

Archivy 2m dalekohledu

Od Reticonu k umělé inteligenci

Virtuální observatoř a Astroinformatika

Petr Škoda

Astronomický ústav AVČR
Ondřejov ČR

Posezení k 50. výročí Perkova dalekohledu
Sídlo AVČR, Národní třída, Praha

Sálové počítače

- Teorie + zpracování měření EC1040 (<1MB)
- EC 1045 – 4 MB – terminál sítě – TSO
- Děrné štítky , pásky, mg pásky
FORTRAN66 (EXEC FORTGCLG, FORT SYSIN DD *)



Přechod k PC

- 1988 rekonstrukce TCS 2m – Villati - 1. PC !
 - PC AT 6/10MHZ , 1MB RAM, EGA 640x350, 16c
- RS232 – komunikace s VILATI (TMS9900)
- Disky 8“ 360KB
- Totéž řízení 2m (ASCOL)

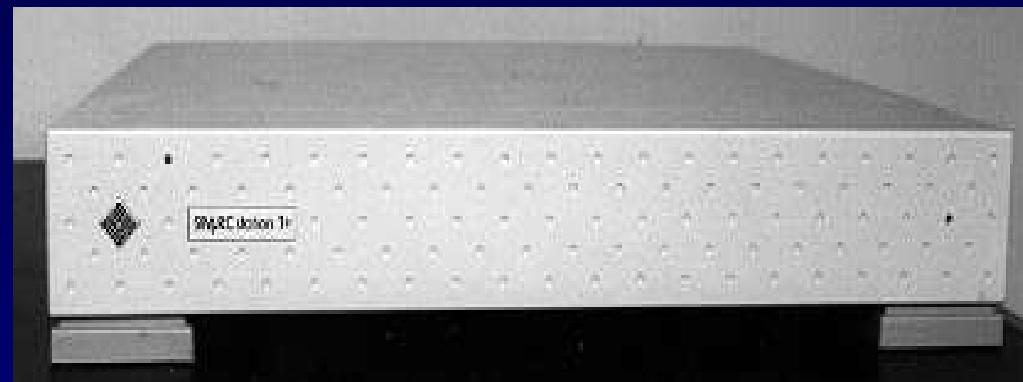


Rozvoj malé výpočetní techniky na AÚ

- 1989 (AÚ – Několik AT, XT, PP06), MS XENIX
- 1990 Několik AT + 386/33 MHz, 4M,
 - 80MB disk
 - Fireball (Ceplecha), PCMOS, Windows 2.0
 - LaTeX (ChiWriter, T602)
- Stelární - první LAN – 10Net (2linka 1Mb)
 - Mezi 3 PC AT

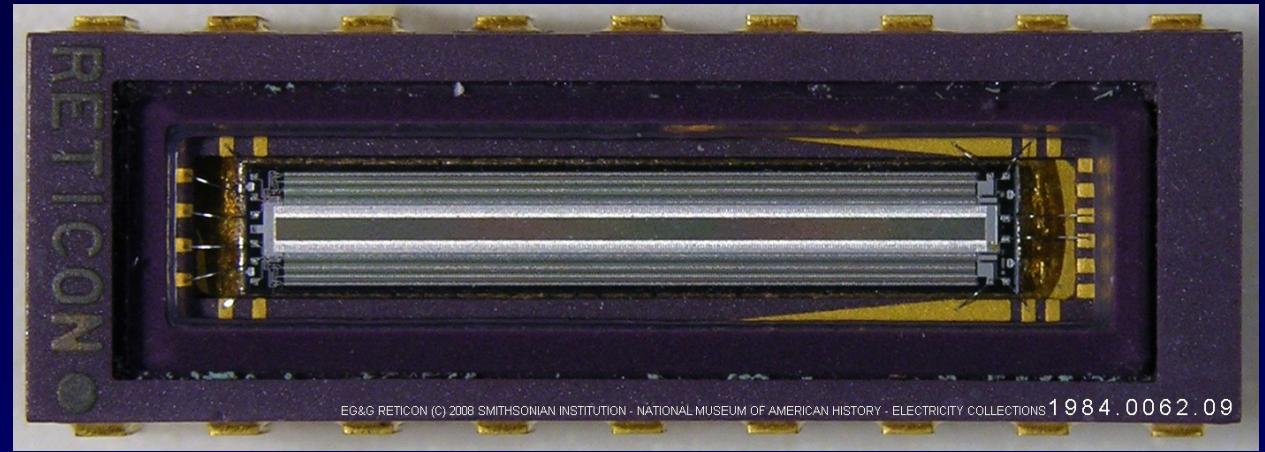
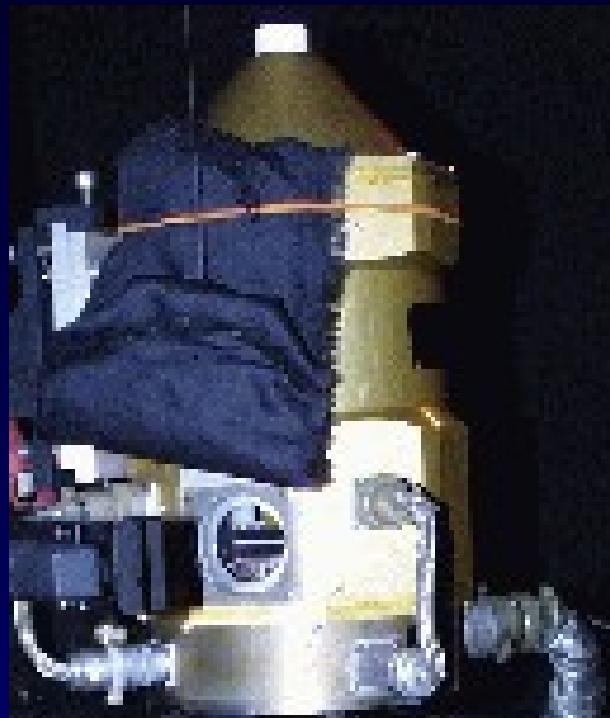
Rozvoj IT 2m - UNIX

- 1991 1 UNIX WS = Sun SPARC 1 CompuAdd
 - SPARC 20MHz, 16MB RAM, 210 MB SCSI
 - Ethernet – WD8003, 10Mb, koax , T 50 Ohm
 - PCTCP (DOS) + Syntax Lan Manager (SMB)
 - TCPIP LAN
 - Několik 486/66 + modem 2400 baud
 - Mail server – e-MAIL na IBM VM/SP Dejvice
 - EARN



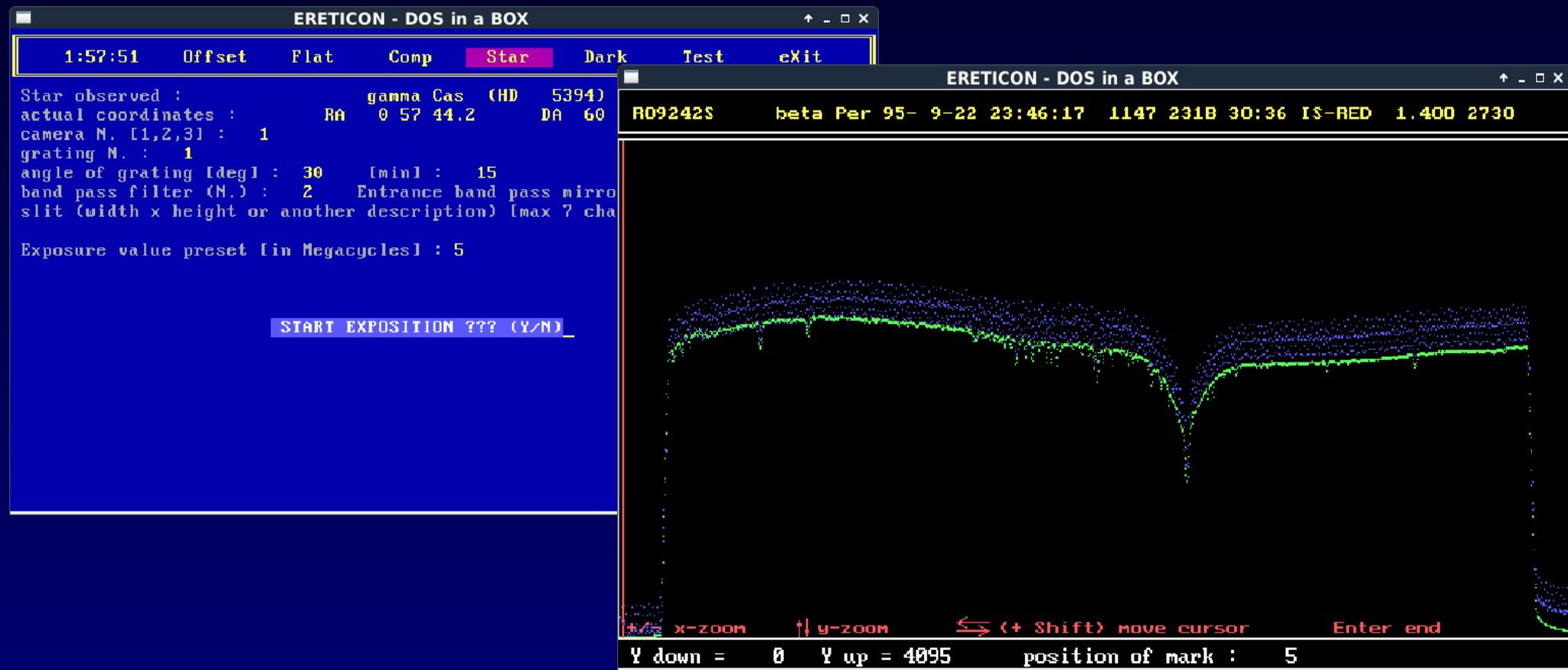
Začátek digitální éry 2m

- 1992 Reticon 1872 AF (Lickova obs., S Vogt)



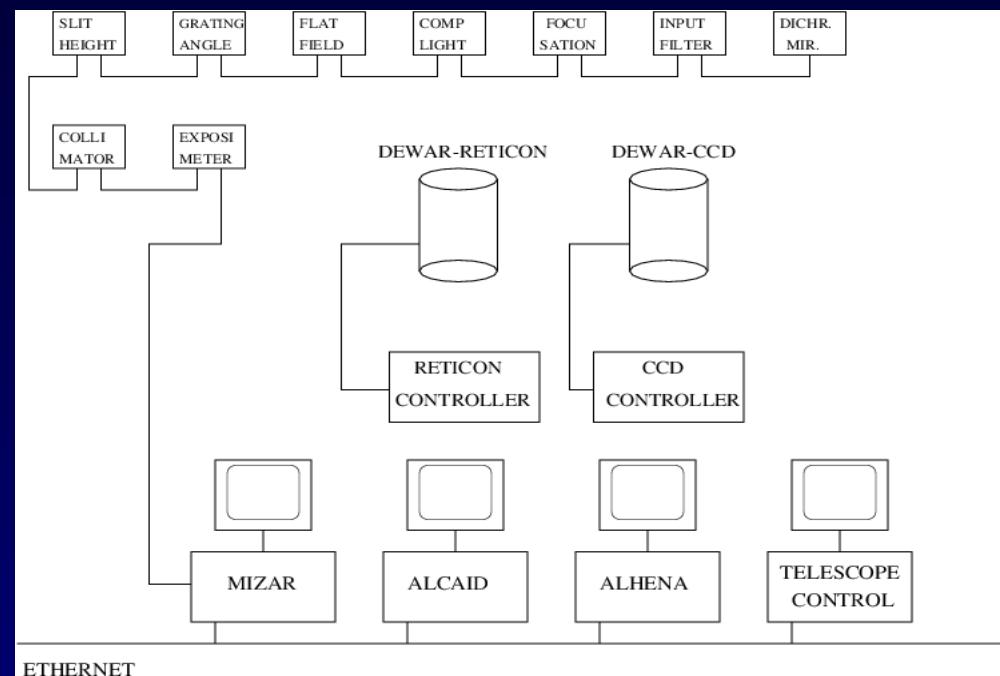
Řídící SW Reticonu

- Modifikace SPEFO (1988)
- Horn - RETICON.EXE, prosinec 1994 umírá
- ERETICON.EXE



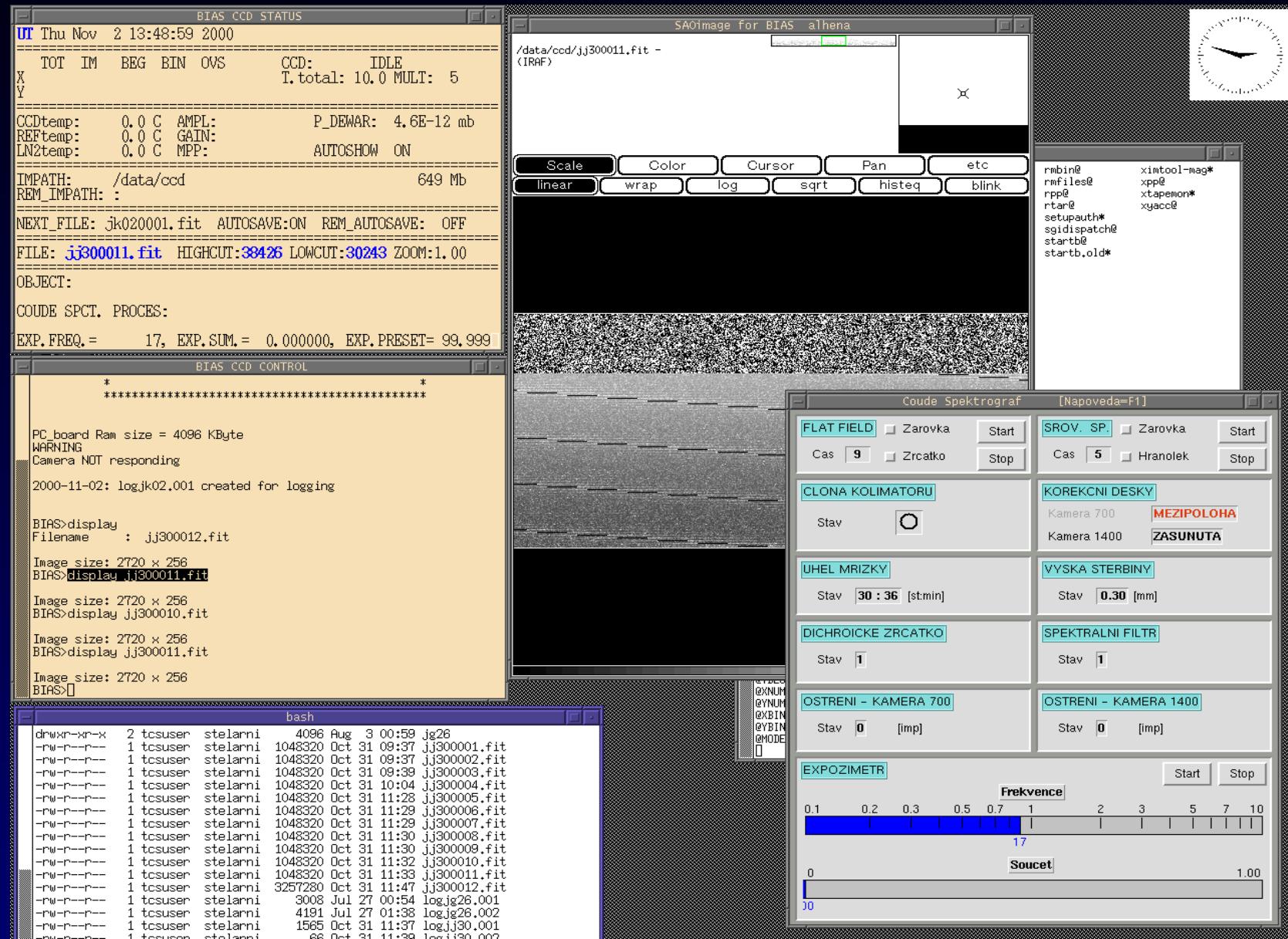
Robotizace spektrografu 1996

- CRETICON.EXE
 - řízení coudé krabiček – Ing. Wudia
 - váhovaní seeingu , zaostření cross-korelací komunikace DOS-UNIX , centrální archív



BIAS -CCD LORAL 2720x256 – od 1996

SiTe 2000x800 -od 2000



Archiv součást RETICONu

ERETICON - DOS in a BOX														
Select star for which you wish the list of observations :														
Current obs log file : Reticon.obs														
Esc : select no star														
alfa And	358	7	2849	9256	1950	0	05	48	28	48	54			
hd 358	358	13	5899	5960										
34 Psc	560	7	7147	9117	2000	0	10	02	11	08	45			
gamma Peg	886	15	732	892	2000	0	13	14	15	11	0			
zeta Cas	3360	2	6508	8594	2000	0	36	58	53	53	49			
alpha Cas	3712	70	2525	9229	2000	0	40	31	56	32	0			
omicron Cas	4180	16	315	8756	1950	0	41	56	48	00	40			
gamma Cas	5394	188	3022	9270	1950	0	53	40.3	60	26	47			
phi Per	10516	69	694	9223	1950	1	40	30.8	50	26	16			
1 Per	11241	10	3560	7572	2000	1	51	59	55	8	51			
beta Ari	11636	5	2859	6432	1900	1	49	06	20	19	00			
alpha Ari	12929	9	6512	7391	2000	2								
4 beta Tri	13161	2	2855	6117	1900	2								
HR 894	18552	3	8727	9164	1950	2								
beta Per	19356	54	3517	9242	1950	3								
29 Per	20365	2	2894	2948	1950	3								
psi Per	22192	4	5688	8823	2000	3								

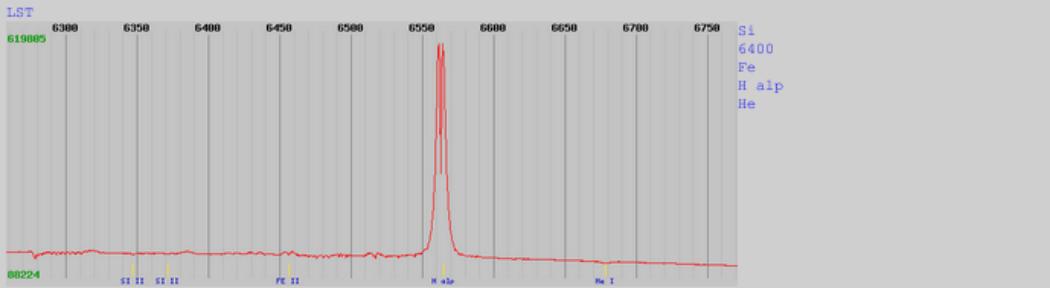
ERETICON - DOS in a BOX														
List of observations of star 1 Per (HD 11241)														
RA(2000) 1 51 59.0 Dec(2000) 55 8 51														
N.	date & UT start	exp[s]	angle	slit	ADU	J.D.hel.	RVcorr	observ						
3560	93-12- 5 18:44:47	10000	231B	30:35	IS-RED	1598	49327.3429	-9.99	juzka					
3615	93-12-29 21: 8: 4	6791	231B	30:35	IS-RED	647	49351.4227	-17.79	HorKa					
6047	94- 9-23 21:57:39	11670	231B	30:35	IS-RED	2210	49619.4855	16.59	KouHa					
7027	94-11-15 17: 8:10	6345	231B	30:36	IS-RED	1530	49672.2551	-2.00	SkoTl					
7112	94-12- 1 16:30:18	10375	231B	30:36	IS-RED	1042	49688.2518	-8.28	SimHa					
7203	94-12- 6 18: 2:58	7422	231B	30:36	IS-RED	812	49693.2989	-10.20	KouTl					
7256	94-12- 8 17:39: 4	3138	231B	30:36	IS-RED	222	49695.2575	-10.87	HecHa					
7303	94-12-15 18:30:21	7455	231B	30:36	IS-RED	269	49702.3178	-13.38	HecKa					
7508	95- 2-12 21: 4:40	7201	231B	30:36	IS-RED	649	49761.4195	-23.12	Koutl					
7572	95- 2-18 17:37: 9	5945	231B	30:36	IS-RED	570	49767.2677	-22.75	HecHa					

List saved as file "1Per.lst" press any key_

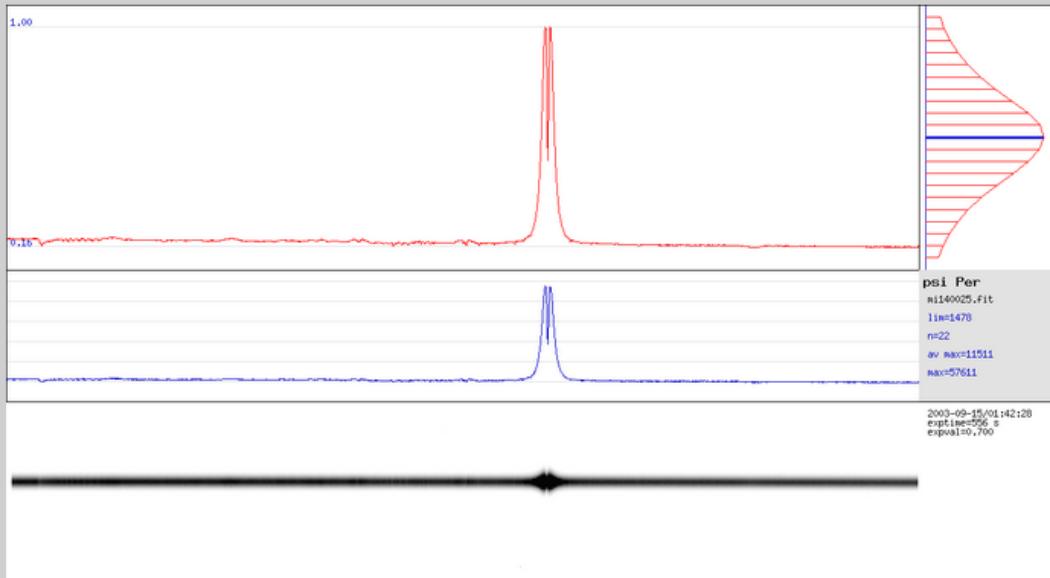
Spectra archive – raw + reduced

H alpha	2008-09-19/21:17	tj180005	a 30:15 2/1	730	15.0888 37-39	20100718	psi Per	[Ceniga, Fuchs]
H alpha	2010-07-19/01:43	tj180048	a 30:15 2/1	150	4.388 42-43	20100718	psi Per	[Kraus, Kotkova]
gaia	2010-09-04/20:17	tj040007	a 35:47 2/3	300	2.789 25-25	20100904	psi Per	[Kraus, Kotkova]
H alpha	2010-09-04/20:44	tj040012	a 30:15 2/1	200	3.755 28-28	20100904	psi Per	[Kraus, Kotkova]
gaia	2011-03-15/18:49	uc150011	a 35:47 2/3	900	16.881 54-52	20110315	psi Per	[Kraus, Sloup, Aret]
Ca II	2011-03-15/19:08	uc150013	a 32:8 2/3	1800	32.073 51-47	20110315	psi Per	[Kraus, Sloup, Aret]
H alpha	2011-12-07/20:21	wl070006	a 30:15 2/1	800	9.562 79-81	20111207	psi Per	[Kraus, Kotkova]
IR	2011-12-07/20:54	wl070023	a 36:0 2/3	1132	11.906 84-87	20111207	psi Per	[Kraus, Kotkova]
H alpha	2012-01-30/17:13	va300019	a 30:15 2/1	130	2.622 82-83	20120130	psi Per	[Koubsky, Fuchs]
H alpha	2014-08-15/09:14	a201408140051	a 30:15 2/1	1201	15.420 45-48	20140814	psi Per	[Kraus, Tlanicha]
H alpha	2014-08-15/09:36	a201408140053	a 30:15 2/1	901	11.149 48-50	20140814	psi Per	[Kraus, Tlanicha]
gaia	2014-08-15/09:54	a201408140055	a 35:47 2/3	1801	19.944 51-55	20140814	psi Per	[Kraus, Tlanicha]

List of files – main metadata



Reduced spectrum (if exists)



Raw spectrum – maximum in COP

Object: 'psi Per'
File: m140025
Date start: '2003-09-15'
UT start: '01:42:28'
Telescope coordinates: '3:36:08.8' '48:12:14.3'
Object coordinates: 3:36:29 48:11:33
Observers: 'Stefl + Tlanicha'
HD: 2452897.5760
Spectrograph configuration: CCD700 30:15 1/1 (10889)
Detector: 'SITe 2000x800'
Spectral range [Å]: 6258 - 6771
Exposure [s]: 596
Expval: 0.700
Integ: 7184
Vhelio [km/s]: 24.78
checked - OK
Observers: Stefli, Tlanicha

Raw spectrum – averaged all COP

Raw spectrum – inverse image

All metadata about frame

Internet v astronomii

- 1992 únor – Internet v ČSFR (EARN-> FERNET,CESNET)
- Na ASU – cca 1993 duben (brána EARN)
- Gopher, WAIS, Veronica, začátek WWW, Telnet
- Duben 1995 CDS Strasbourg

Weaving the Astronomy Web

- Použití WWW pro astronomii
- Databáze ADS
- Astrophysical Journal – on-line
- Proroctví o konci tištěné verze za 20 let
- NCSA Mosaic -> Netscape 1.0

WWW Archiv Reticonu

- ADASS – říjen 1995 – Tucson
 - komunita na celý život -SW, detektory, algoritmy
 - Školení IRAF, CCD laboratoř , prohlídka KPNO
- 1998 GRANT MŠMT INFRA2 LB98251
 - Budování archivu Reticonu
 - SQL Adabas, PostgresQL, WDB (ESO 1995)
 - 2000 - Archiv 2M byl jedním ze 6 světových archivů pozemních dalekohledů na WWW
 - ESO, CFHT, AAO Okyama, LaPalma INT
 - +logy Asiago, OHP ELODIE
- 2005 Šárka Zehnalová UJEP - do Java Struts

Virtual Observatory : Key Definitions

- “*The Virtual Observatory will be a system that allows astronomers to interrogate multiple data centers in a seamless and transparent way, which provides new powerful analysis and visualization tools within that system, and which gives data centers a standard framework for publishing and delivering services using their data*”.
- Standardization of data and metadata, and of data exchange methods.
- Registry, listing available services and what can be done with them.

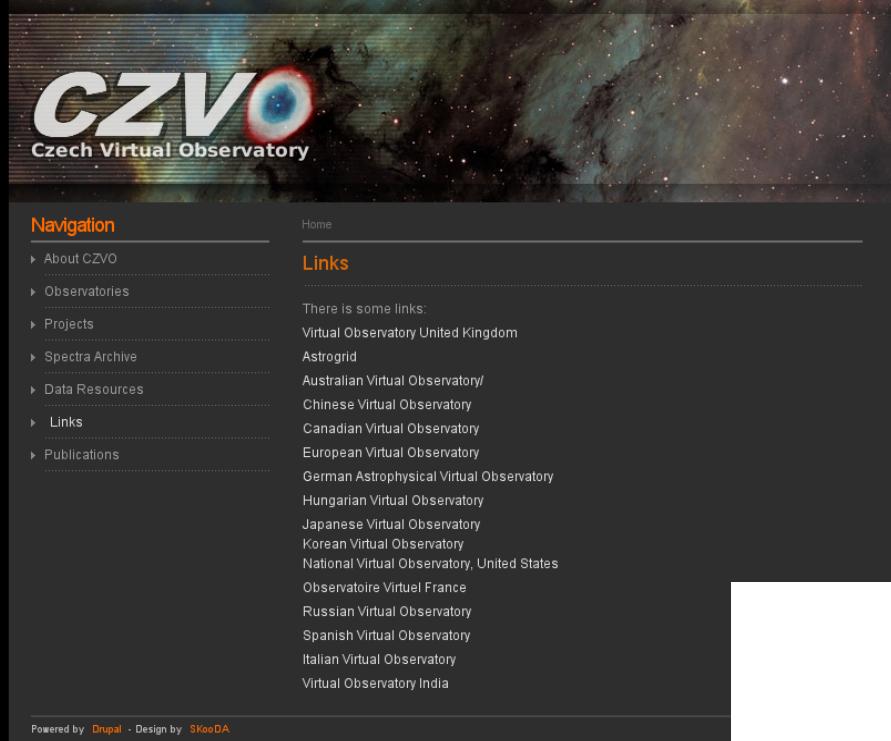
R.J.Hanisch, P.J.Quinn, in “IVOA – Guidelines for participation”

IVOA



Czech VO - CZVO

Od 2006 po IAU



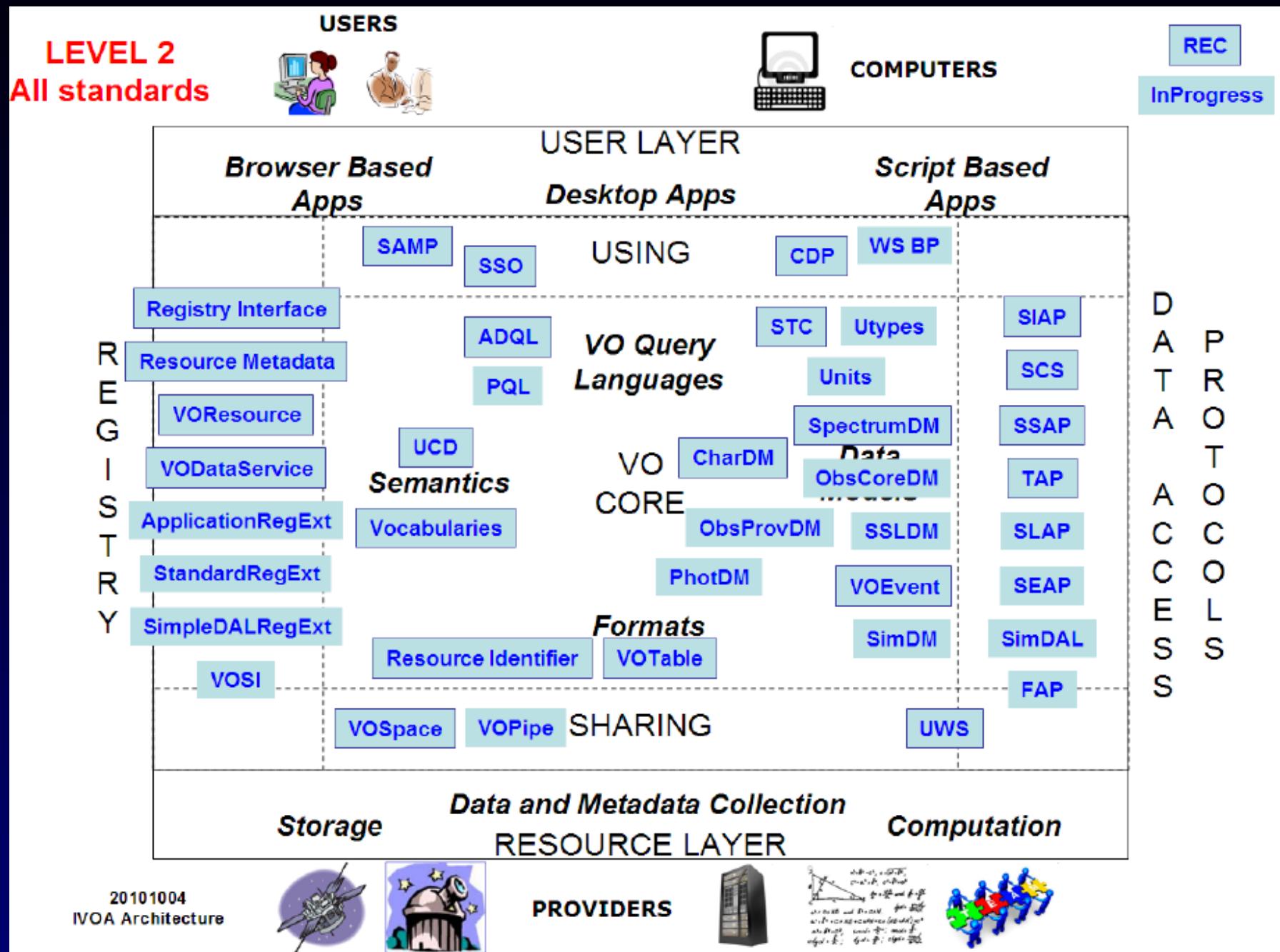
The screenshot shows the CZVO website homepage. At the top, there is a navigation bar with links for "Search" and "Login". Below the header, the CZVO logo is displayed against a background of a star-filled galaxy. The main content area has two columns: "Navigation" on the left and "Links" on the right. The "Navigation" column includes links to "About CZVO", "Observatories", "Projects", "Spectra Archive", "Data Resources", "Links", and "Publications". The "Links" column lists various international Virtual Observatory websites, such as "Virtual Observatory United Kingdom", "Astrogrid", "Australian Virtual Observatory", "Chinese Virtual Observatory", "Canadian Virtual Observatory", "European Virtual Observatory", "German Astrophysical Virtual Observatory", "Hungarian Virtual Observatory", "Japanese Virtual Observatory", "Korean Virtual Observatory", "National Virtual Observatory, United States", "Observatoire Virtuel France", "Russian Virtual Observatory", "Spanish Virtual Observatory", "Italian Virtual Observatory", and "Virtual Observatory India". At the bottom of the page, there is a footer note: "Powered by [Drupal](#) - Design by [SkooDA](#)".

Podpora EURO-VO

2008 HEROS archive
ve VO Prugniel OHP
Cutout server



Ecosystem of VO - level 2



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.

VO access to spectra - Browser



AIASCR VO Services

Welcome to ASU CAS Data Center.

In addition to the services listed below, on this site you probably can access numerous tables using TAP or form-based ADQL.

Please check out our [site help](#).

This project was supported by grant 13-08195S of Czech Science Foundation.

Services Available

[By Title](#)

[By Subject](#)

[By Author](#)

- [CCD700 Spectra Web Interface](#)
- [HEROS Public Spectra Web Interface](#)
- [LAMOST DR1 Spectra Web Interface](#)

Spectra query - web browser

CCD700 Spectra Web Interface

CCD700 public web interface.

Object standard name

No selection matches all, multiple values legal.

Object standard name dropdown menu:

- pmauru
- phicas
- phiper
- piaqr
- psi09aur
- psiper
- pugem
- pgem
- qqgem
- qrval
- nsan

Somewhat cleaned name of the target object as given in the file. The names should either be SIMBAD-resolvable or give a ref they are appreciated.

Location

Coordinates (as h m s, d m s or decimal degrees), or SIMBAD-resolvable object

Search radius for Location [arcmin]

Search radius in arcminutes

Date Obs.

[?date expr.]

Midpoint of exposure

Type

image/fits
 application/x-votable+xml

MIME type of the file served

Table

Sort by Limit to 100 items.

Output format

HTML Pop down field selector

checkbox list:

- accsize -- Size of the data in bytes
- detector -- Detector used to capture spectrum
- embargo -- Date the data will become/became public
- expval -- Photon counts (Mcount)
- grat_angle -- Grating tilt
- instrument -- Instrument used to capture spectrum
- mime -- MIME type of the file served
- obj_name -- Unique identifier of 1 object.
- owner -- Owner of the data
- spg_setup -- Spectrograph setup (Dichroic mirror/Spectral filter)

Spectra - query output and previews

CCD700 Spectra Web Interface

Parameters

- Object standard name: ['psiper']

Result

Matched: 44

[Send via SAMP](#) [Quick Plot](#)

Product key	Object	Raj2000	Dej2000	Band start [Angstrom]	Band end [Angstrom]	Date Obs.	Observer	Exp. Time [s]	MHJD
tg180048.fit	psi Per	03:36:29.380	+48:11:33.40	6261.00	6773.40	2010-07-19T01:43:36Z		150.0	55396.0696765
	psi Per	3:37:01.1	48:12:17.1	6262.34	6774.66	2005-03-22T21:16:04Z	Kubat, Sarounova	899.564	53451.8889388
oc220022.fit									
	Psi Per	03:36:29.380	+48:11:33.40	6252.67	6764.96	2011-12-07T20:21:43Z	Kraus, Kotkova	800.0	55902.8579726
ul070006.fit									
va300019.fit	psi Per	03:36:29.380	+48:11:33.40	6252.09	6764.42	2012-01-30T17:13:10Z		130.0	55956.7202446
mi180074.fit	psi Per	3:37:03.2	48:10:44.2	6259.63	6772.03	2003-09-19T02:34:39Z		60.0	52901.1096669
ng290040.fit	Psi Per	3:36:46.8	48:11:39.4	6261.83	6774.26	2004-07-30T00:40:28Z	Kubat, Kalas	600.0	53216.0292701
	psi Per	3:36:42.4	48:13:02.0	6264.34	6776.77	2004-08-11T01:02:47Z	Stefl + Rezna	218.373	53228.0435205
nh100015.fit									
	Psi Per	3:36:46.3	48:13:02.4	6264.01	6776.40	2004-09-01T00:22:31Z	Libich, Sarounova	60.0	53249.016454
nh310030.fit									

[Try ADQL](#) to query our data.

Please report errors and problems to the [site operators](#). Thanks.

[Privacy](#) | [Disclaimer](#)

[Log in](#)

Spectra in SPLAT-VO - query

Starlink SPLAT-VO: Query VO for Spectra

Service selection options:

- Observed data Theoretical data
- Radio Millimeter Infrared
- Optical UV EUV
- X-ray Gamma-ray ALL

Search parameters:

Simple Query

Object:	psi per	Lookup	Clear
RA:	03:36:29.38	Dec:	+48:11:33.48
Radius:	10.0	MAXREC:	
Band:	6530e-10	/6580e-10	
Time:			
Query Format:	votable		
Wavelength calibration:	None		
Flux calibration:	None		

Optional Parameters:

Use	Name	Value	UCD
<input type="checkbox"/>	REDSHIFT		src.redshift
<input type="checkbox"/>	TARGETCLASS		src.class
<input type="checkbox"/>	MTIME		
<input type="checkbox"/>	SPECRP		spect.resolution;em.wl
<input type="checkbox"/>	SPATRES		pos.angResolution
<input type="checkbox"/>	PUBBDID		
<input type="checkbox"/>	CREATORID		meta.id

Buttons:

- Select all
- Deselect all
- Update
- SEND QUERY

Query:

```
<SERVER>?REQUEST=queryData&POS=54.12241666666666,48.19263333333333&FORMAT=votable&SIZE=0.1666666666666666&BAND=6530
```

Query results:

CCD700-voarchive

I...	ssa_specstart	ssa_specend	ssa_dstitle	ssa_targname	ssa_dateObs	ssa_timeExt	ssa_snr	ssa_length	acref
12	6.26183E-7	6.77426E-7	ccd700/data/psiper/6255-676...	Psi Per	53216.0049	1800.		1997	http://voarchive.asu.cas.cz/ge... appli
27	6.24978E-7	6.76217E-7	ccd700/data/psiper/6255-676...	Psi Per	54701.97209	1333.19		1997	http://voarchive.asu.cas.cz/ge... appli
13	6.25854E-7	6.77081E-7	ccd700/data/psiper/6255-676...	Psi Per	52982.77964	1300.		1997	http://voarchive.asu.cas.cz/ge... appli
31	6.26359E-7	6.73614E-7	ccd700/data/psiper/6260-673...	Psi Per	56884.01015	1201.		2047	http://voarchive.asu.cas.cz/ge... appli
30	6.26359E-7	6.73614E-7	ccd700/data/psiper/6260-673...	Psi Per	56884.02565	901.		2047	http://voarchive.asu.cas.cz/ge... appli
2	6.26234E-7	6.77466E-7	ccd700/data/psiper/6255-676...	Psi Per	53451.88616	899.564		1997	http://voarchive.asu.cas.cz/ge... appli
28	6.26587E-7	6.77822E-7	ccd700/data/psiper/6255-676...	Psi Per	54209.80019	899.214		1997	http://voarchive.asu.cas.cz/ge... appli
3	6.25267E-7	6.76496E-7	ccd700/data/psiper/6255-676...	Psi Per	55902.84843	800.		1997	http://voarchive.asu.cas.cz/ge... appli
10	6.25163E-7	6.76402E-7	ccd700/data/psiper/6255-676...	Psi Per	54757.88729	730.		1997	http://voarchive.asu.cas.cz/ge... appli
6	6.26183E-7	6.77426E-7	ccd700/data/psiper/6255-676...	Psi Per	53216.02811	600.		1997	http://voarchive.asu.cas.cz/ge... appli
14	6.24978E-7	6.76217E-7	ccd700/data/psiper/6255-676...	Psi Per	54701.96266	600.		1997	http://voarchive.asu.cas.cz/ge... appli
9	6.25809E-7	6.77051E-7	ccd700/data/psiper/6255-676...	Psi Per	52897.07116	556.599		1997	http://voarchive.asu.cas.cz/ge... appli
18	6.26587E-7	6.77822E-7	ccd700/data/psiper/6255-676...	Psi Per	54209.81274	450.		1997	http://voarchive.asu.cas.cz/ge... appli
20	6.25854E-7	6.77081E-7	ccd700/data/psiper/6255-676...	Psi Per	52982.79788	350.		1997	http://voarchive.asu.cas.cz/ge... appli
16	6.26235E-7	6.77492E-7	ccd700/data/psiper/6255-676...	Psi Per	53475.85182	300.043		1997	http://voarchive.asu.cas.cz/ge... appli
19	6.26401E-7	6.77640E-7	ccd700/data/psiper/6255-676...	Psi Per	53249.01697	300.		1997	http://voarchive.asu.cas.cz/ge... appli
23	6.25161E-7	6.76378E-7	ccd700/data/psiper/6255-676...	Psi Per	54519.80405	300.		1997	http://voarchive.asu.cas.cz/ge... appli
24	6.25161E-7	6.76378E-7	ccd700/data/psiper/6255-676...	Psi Per	54519.7984	300.		1997	http://voarchive.asu.cas.cz/ge... appli
7	6.26434E-7	6.77677E-7	ccd700/data/psiper/6255-676...	Psi Per	53228.0436	218.373		1997	http://voarchive.asu.cas.cz/ge... appli
11	6.25641E-7	6.76875E-7	ccd700/data/psiper/6255-676...	Psi Per	55443.86444	200.		1997	http://voarchive.asu.cas.cz/ge... appli
22	6.25161E-7	6.76378E-7	ccd700/data/psiper/6255-676...	Psi Per	54519.81162	200.		1997	http://voarchive.asu.cas.cz/ge... appli
25	6.25908E-7	6.77147E-7	ccd700/data/psiper/6255-676...	Psi Per	52901.91749	194.574		1997	http://voarchive.asu.cas.cz/ge... appli
15	6.25937E-7	6.77178E-7	ccd700/data/psiper/6255-676...	Psi Per	52904.03609	180.		1997	http://voarchive.asu.cas.cz/ge... appli
1	6.26100E-7	6.77340E-7	ccd700/data/psiper/6255-676...	Psi Per	55396.07196	150.		1997	http://voarchive.asu.cas.cz/ge... appli
4	6.25209E-7	6.76442E-7	ccd700/data/psiper/6255-676...	Psi Per	55055.77740	120.		1997	http://voarchive.asu.cas.cz/ge... appli

Buttons:

- Display selected
- Display all
- Download selected
- Download all
- Deselect table
- Deselect all
- DataLink Services

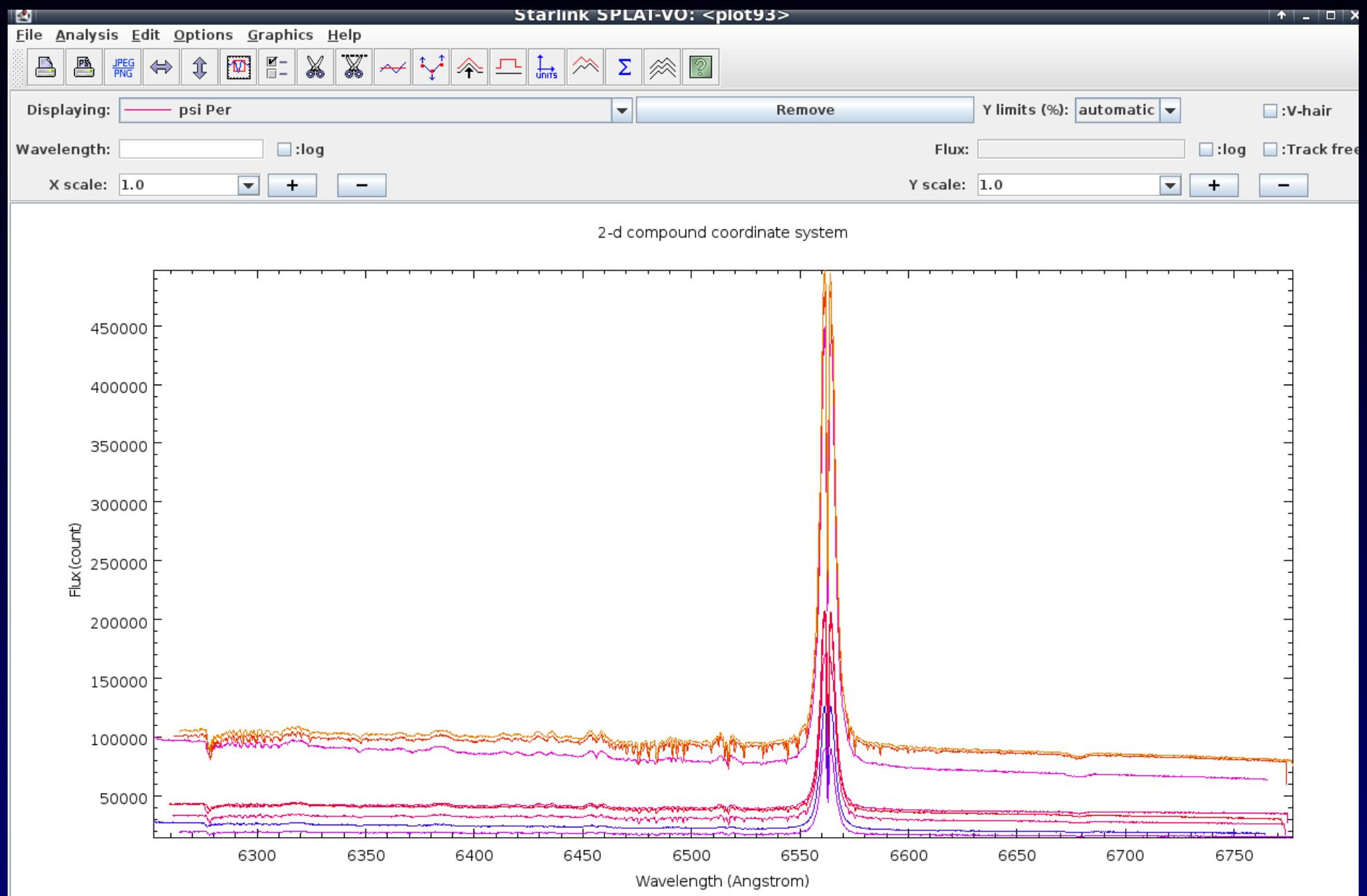
File Buttons:

- Save query results
- Restore query results
- Close

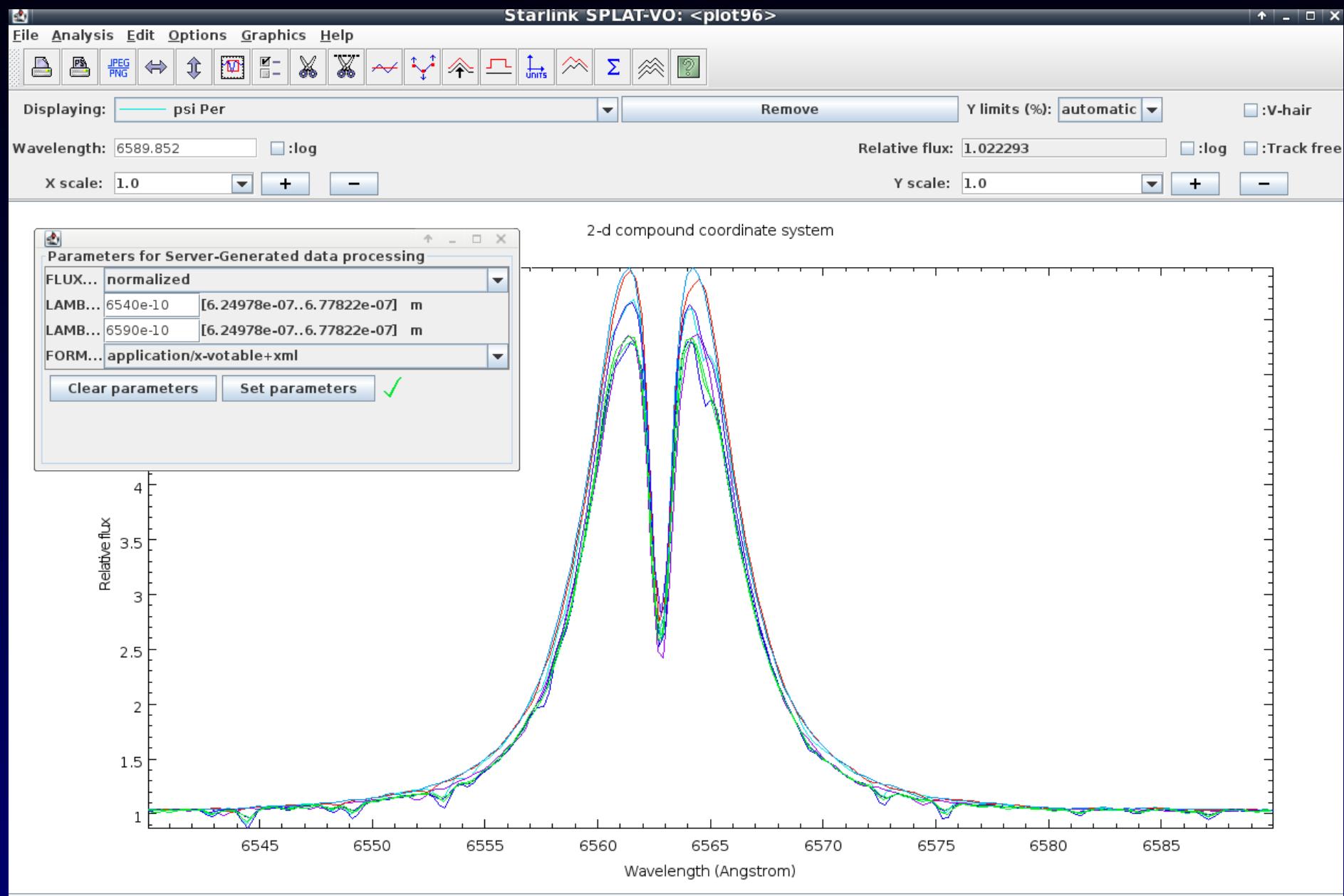
Bottom Buttons:

- Select all
- Deselect all
- Query registry
- Add New Server

Spectra in VO - direct access plot



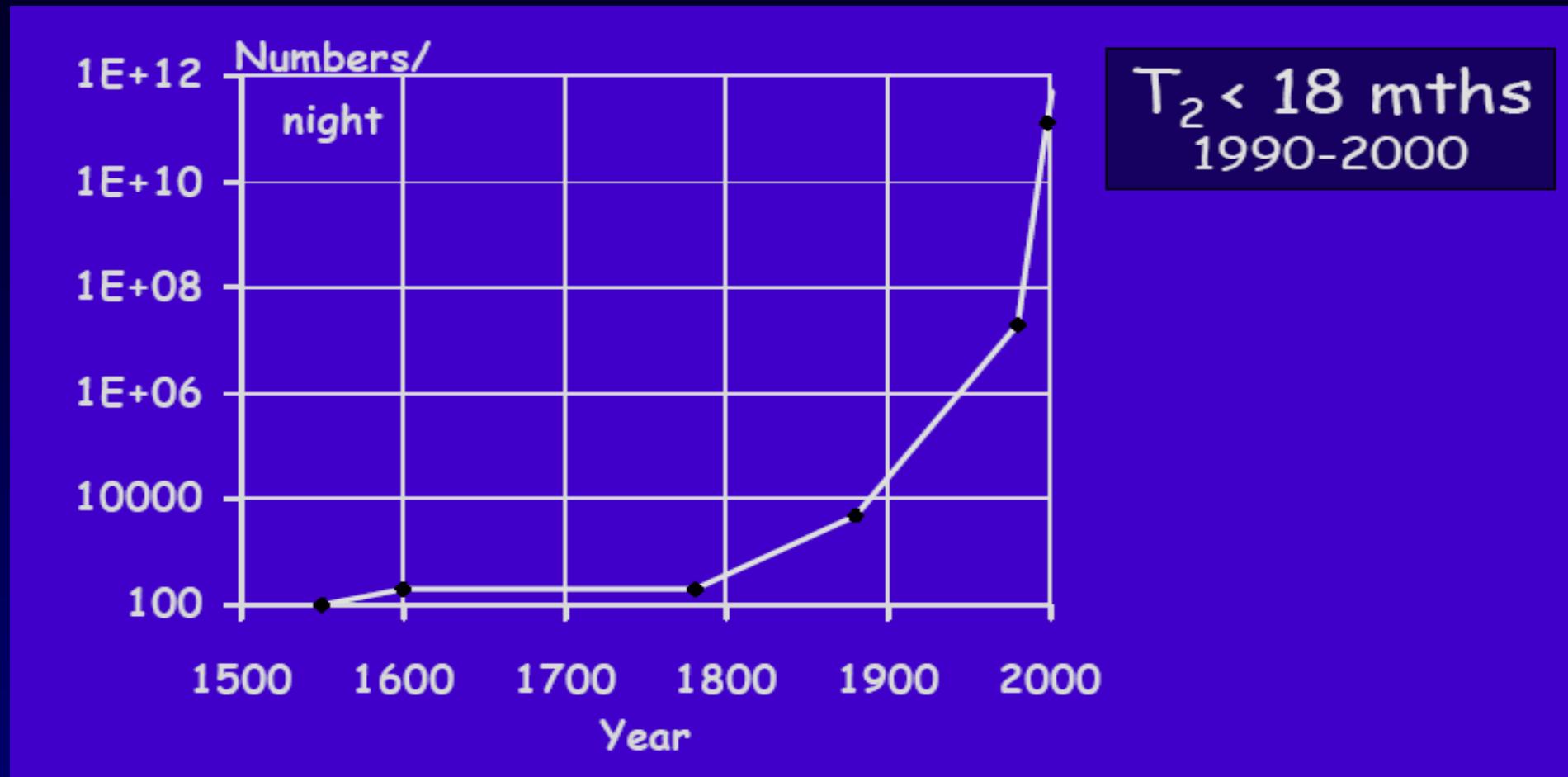
Spectra in SPLAT-VO - DataLink



Lavina dat

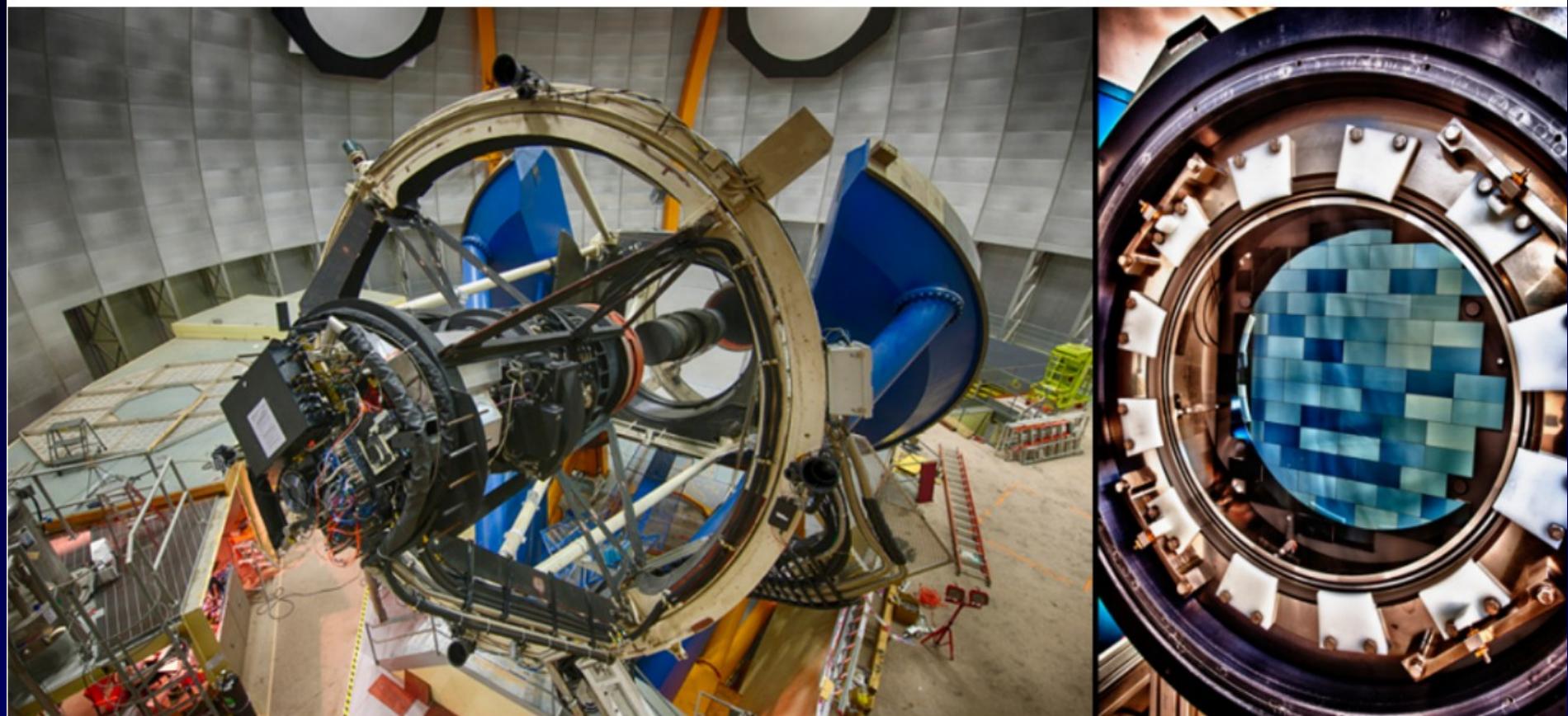
Moorův zákon chipy – zdvojení 1.5 roku (1000/10 let)

Data v astronomii – zdvojení < 1 rok !



Dark Energy Survey Camera

Dark Energy Camera (DECam)



~0.4 PB/yr

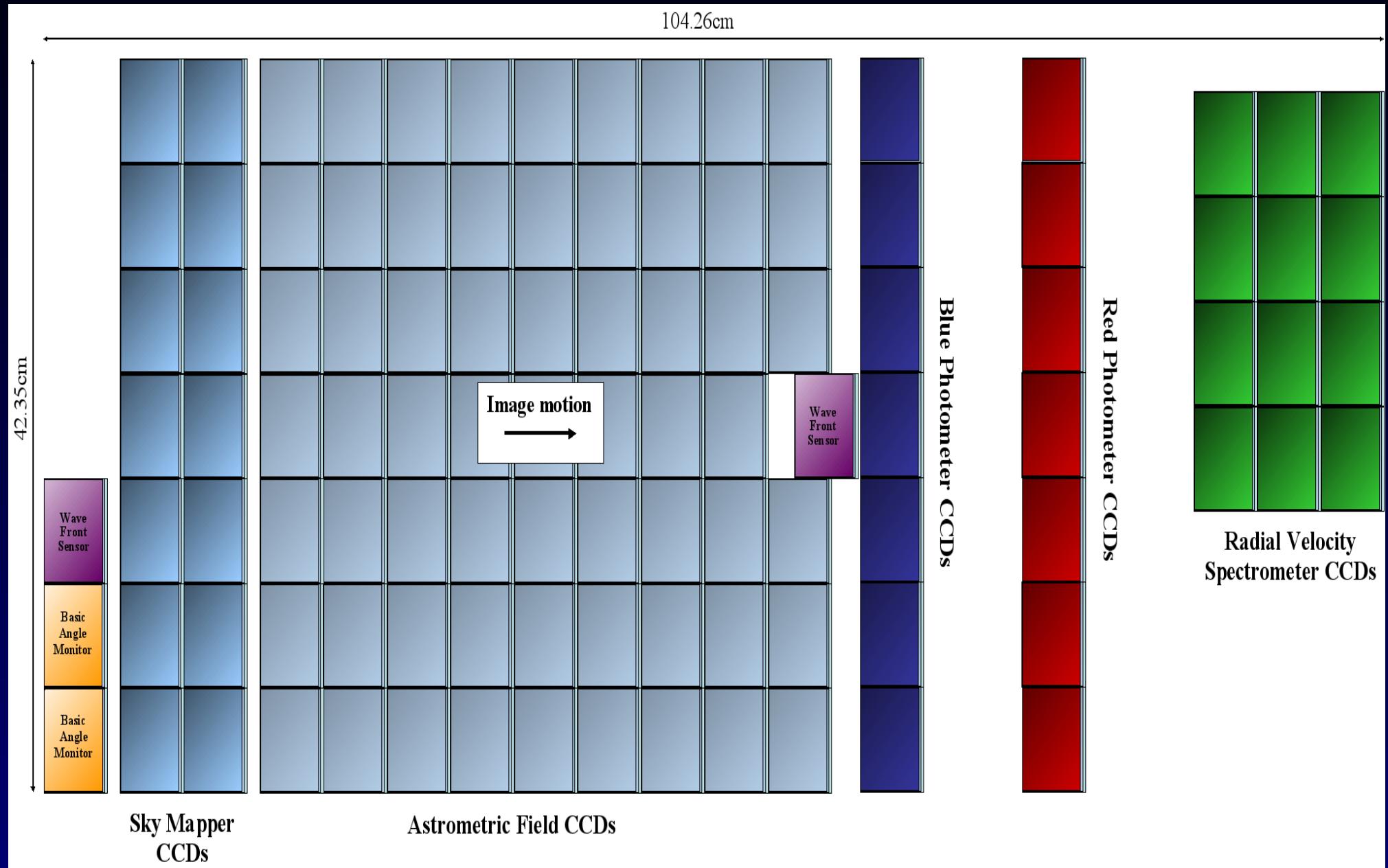
74 chips – 570 Mpixels - 4m Cerro Tololo

GAIA CCDs

106 CCDs

938 Mpix

2800cm²

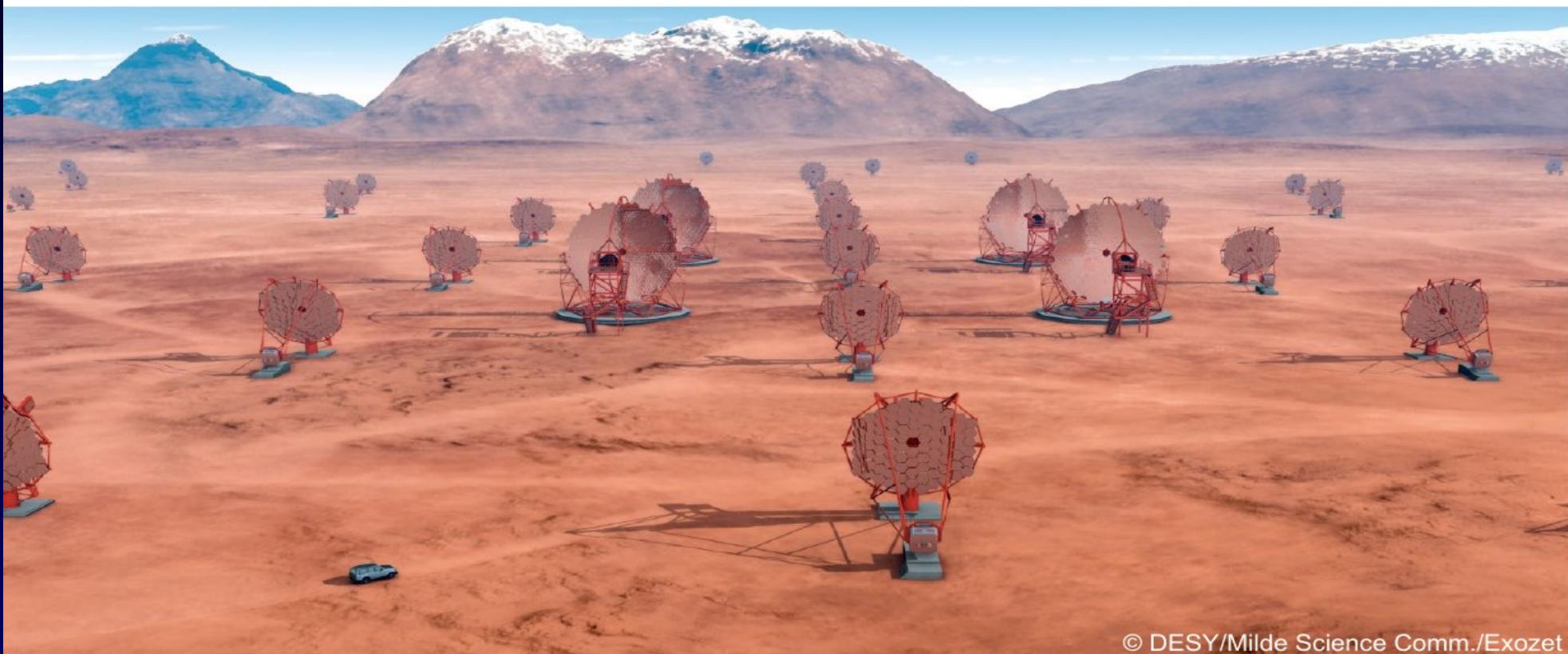


Cherenkov Telescope Array

Cherenkov Astronomy and CTA



- ◆ Two arrays of 100 (South) et 20 (North) telescopes
- ◆ July 2015: sites selection, Chile (ESO) and La Palma
- ◆ 2016: pre-production phase
- ◆ 2018-2013: production phase
- ◆ Observatory open to the community



© DESY/Milde Science Comm./Exozet

LOFAR network



LOFAR

SKA

Raw Telescope	112 PB/yr	60 EB/yr
Archive Rate	6 PB/yr	100 PB/yr

SKA



Dishes

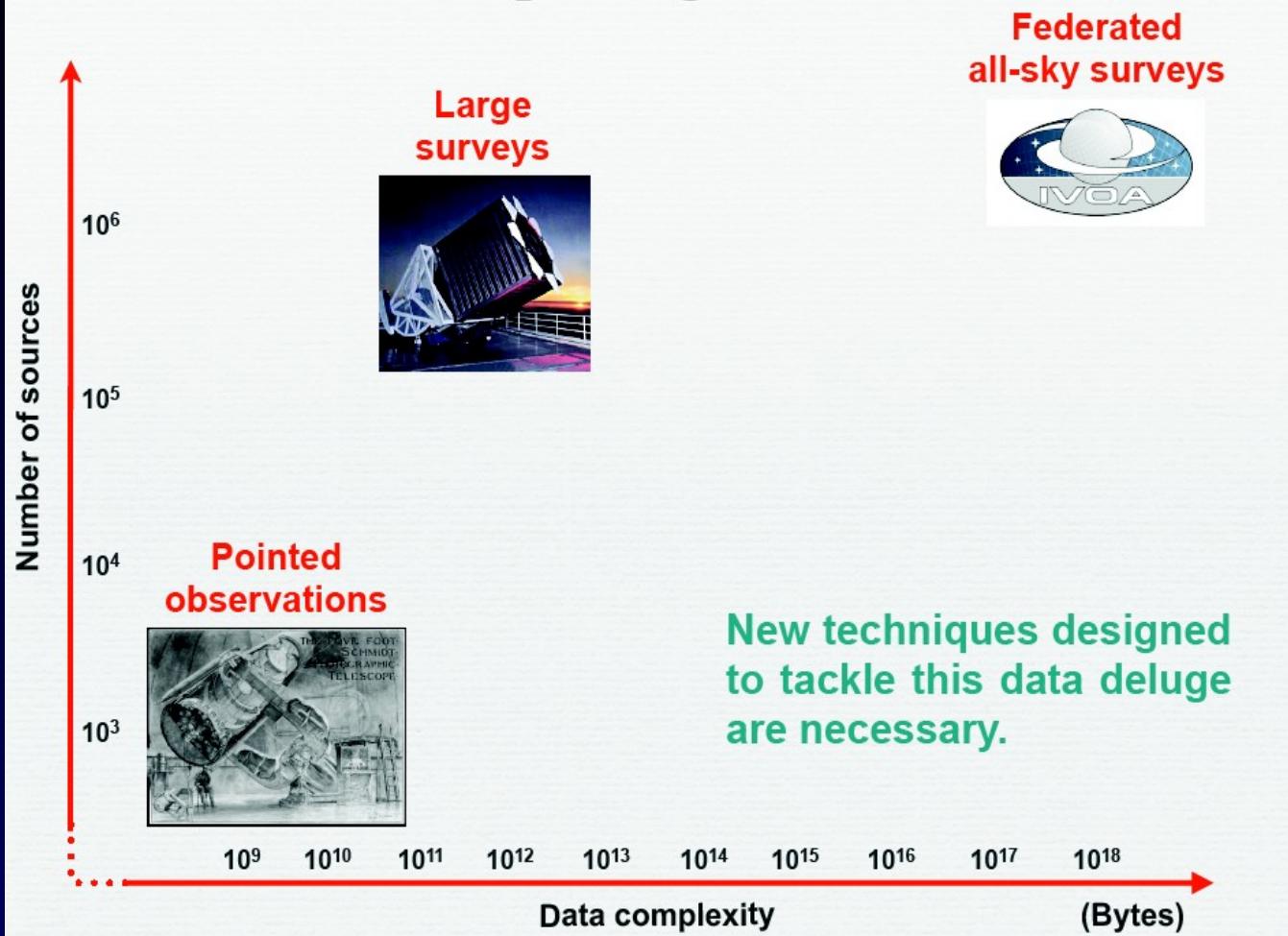
SPDO/Swinburne

SKA Archive Volumes

- ~0.5 – 10 PB/day of image data
- Source count ~ 10^6 sources per square degree
- ~ 10^{10} sources in the accessible SKA sky, 10^4 numbers/record
- ~1 PB for the catalogued data

100 Pbytes – 3 EBytes / year of fully processed data

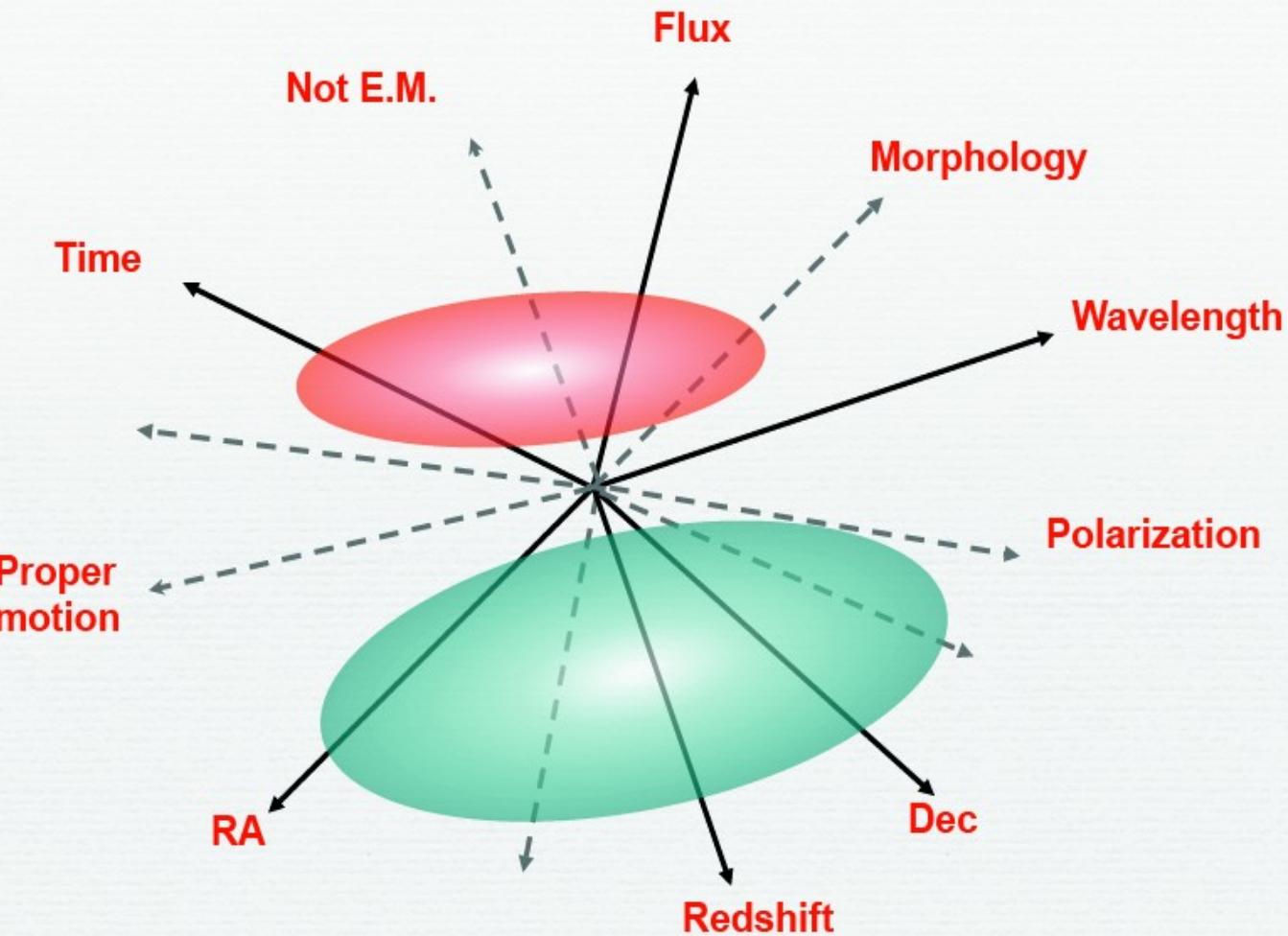
A paradigm shift



D'Abrusco 2010

Data nutno analyzovat na místě uložení
Move processing = not data !

A growing parameter space



**Most discoveries were made in small regions
of subspaces or along some of these axes**

Need for a new science: Astroinformatics

Knowledge Discovery in Databases

Data Gathering (e.g., from sensor networks, telescopes...)



→ Data Farming:

Storage/Archiving
Indexing, Searchability
Data Fusion, Interoperability, ontologies, etc.

Database technologies

→ Data Mining (or Knowledge Discovery in Databases):

Pattern or correlation search
Clustering analysis, automated classification
Outlier / anomaly searches
Hyperdimensional visualization

Key mathematical issues

→ Data understanding

Computer aided understanding
KDD
Etc.

Ongoing research

→ New Knowledge



Data Driven Science

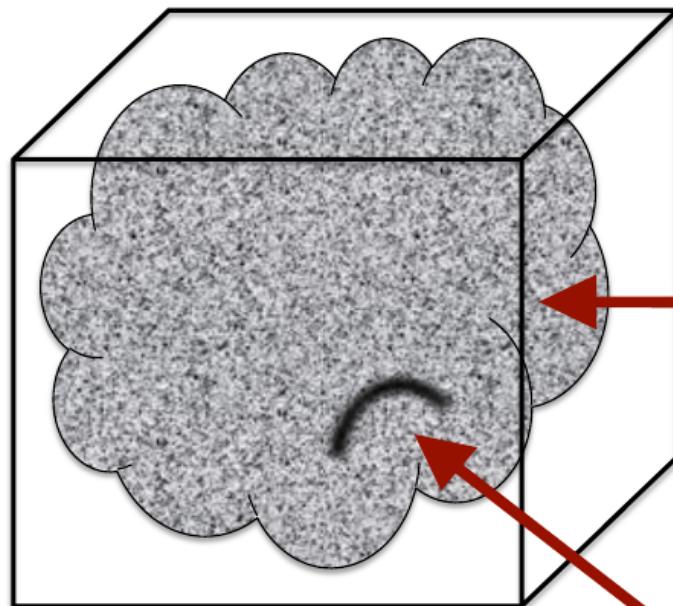
What is Fundamentally New Here?

- The *information volumes and rates* grow exponentially
 - **Most data will never be seen by humans**
- A great increase in the data *information content*
 - **Data driven vs. hypothesis driven science**
- A great increase in the *information complexity*
 - **There are patterns in the data that cannot be comprehended by humans directly**



Hidden Patterns in Data

Pattern or structure (Correlations, Clustering, Outliers, etc.) Discovery in High-Dimensional Parameter Spaces



$D \gg 3$ parameter space hypercube

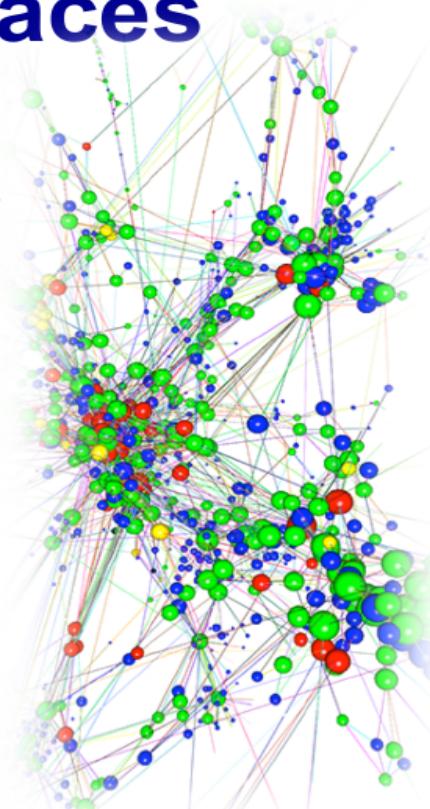
High-D data cloud:
mostly noise, of an arbitrary distribution

But in some corner of some sub-D projection of this data space, there is ***something ≠ noise***

Visualization in Machine Learning

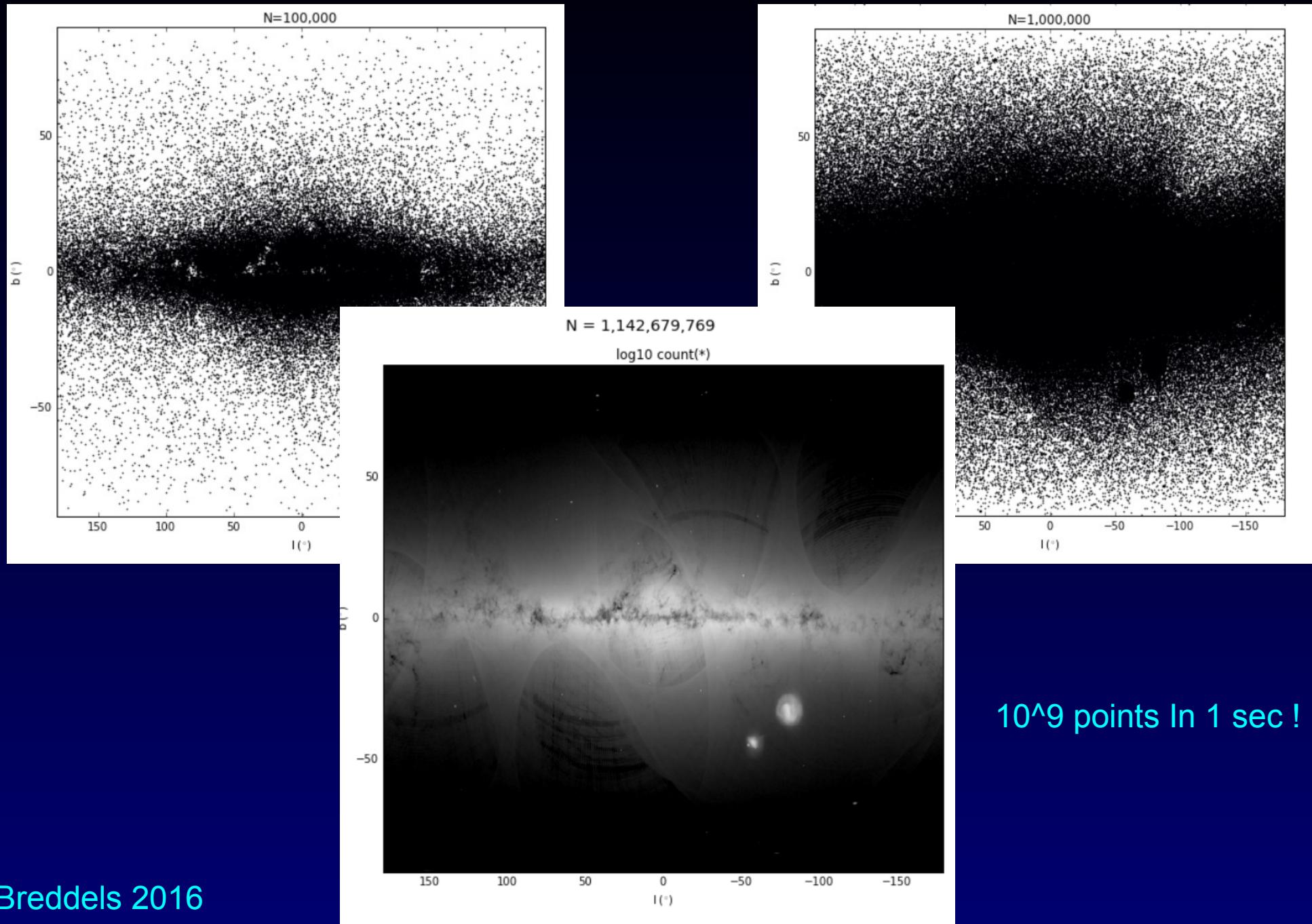
A Key Challenge: Visualising Multidimensional Data Spaces

- Hyperdimensional structures (clusters, correlations, etc.) may be present in many complex data sets, whose dimensionality may be $D \sim 10^2 - 10^4$, or higher
- It is a matter of ***data understanding***, choosing the right data mining algorithms, and interpreting the results
- We are biologically limited to perceiving up to $\sim 3 - 12(?)$ dimensions

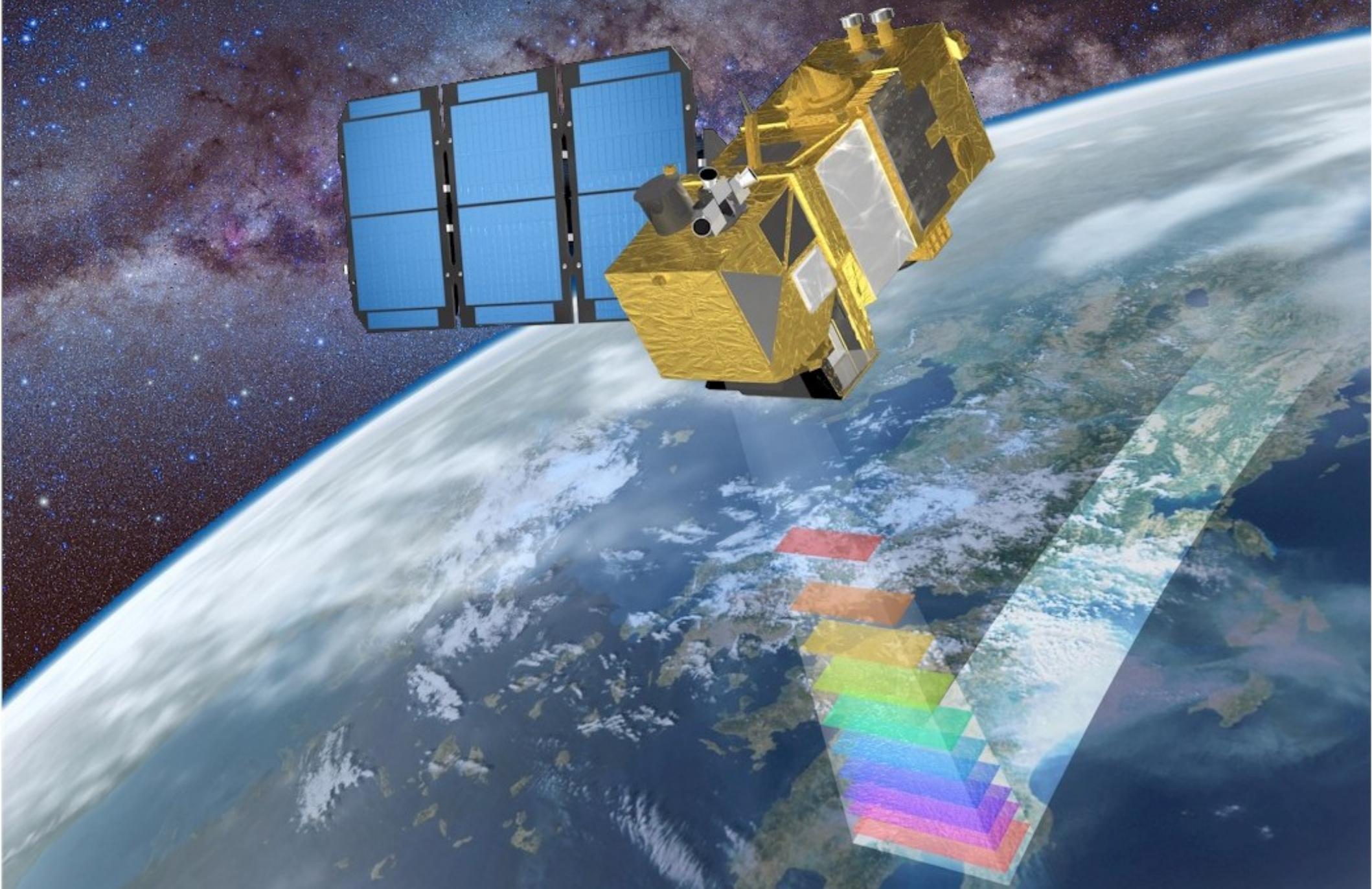


What good are the data if we cannot effectively extract knowledge from them?

Visualization of 1 B points – Gaia DR1



Big Data Era in Sky and Earth Observation – TD 1403 COST action



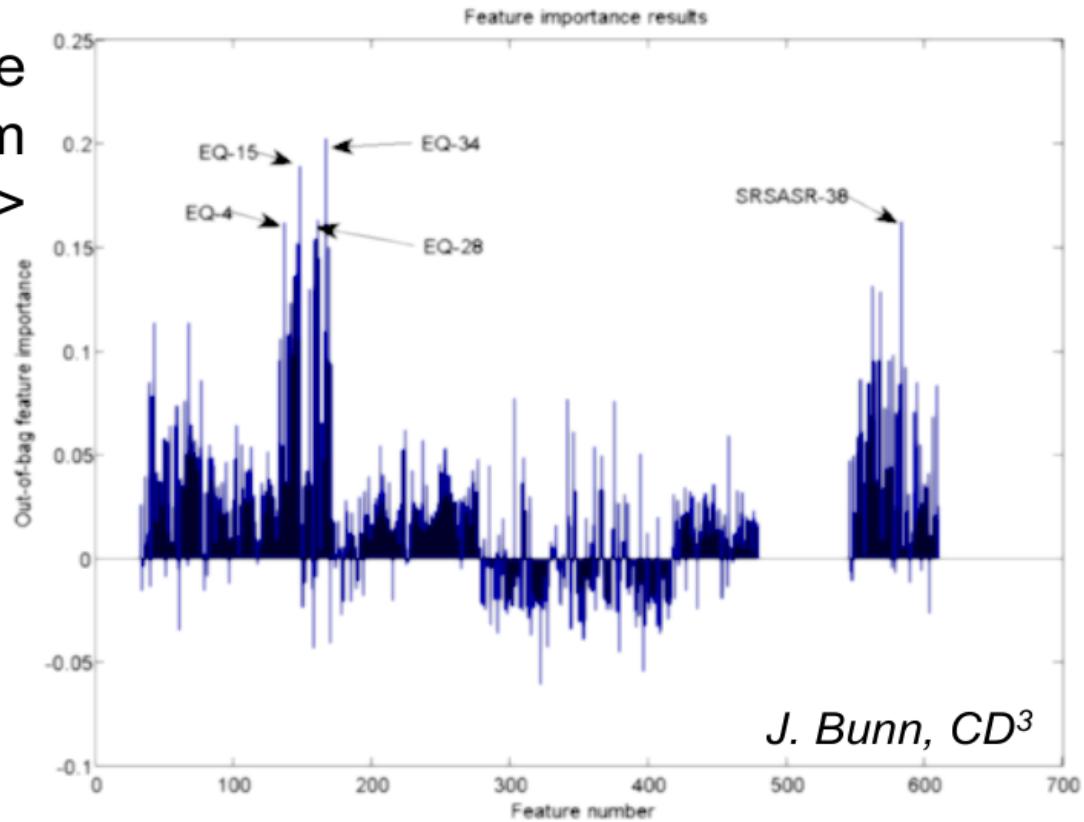
Astro-Neurology

From Sky Surveys to Neurobiology

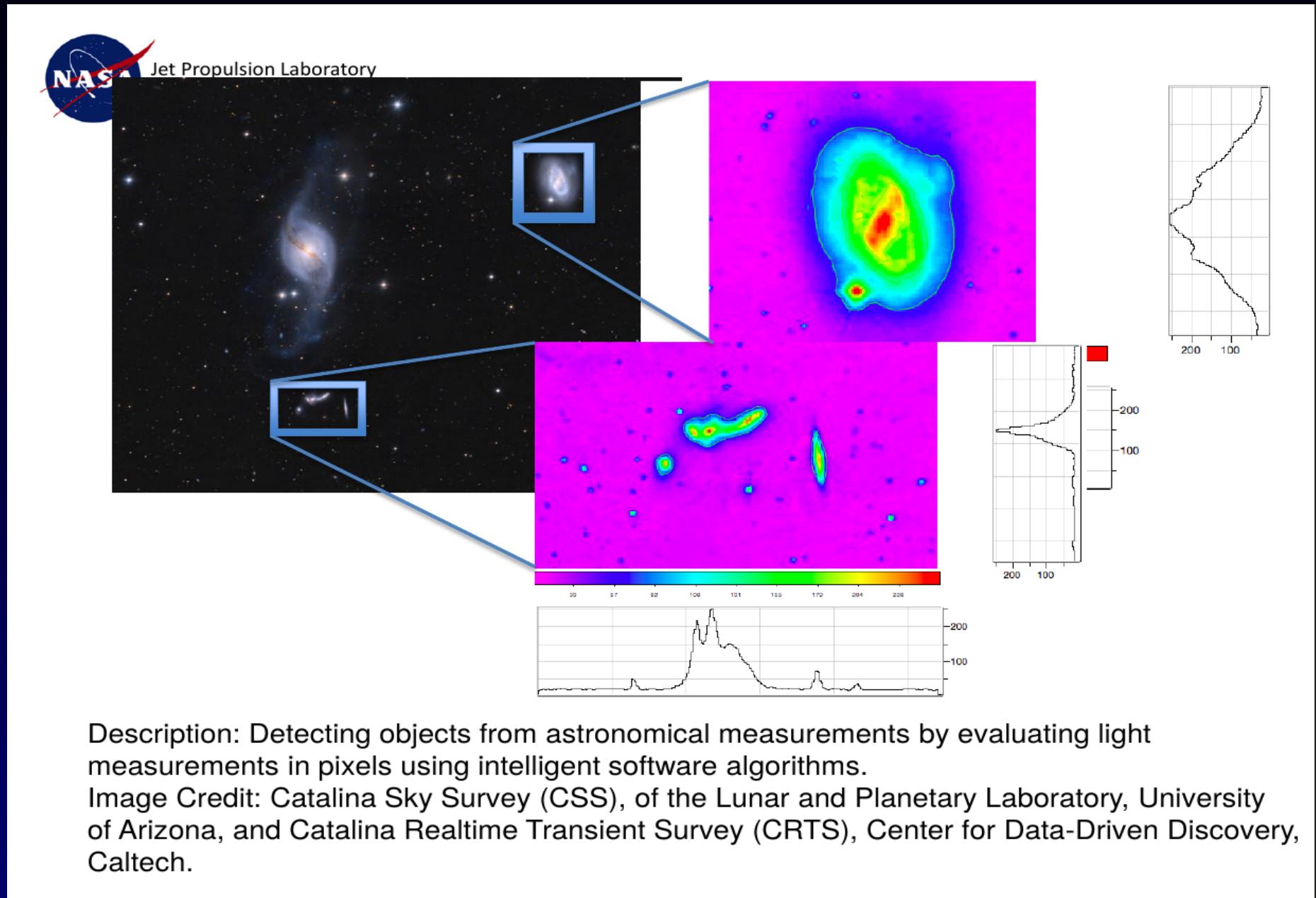
- Using the data analytics tools based on ML, developed for the analysis of sky surveys, to design a better diagnostics for autism
- Feature importance using random forests =>
 - Next: correlate with MRI scans

(with R. Adolphs et al.)

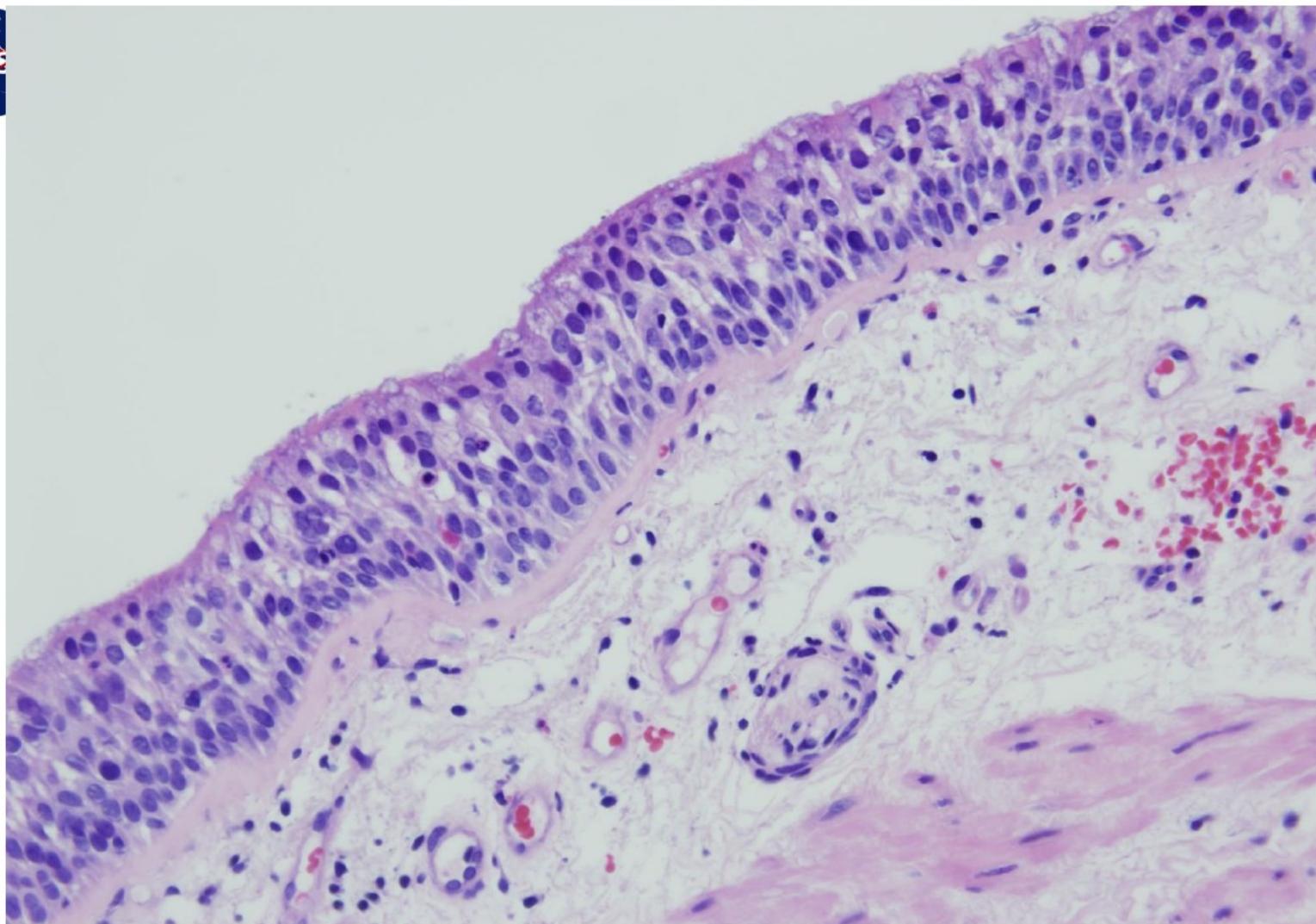
Djorgovski



Finding Galaxies by Shape NASA



Finding Cancer Signatures NASA



Description: Detecting objects from oncology images using intelligent software algorithms transferred to and from space science.

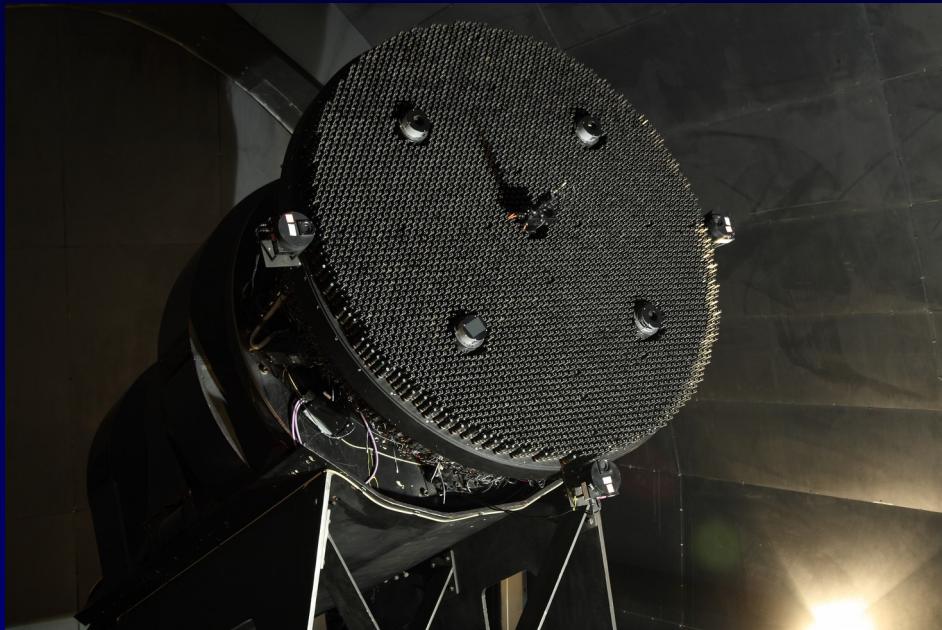
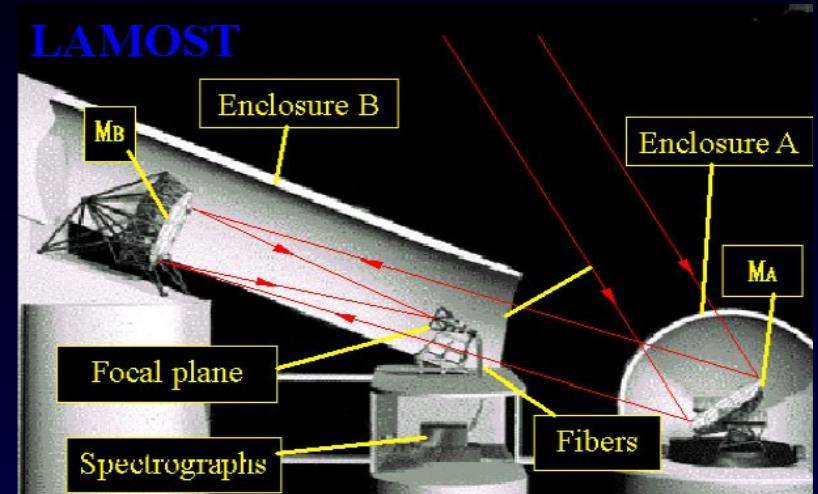
Image Credit: EDRN Lung Specimen Pathology image example, University of Colorado

BP a DP na FITs z astroinformatiky a VO

- FIT VUT Brno
 - 2011 1 BP (Random Forests in Astronomy)
 - 1 PhD – Wavelets Dimensionality Reduction (pending)
- VŠB-TU Ostrava
 - 2013+2015 1 BP + 1 DP - SPLAT-VO
- FIT ČVUT
 - 2012 2 BP (VO-Korel+SSA proxy)
 - 2013 2 BP (OSPS Image + Catalogue Server)
 - 2014 2 BP (Random Forests + SOM)
 - 2015 1 BP (VO-Cloud)
 - 2 DP (Clustering OSPS + Deep Learning)
 - 2016 2 DP (Semisupervised learning + Outlier finding)
 - 2017 1 DP (VO Cloud) + 1 BP (deep learning)+ 1 PhD (VO light curve)

LAMOST (Guoshoujing)

Xinglong- China
4m mirror (30 deg meridian)
4000 fibers
10 mil spectra / 5 yr
Automatic RV-z



LAMOST Spectral Surveys

DR1 (end 2013) **2 204 860** spectra
 1 085 404 stars

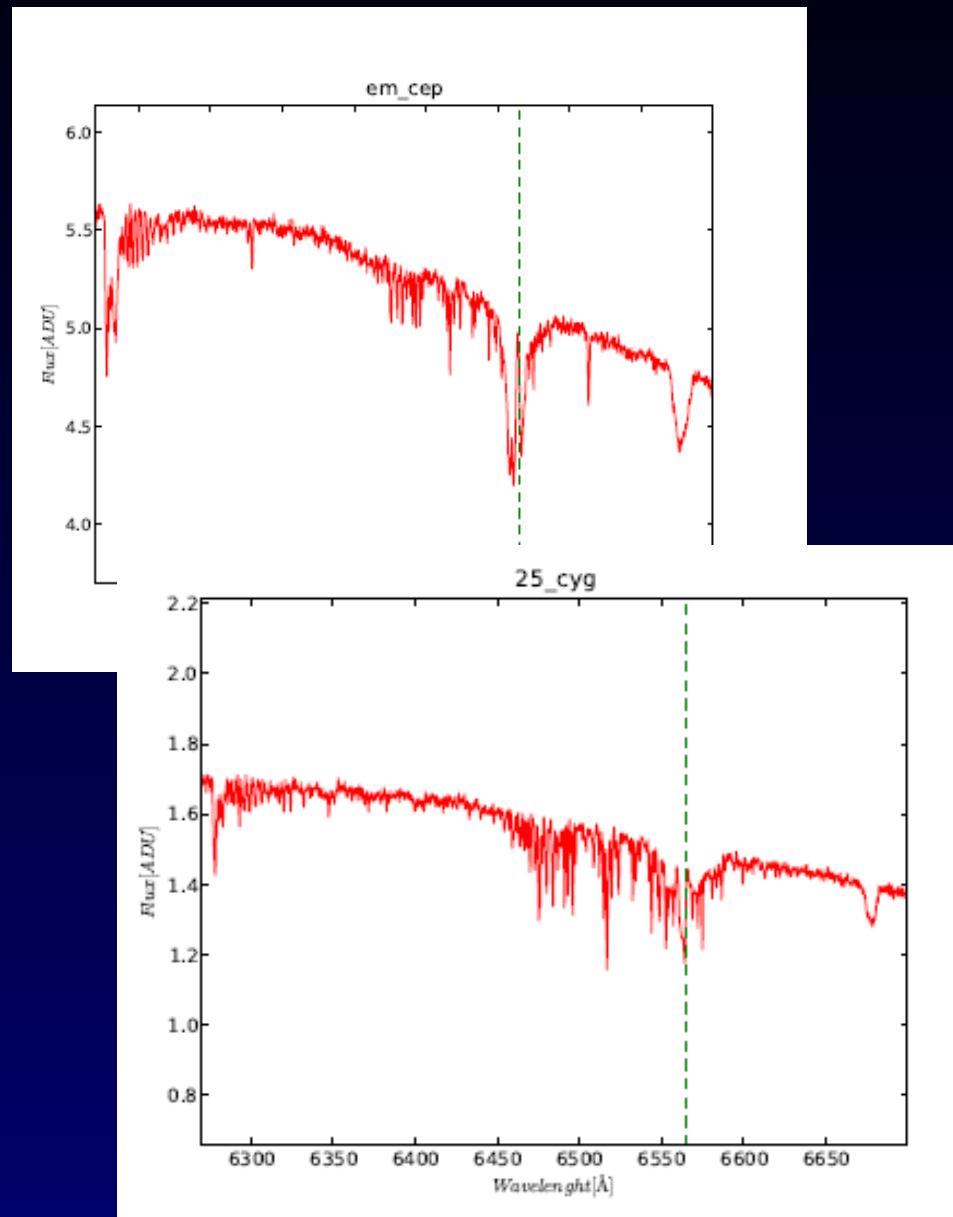
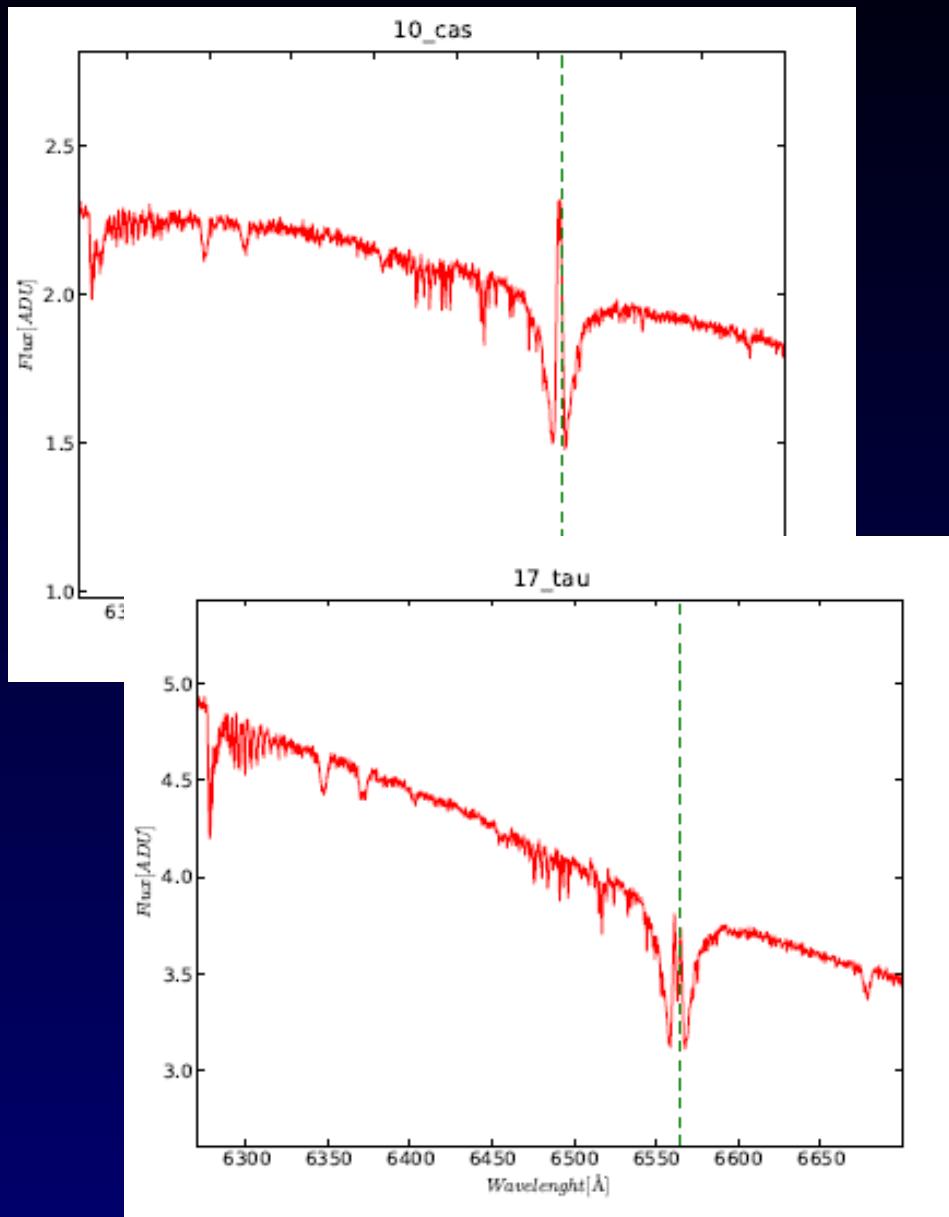
DR3 (half 2015) **5 755 126** spectra
DR4 (Feb 2016) **+ 741 522**

Each Fiber – 2 motors
double arm 33mm circle

Fibre collects light from
3.3 arcsec circle on sky

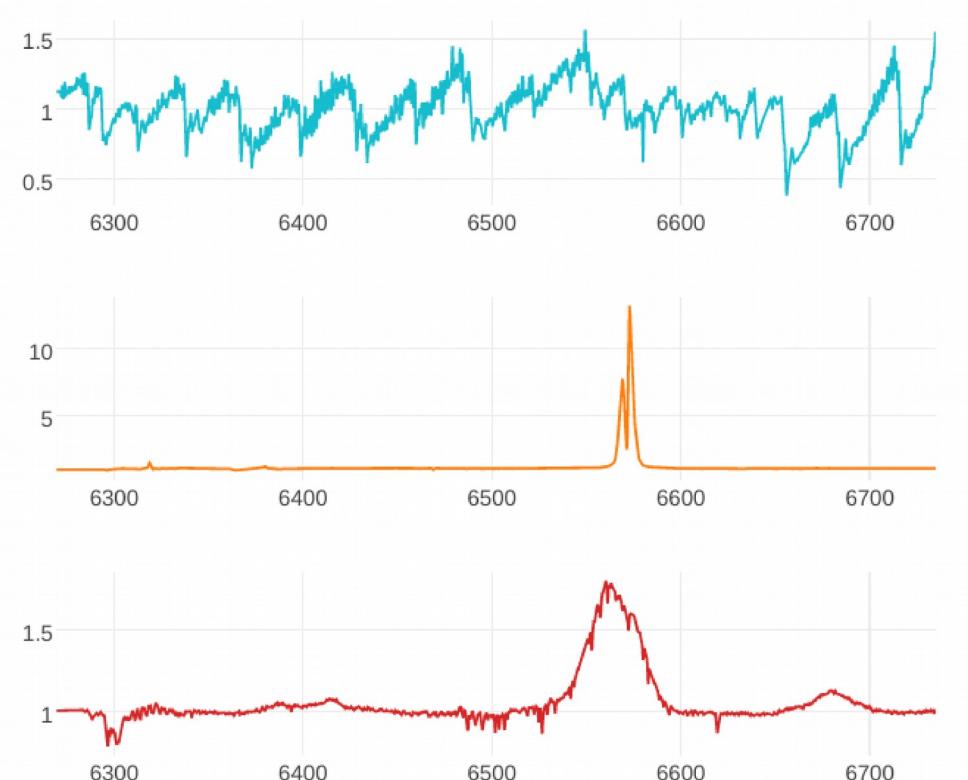
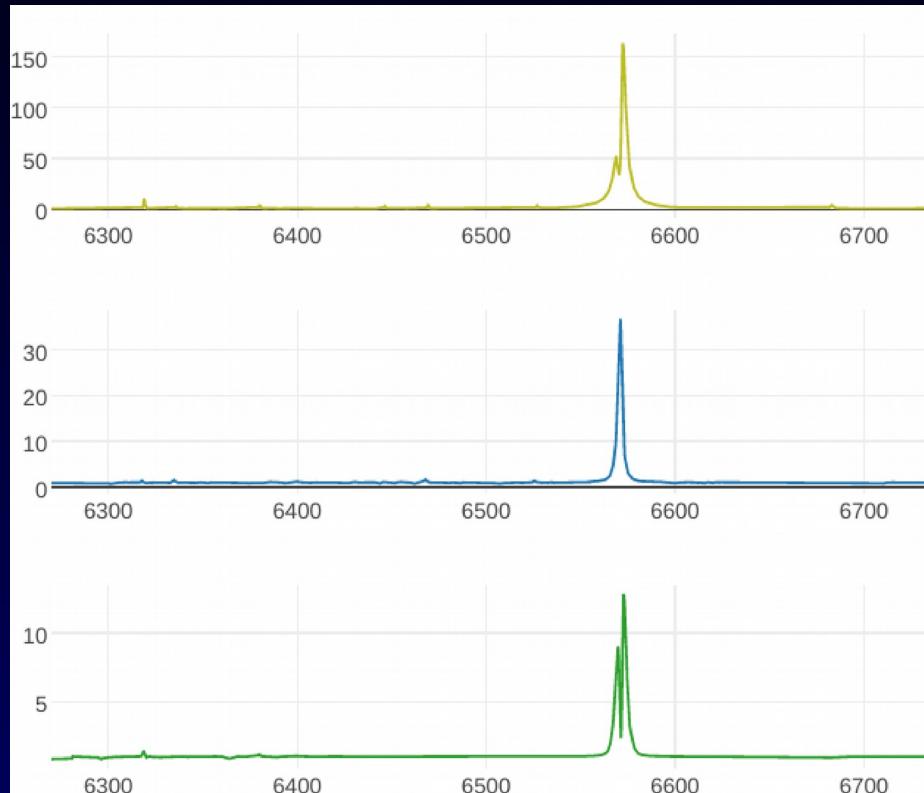


Be Stars : Emission in absorption



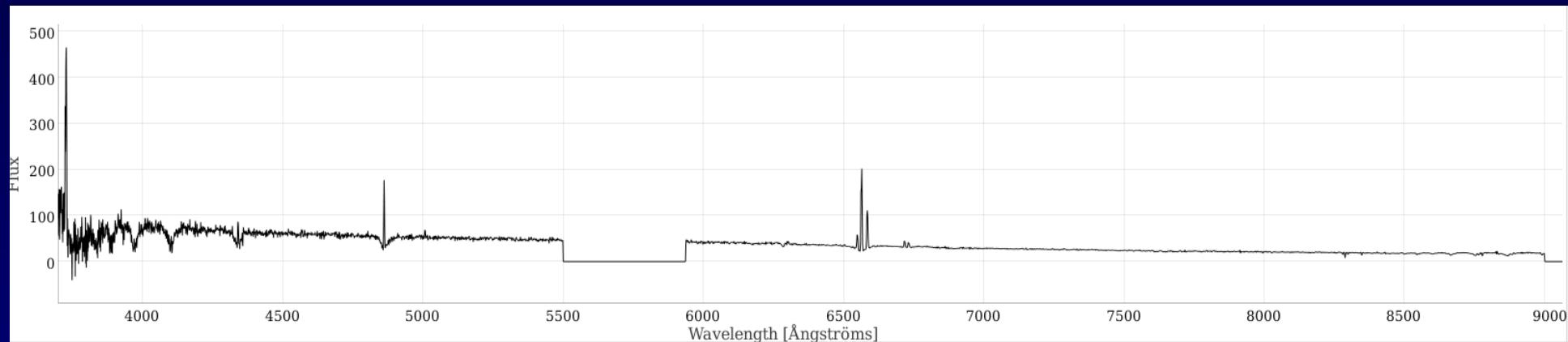
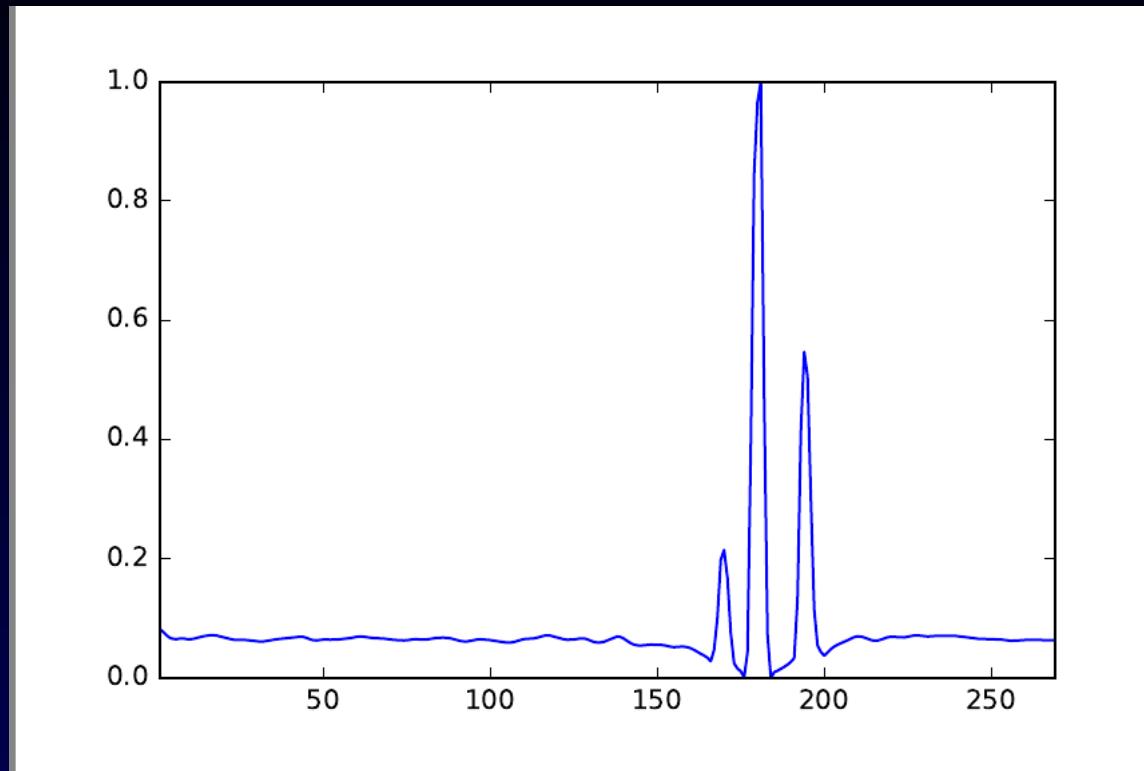
CCD700 Outliers

Unsupervised learning – Local Outlier Factor - LOF

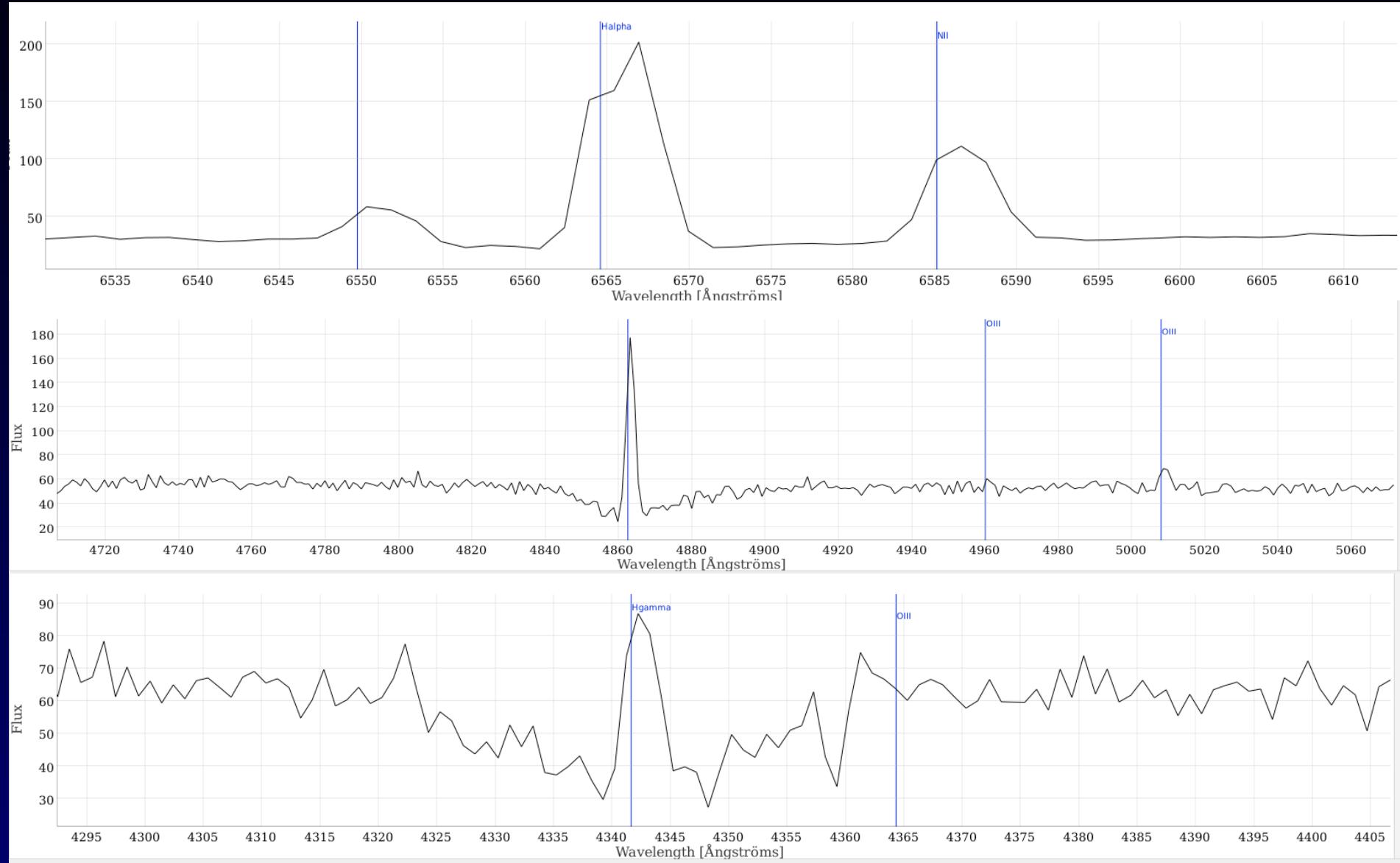


Shakurova 2016

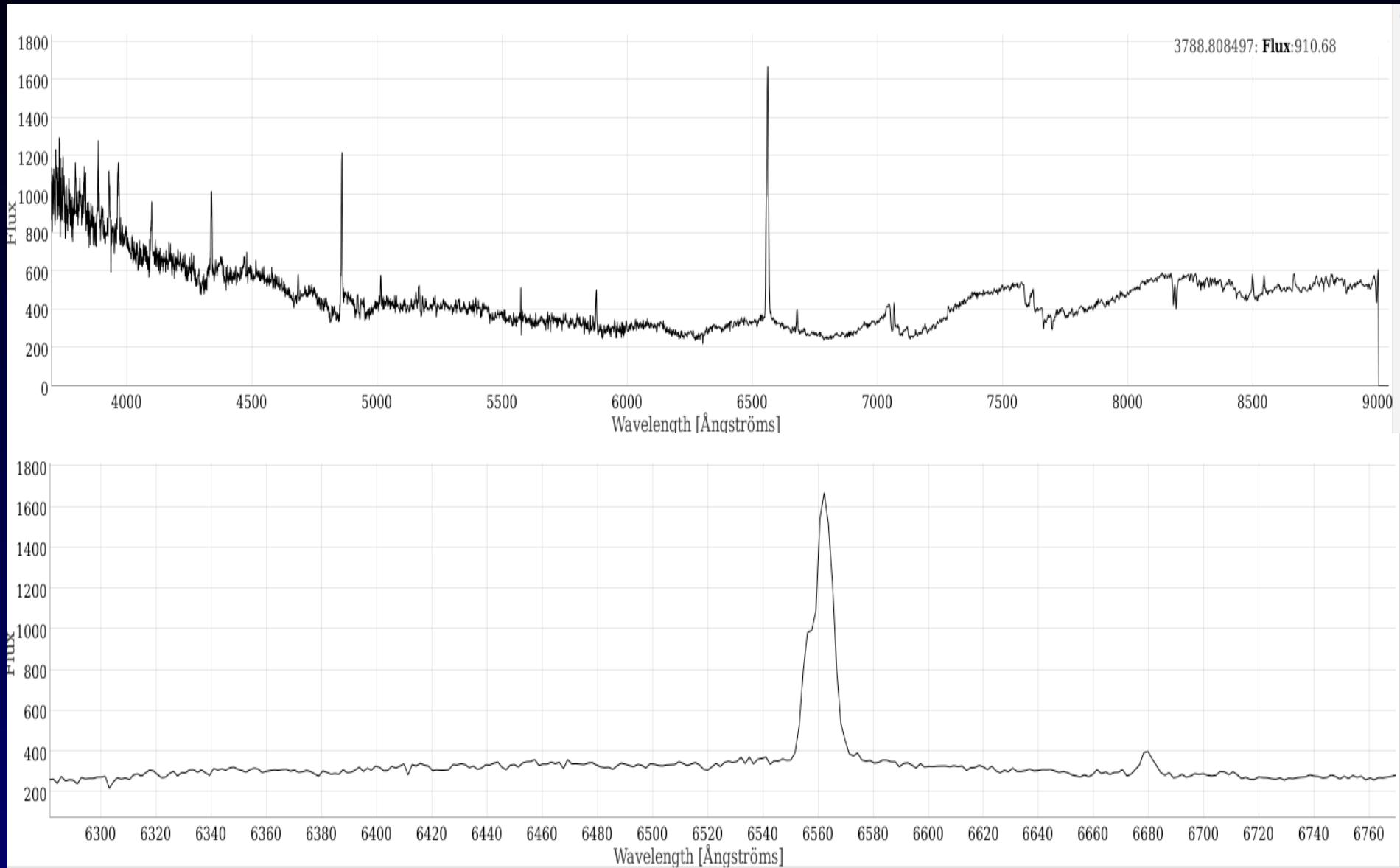
Be Candidates Found



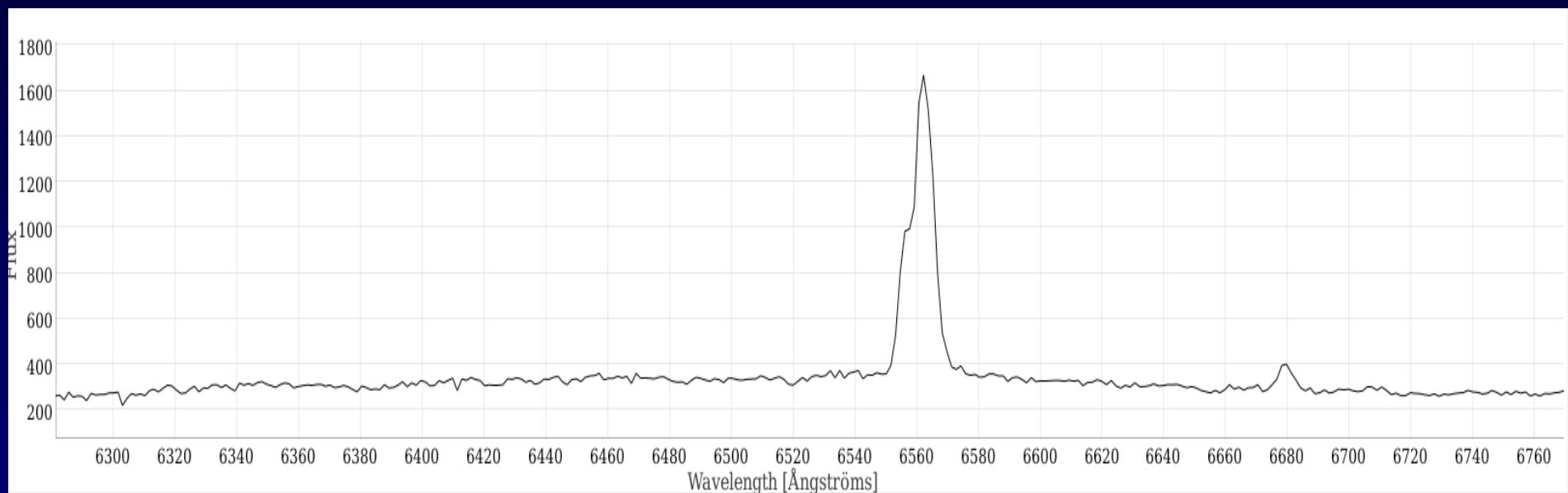
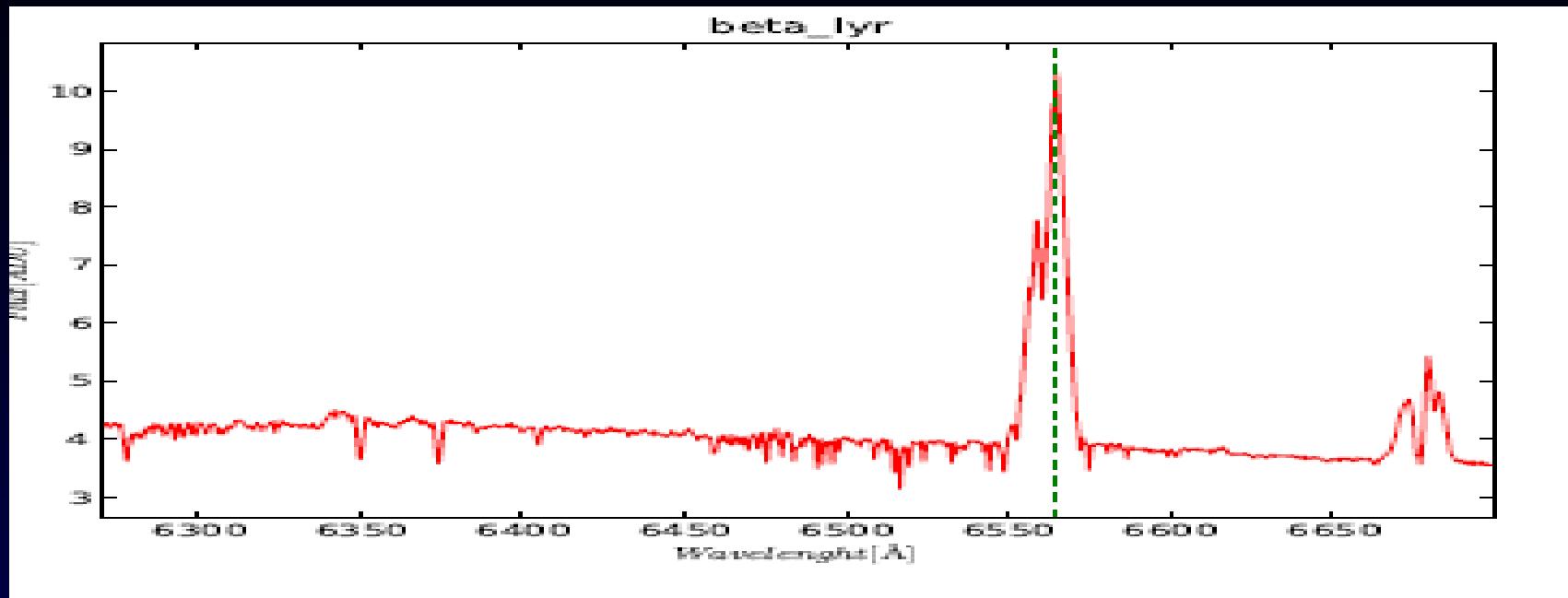
Be Candidates Found



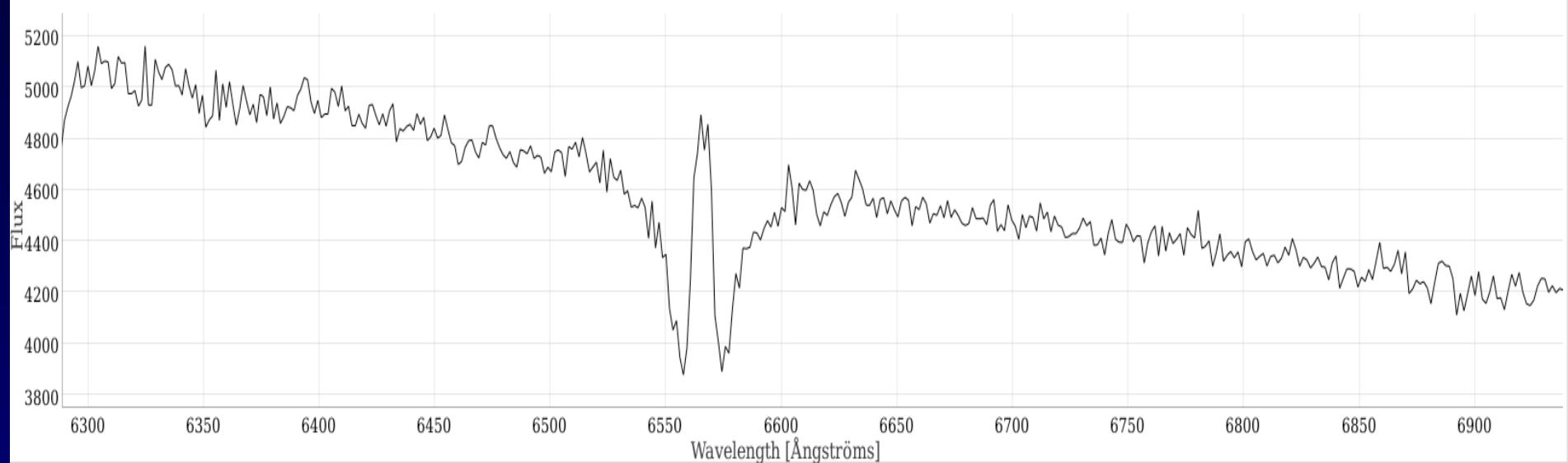
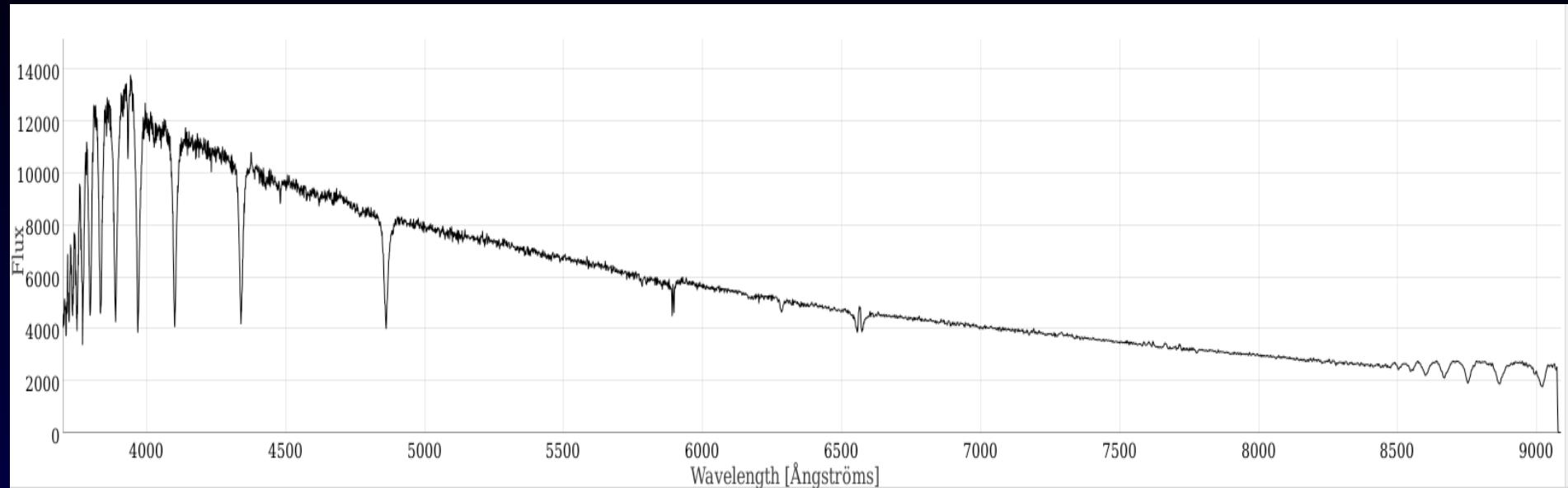
Be Candidates Found



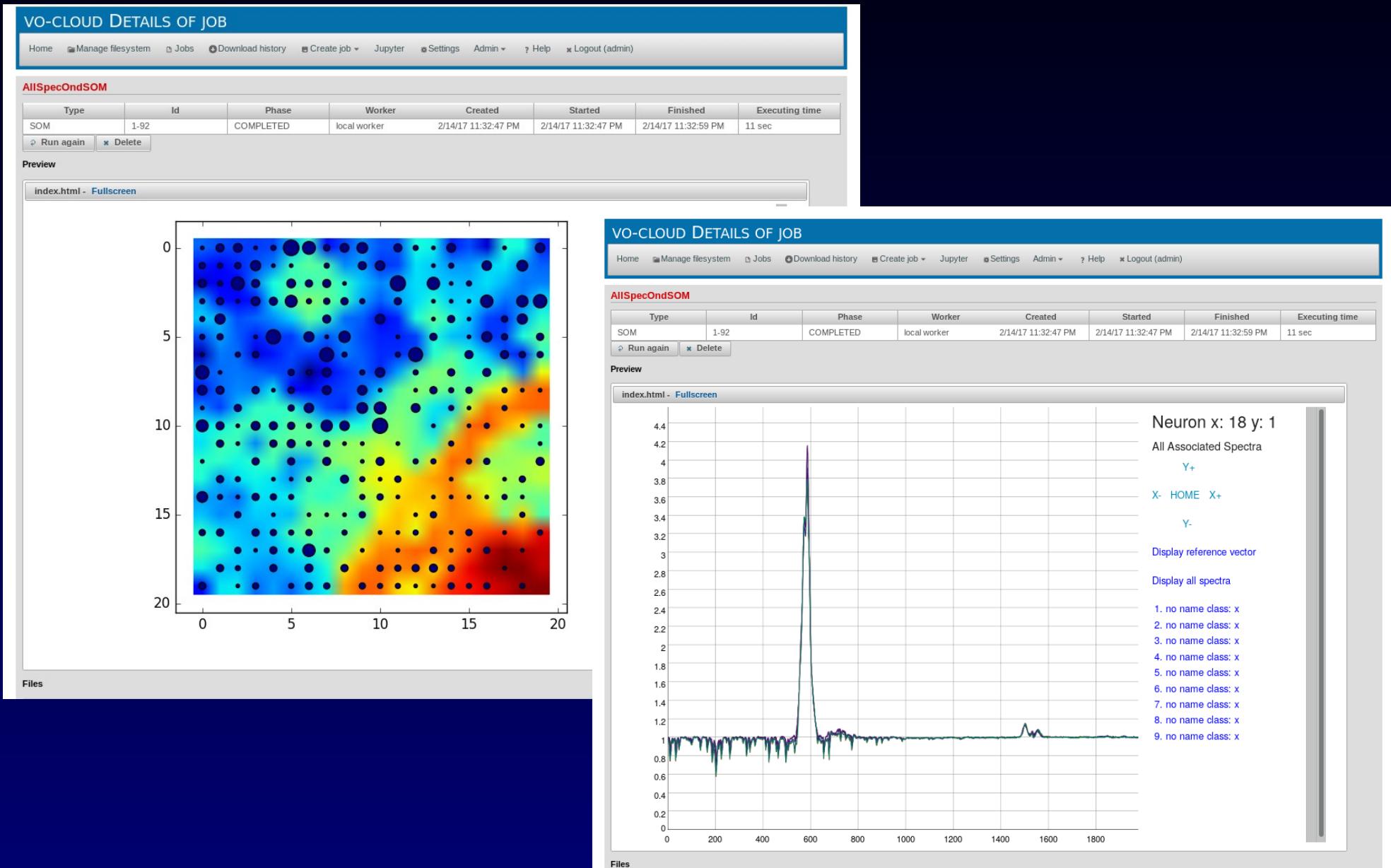
Be Candidates Found



Yet Unknown Be Star



Samoorganizační mapa (Kohonen)



VO-CLOUD vizualizace spekter

Budoucnost 2m

Určitě velmi moderní

V souladu s dobovými trendy

Moderní technologie

Mnoho produkce z archivu – followup, hledání zpět