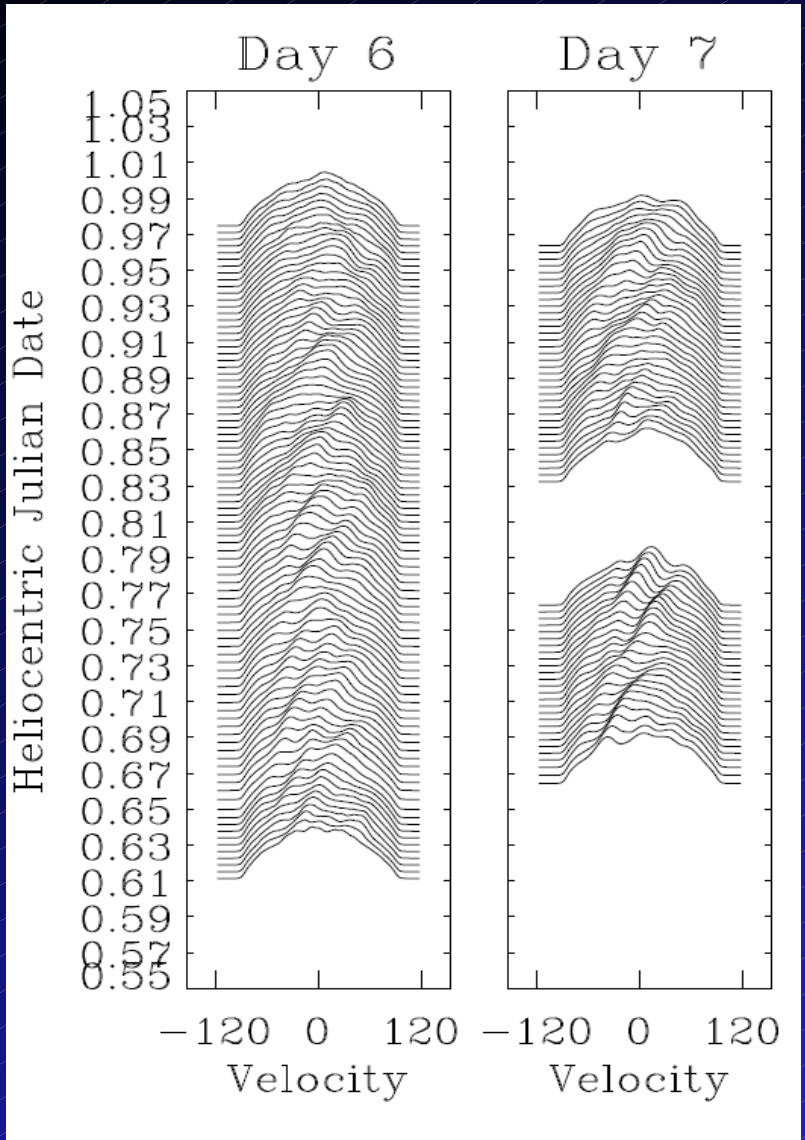
CZVO

- VO-KOREL (web services)
  - parallel run of many jobs – more users
  - web service using VO Universal Worker Server
  - job control, queuing, jobs results polling
  - will be integrated in VODesktop
- 1D spectra cutout server (HEROS)
  - SSA access to 1D spectra + cutout of regions (lines)
  - need normalization, rebinning , convolution (resolution) on server
  - other 1D FITS spectra easy (Rozhen ?)
- Data mining – chaos, NN

Measured Pulsations



Rho Pup – Aerts



Eps Cep - Aerts

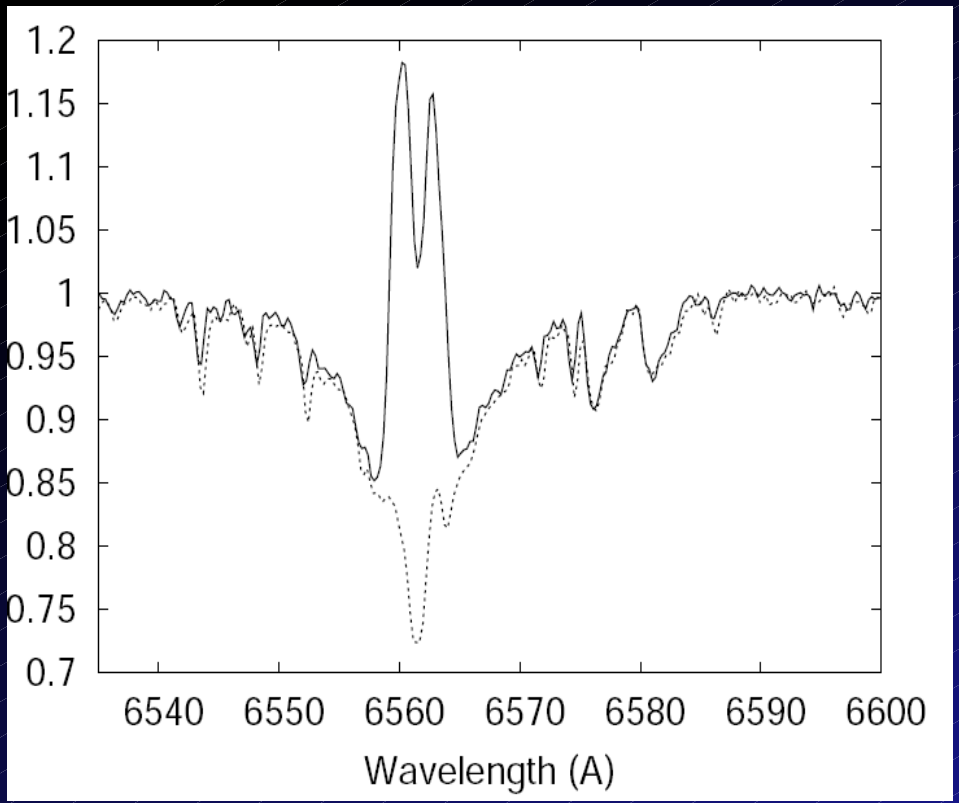
Changes of Line Profiles in Time

Blind comparison of different exposures

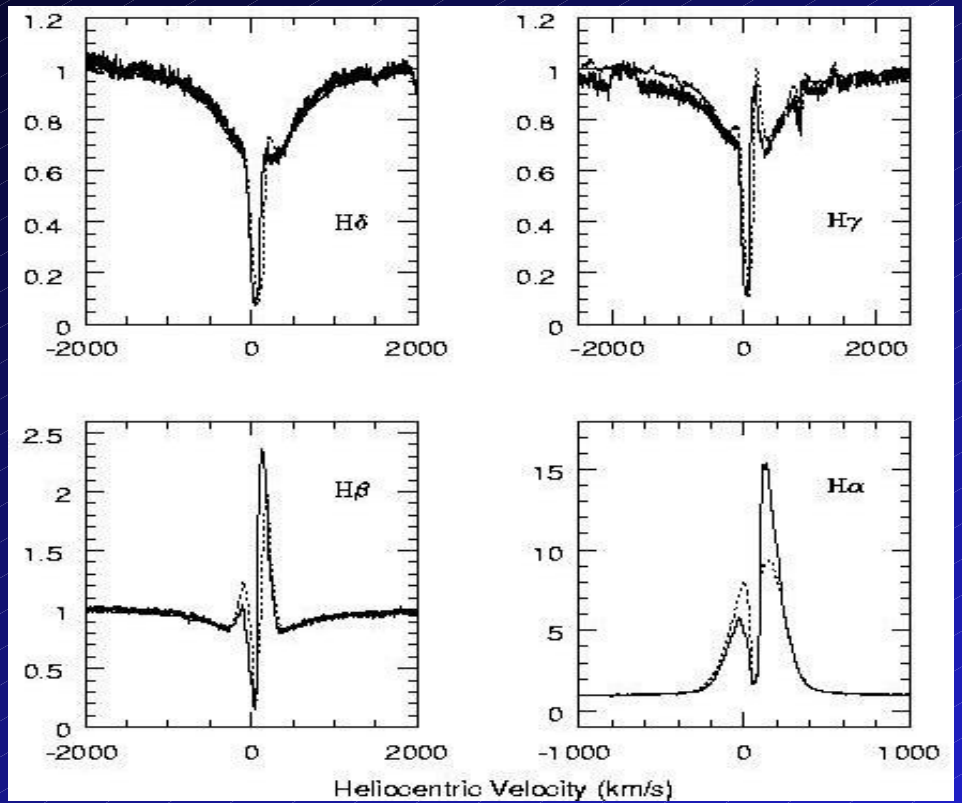
Emission/absorption, shell phases

Time evolution of object – mass transfer,

V/R variations



HD6226 : Slechta and Skoda 2004



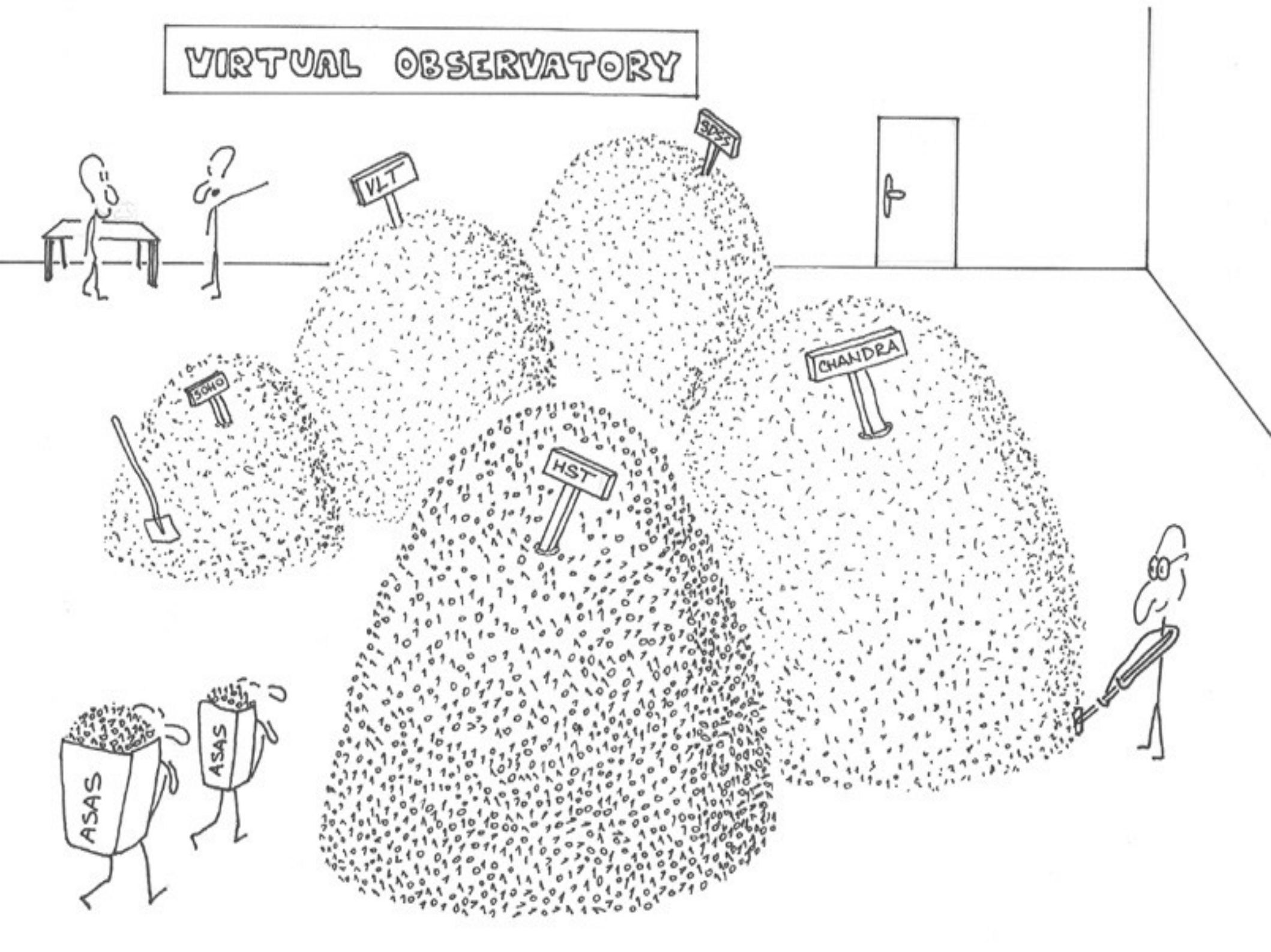
Borges et al. 2008

Spectra postprocessing

- cutout service (ranges in SSAP)
- Rebinning (change od dlambda in echelle)
- Instrument profile (de)convolution
- Broadening functions (rotation, limb dark)
- RV shift
  - on client side (memory, size)
  - on server side (Pleinpot WWW pipeline)

Killer spectral applications

- Use VO to find all stars with emission in given line (EW<0) – find the time when it was in em.
- Use VO to get 1000 spectra of the given object cut out regions arround given lines, plot the lines, make a gray dynamic spectrum folded in time
- The same – search period, fold by period
- Get the unknown line ID of piece of spectra from SLAP overploted over SSA data
- Create Light and RV curve for given period
- Fit the grid of models (Teff, log g) to the observed spectrum – for many stars


VIRTUAL OBSERVATORY
VLT
SDSS
SOHO
HST
CHANDRA
ASAS
ASAS