Global Scientific Data Sharing and Big Data Knowledge discovery

* * * * * *

Czech Virtual Observatory and COST action Big Sky Earth

Petr Škoda

Astronomical Institute of the Czech Academy of Sciences

Supported by grant COST LD-15113 of the
Czech Ministry of Education Youth and Sports
And COST Action TD1403 Big Sky Earth

RDA Meets Czech Researchers
Prague, Czech Republic, 27th October 2017

Credits

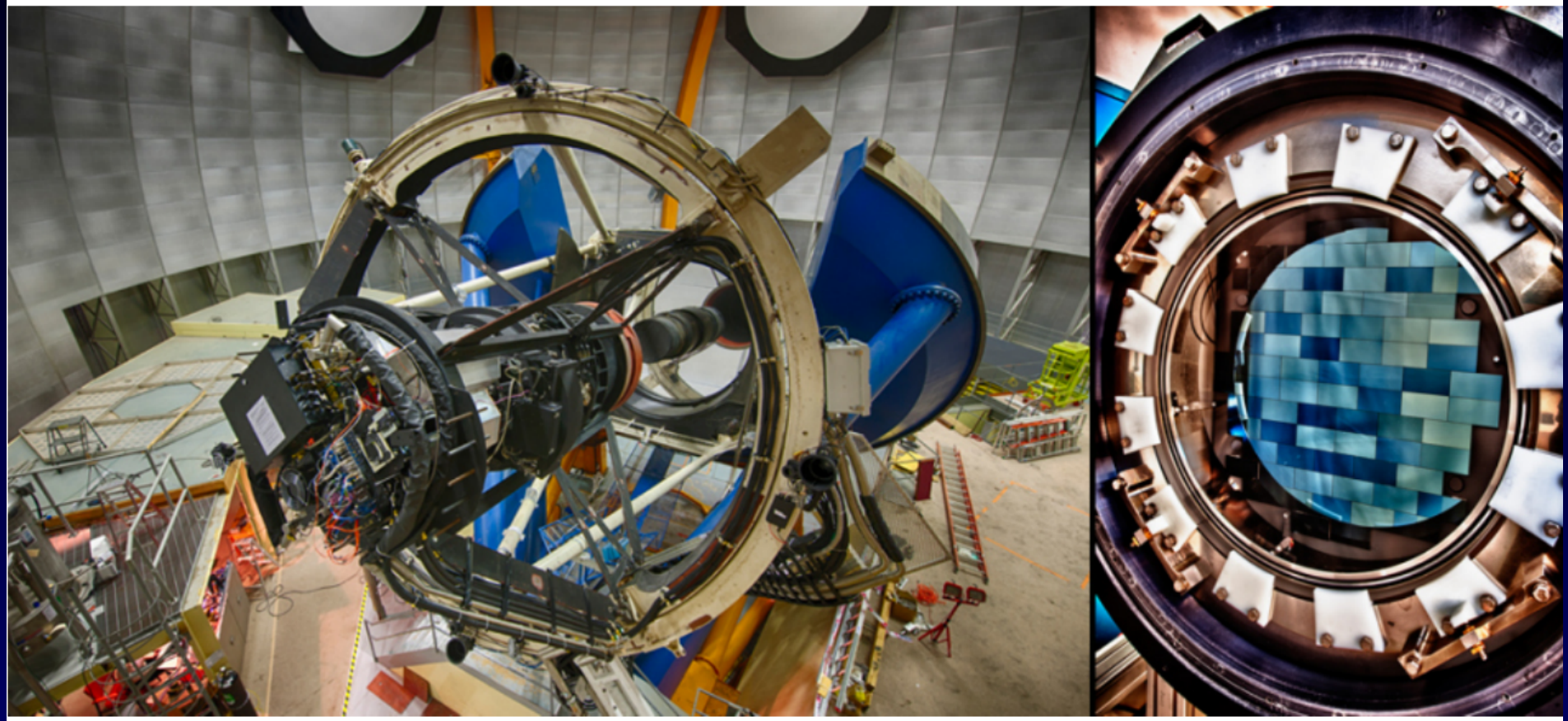
- The presentation is based on many different sources – mainly the on-line published slides from IVOA meetings, slides from Astroinformatics and COST meetings or pictures found on Internet.
- We acknowledge namely materials of B.Hanish, G. Djorgovski, G. Longo, T. Hey and M. Breddels, D Vinkovic, P. Baumann, A Nina and presentations from AI2016 in Sorrento

Outline of the Talk

- Data Avalanche in astronomy
- Virtual Observatory
- Astroinformatics
- Big Sky Earth
- Transfer of technology
- CZVO
  - Be stars discovery in LAMOST surveys
  - Ondrejov Southern Photometry Survey
- Future

Dark Energy Survey Camera

Dark Energy Camera (DECam)



~0.4 PB/yr

74 chips – 570 Mpixels - 4m Cerro Tololo

Large Synoptic Survey Telescope

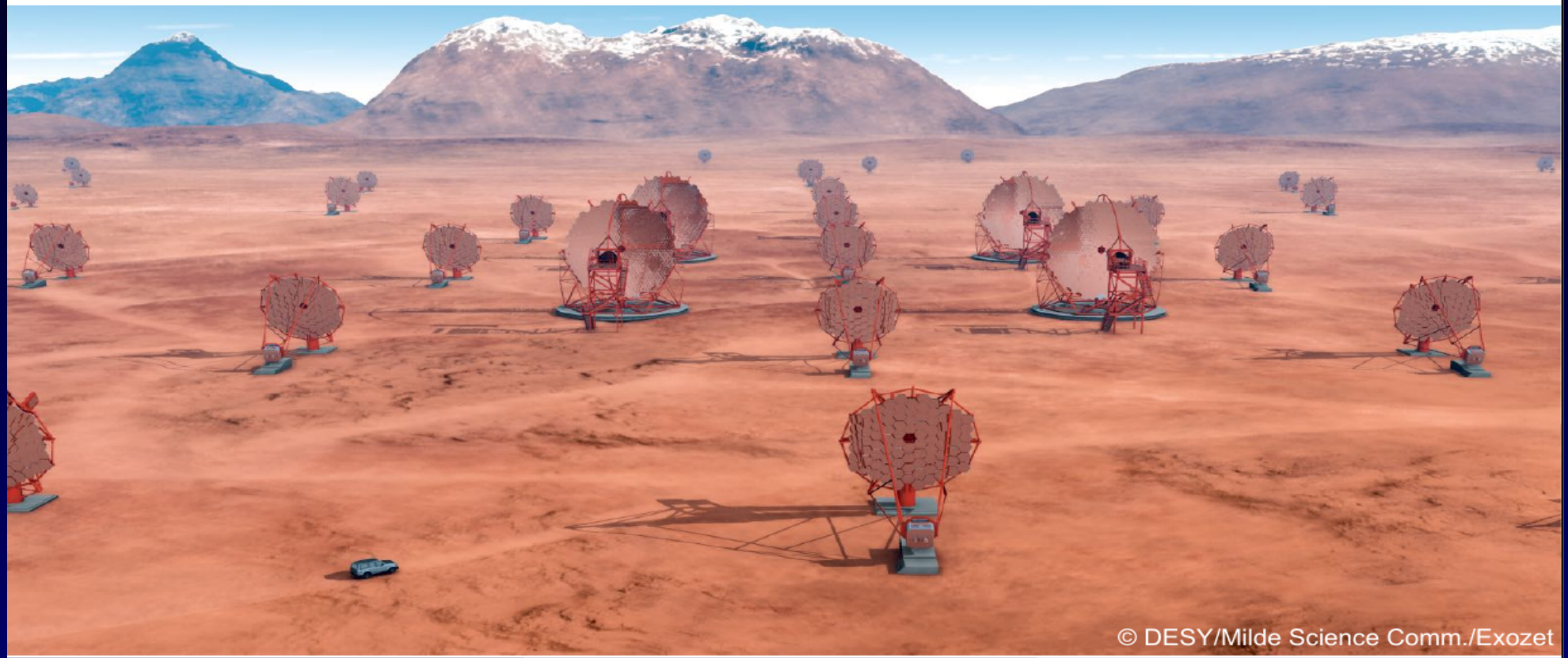
201 CCD 4kx4k,
3.2 Gpix every 20 sec
3.5 deg FOV (64cm)
20 TB/day=6 PB/yr RAW
1.5 PB catalogue !!!
detection of changes 60s!

38 billion objects x 1000
32 tril. meas. -5 PB table

Cherenkov Telescope Array

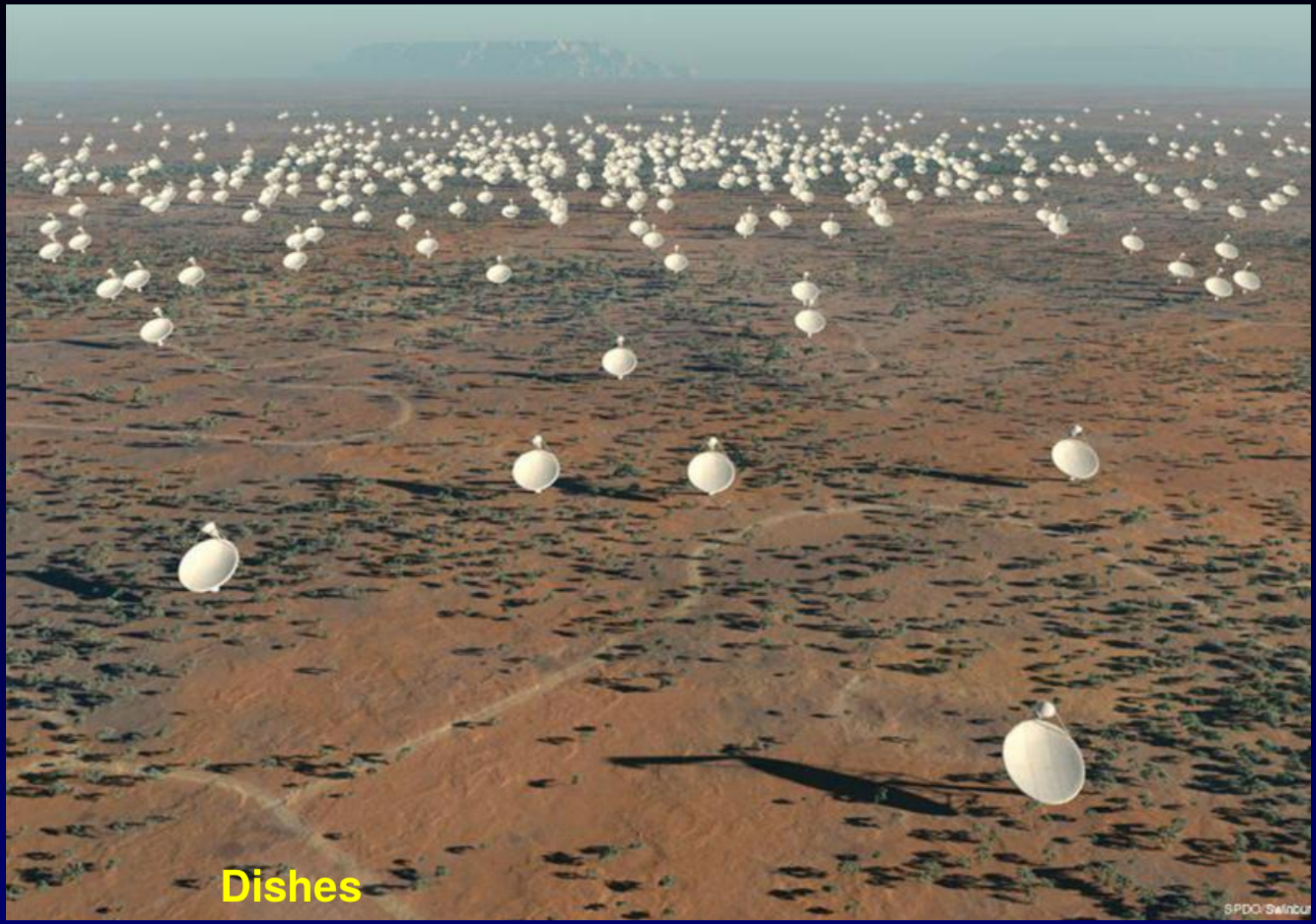
cta
cherenkov telescope array
Observatory

- 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



27 PB/year in archive

SKA



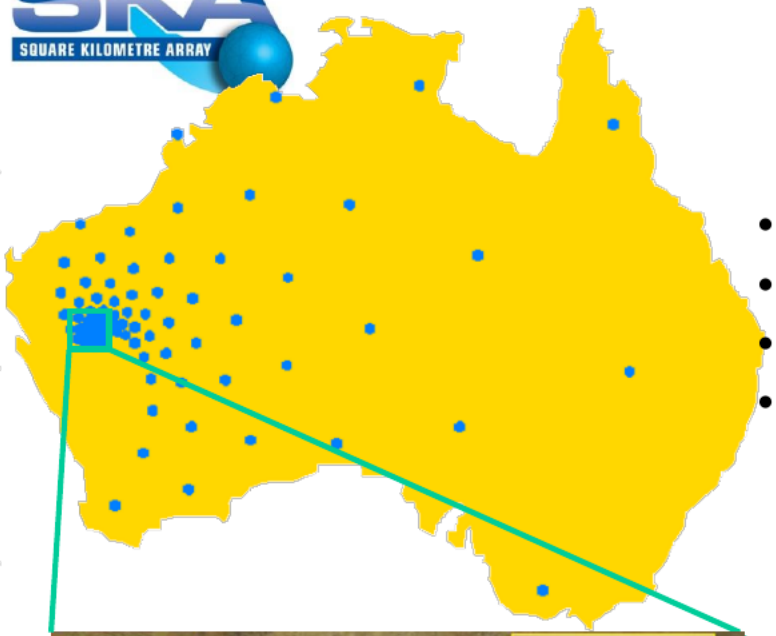
SKA

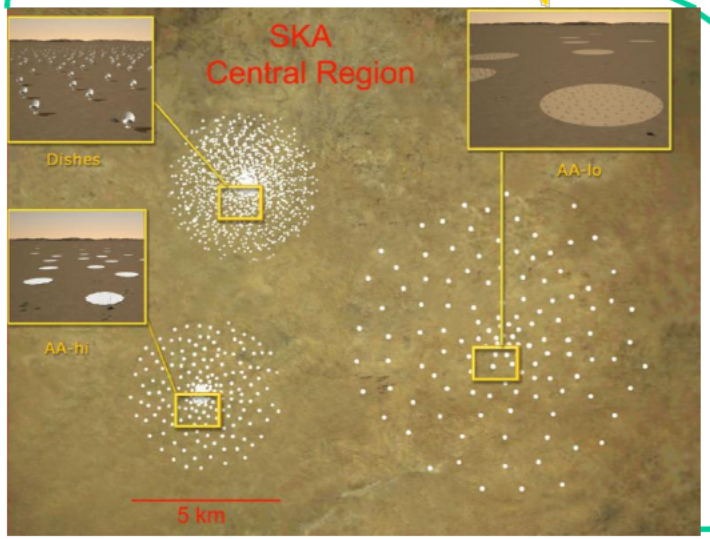
SKA SQUARE KILOMETRE ARRAY

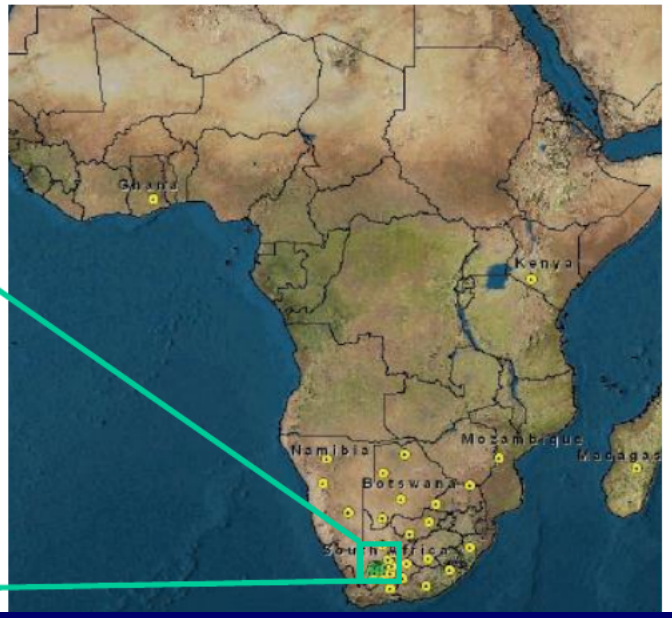
UNIVERSITY OF CAMBRIDGE

also a Continental sized Radio Telescope

- Need a radio-quiet site
- Very low population density
- Large amount of space
- Possible sites (decision 2012)
  - Western Australia
  - Karoo Desert RSA







LOFAR network

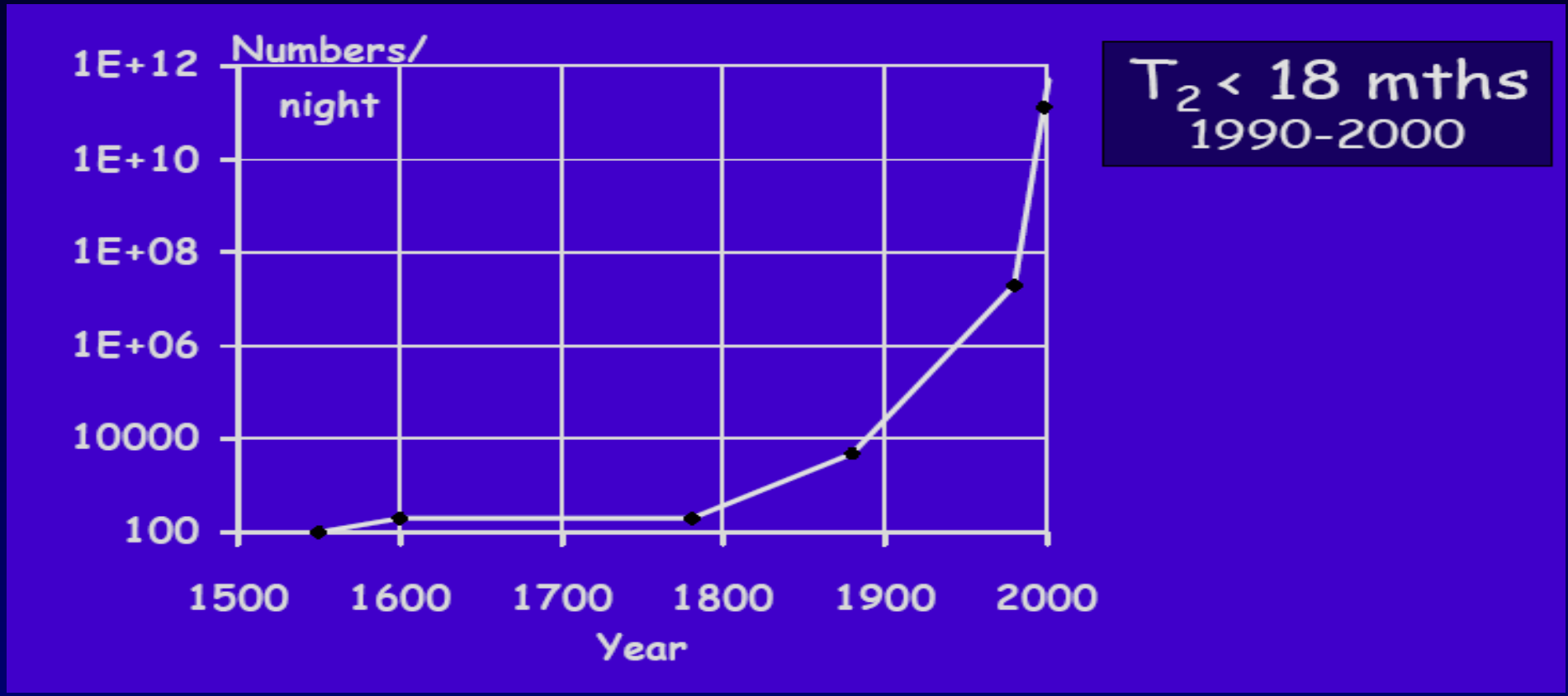
	LOFAR	SKA
Raw Telescope	112 PB/yr	60 EB/yr
Archive Rate	6 PB/yr	100 PB/yr

Data Avalanche

Moore law for chips –doubling 1.5 year

Data in astronomy – doubling < 1 yr ! (1000/10 yr)

====> Processing to data



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.

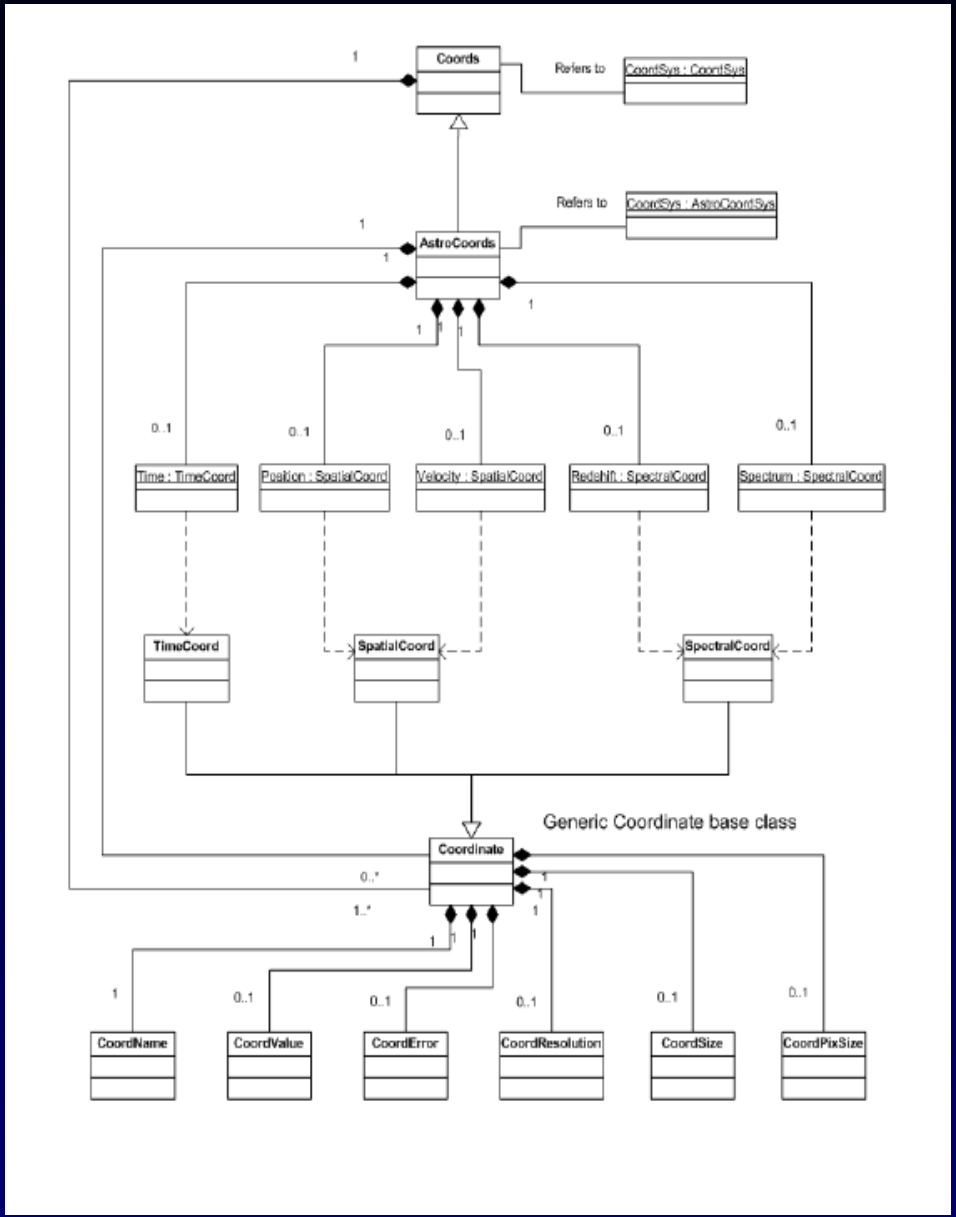
R.J.Hanisch, P.J.Quinn, in "IVOA – Guidelines for participation"

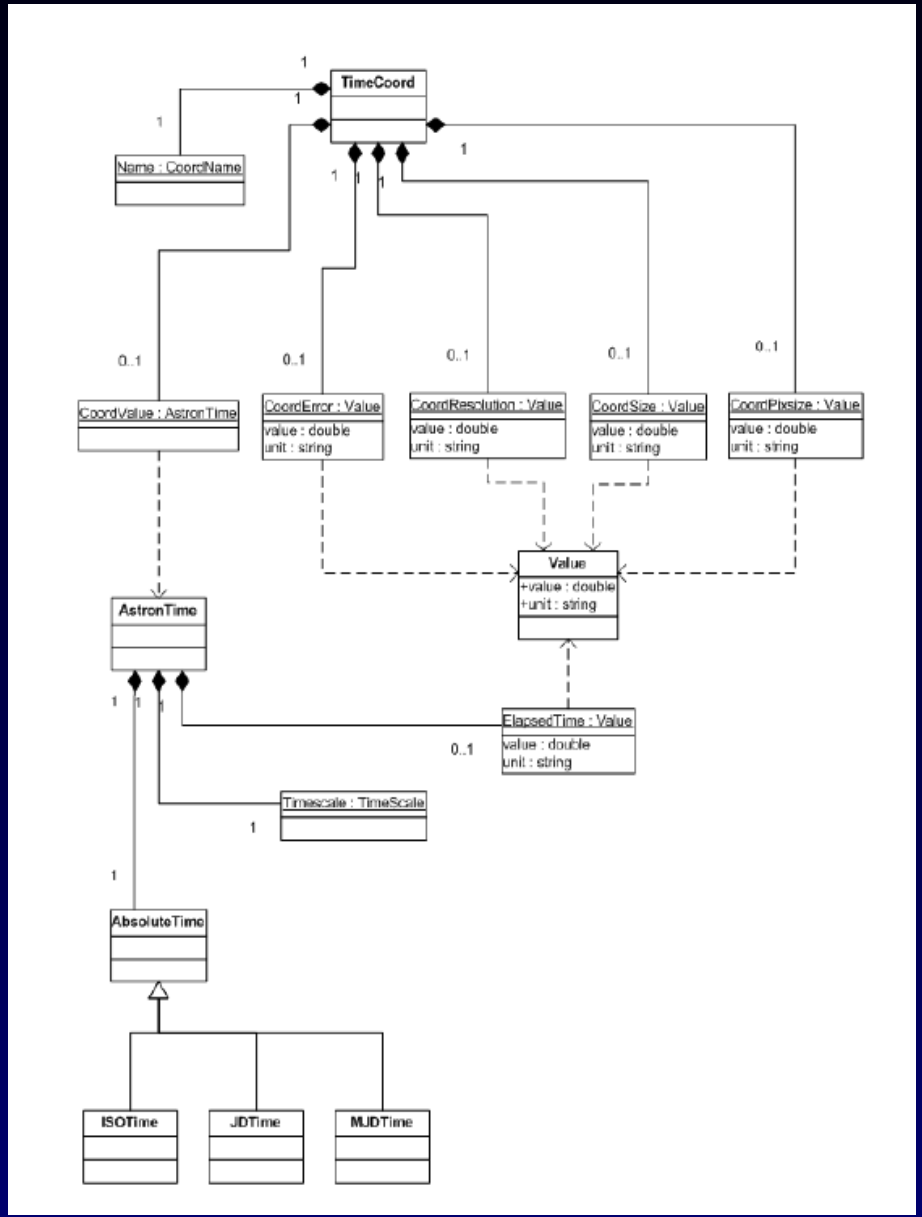
Start of VO – early 2000

IVOA

Astro Grid
ArVO
American Virtual Observatory
Aus-VO
BRAVO
China-VO
CVO
esa
France
ESO
IVOA
GERMAN ASTROPHYSICAL
GAVO
VIRTUAL OBSERVATORY
EURO VO
HVO
Hungarian Virtual Observatory
JVO
NOVA
RVO
SVO
vo-i
VAO
VIRTUAL ASTRONOMICAL OBSERVATORY
VObs.it

Space-Time-Coordinate Data Model





Technology of VO

Unified data format– VOTable, UCD (semantics)

Web services (WS)

VOregistry – Google for data+WS

protocols

- ConeSearch (searching in circle on sky)
- SIAP (Simple Image Access Protocol)
- SSAP(Simple Spectral Access Protocol)
- SLAP(Simple Line Access Protocol) - VAMDC
- TAP (Table Access Protocol) – query e.g. whole SDSS
- VOEVENT (transients, robotic telescopes,Sun)
- datacubes, DATALINK on-the-fly data processing

Technology of VO

ADQL (Astronomical Data Query Language)

- XMATCH, REGION (2 catalogues – shifted)

Application interoperabilty – SAMP

- Allows develop applications as bricks
- sending VOTABLES (catalogue-spectra-images)

Big Data handling

VO Space	Moving big tables across (load only results)
SSO	Authentication, authorization, groups and consortia
UWS	Universal worker service (job synch, asynch)
PDL	Parameter Description Language
SIM-DB	Simulations, theory data

Simple Spectra Access Protocol
Spectral Data Model

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:

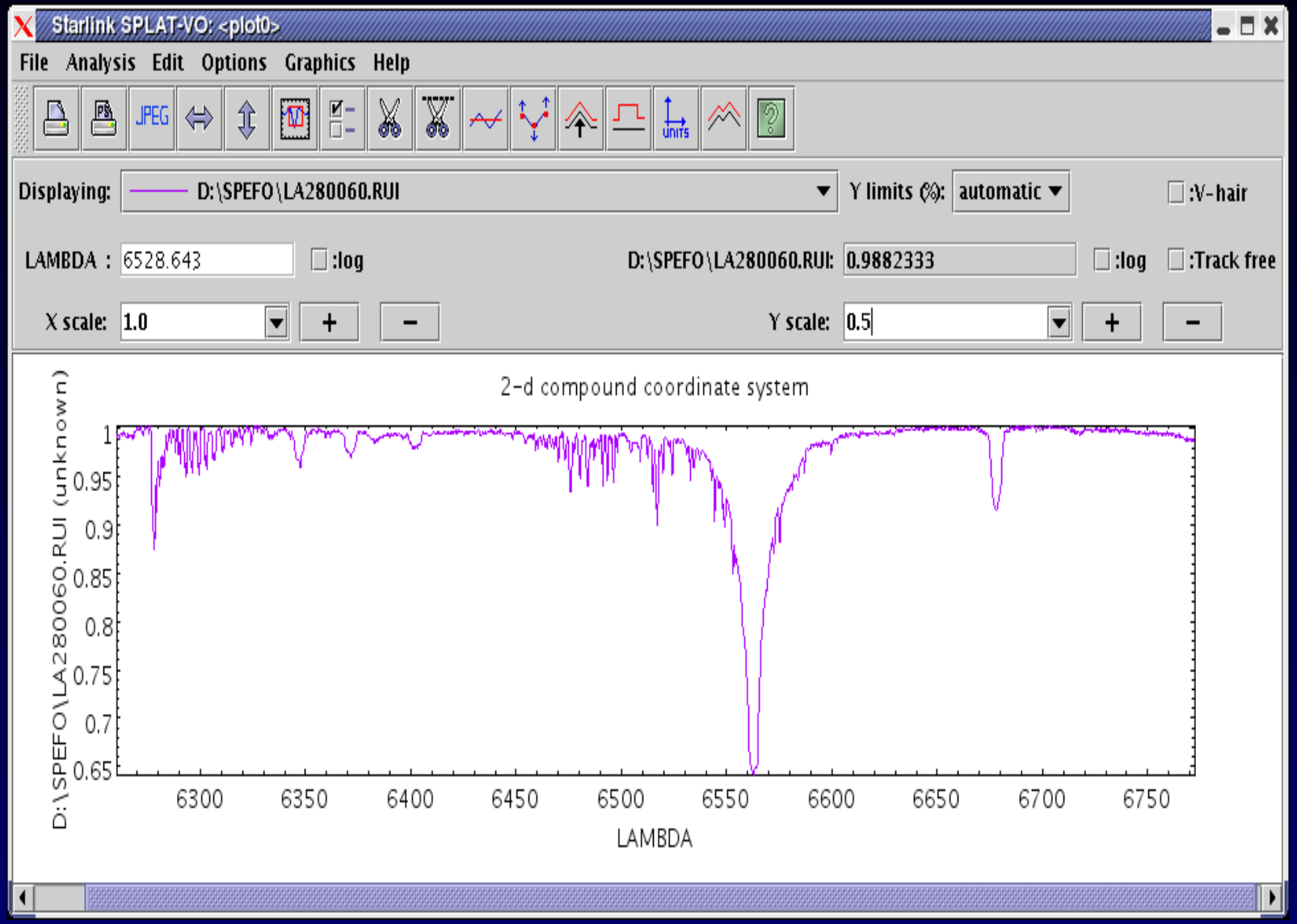
Parameter	*Sample value*	*Physical unit*	*Datatype*
POS	52,-27.8	degrees; defaults to ICRS	string
SIZE	0.05	degrees	double
BAND	2.7E-7/0.13	meters	string
TIME	1998-05-21/1999	ISO 8601 UTC	string
FORMAT	votable	-	string

4.1.2 Recommended and Optional Query Parameters

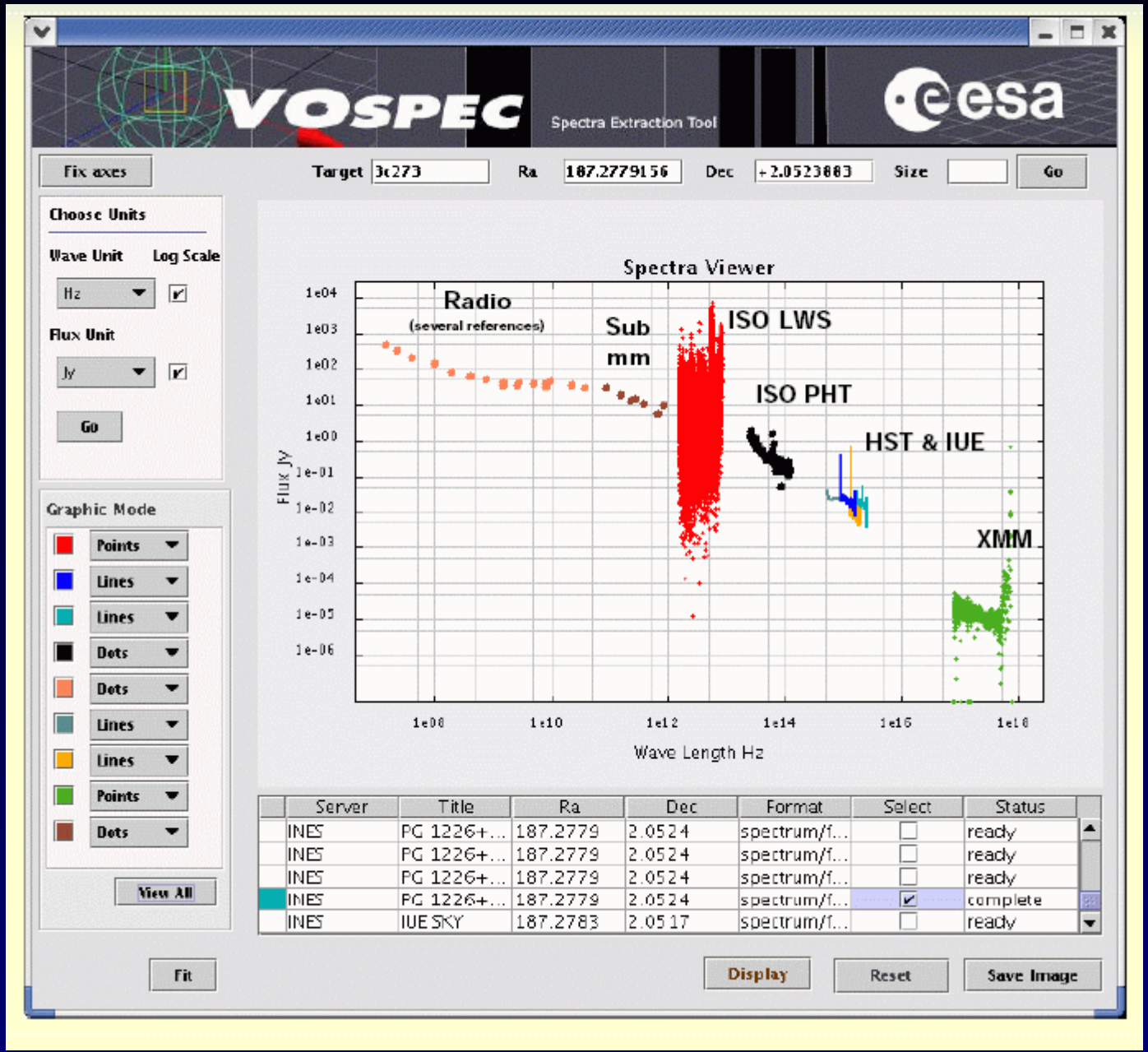
Parameter	Sample value	Unit	Req	Datatype
APERTURE	0.00028 (=1")	degrees	OPT	double
SPECRP	2000	λ/dλ	REC	double
SPATRES	0.05	degrees	REC	double
TIMERES	31536000 (=1yr)	seconds	OPT	double
SNR	5.0	dimensionless	OPT	double
REDSHIFT	1.3/3.0	dimensionless	OPT	string
VARAMPL	0.77	dimensionless	OPT	string
TARGETNAME	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

SPLAT-VO
(Starlink, Heidelberg, Ondrejov)



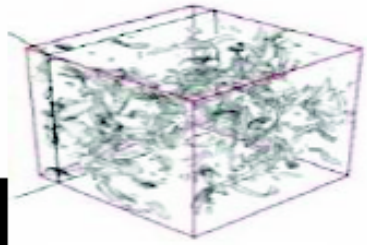
VOspec (ESAC)

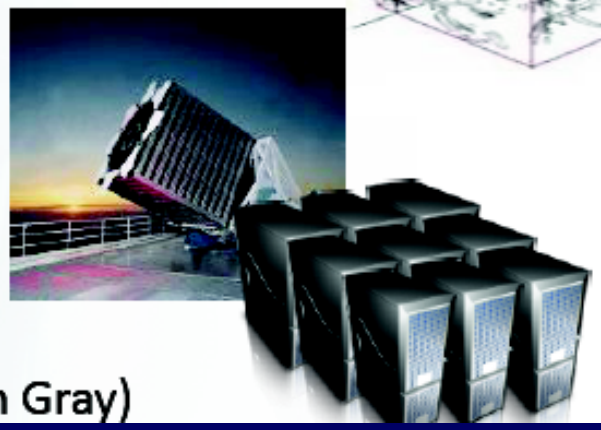

VOSPEC
Spectra Extraction Tool
esa
Fix axes
Target 3c273
Ra 187.2779156
Dec +2.0523883
Size
Go
Choose Units
Wave Unit
Log Scale
Hz
Flux Unit
Jy
Go
Graphic Mode
Points
Lines
Lines
Dots
Dots
Lines
Lines
Points
Dots
View All
Spectra Viewer
Radio
(several references)
Sub
mm
ISO LWS
ISO PHT
HST & IUE
XMM
Flux Jy
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

Emergence of a Fourth Research Paradigm

1. Thousand years ago – **Experimental Science**
   - Description of natural phenomena
2. Last few hundred years – **Theoretical Science**
   - Newton's Laws, Maxwell's Equations…
3. Last few decades – **Computational Science**
   - Simulation of complex phenomena
4. Today – **Data-Intensive Science**
   - Scientists overwhelmed with data sets from many different sources
     - Data captured by instruments
     - Data generated by simulations
     - Data generated by sensor networks
   - ➢ eScience is the set of tools and technologies to support data federation and collaboration
     - For analysis and data mining
     - For data visualization and exploration
     - For scholarly communication and dissemination

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G\rho}{3} - \mathrm{K}\frac{c^2}{a^2}$$



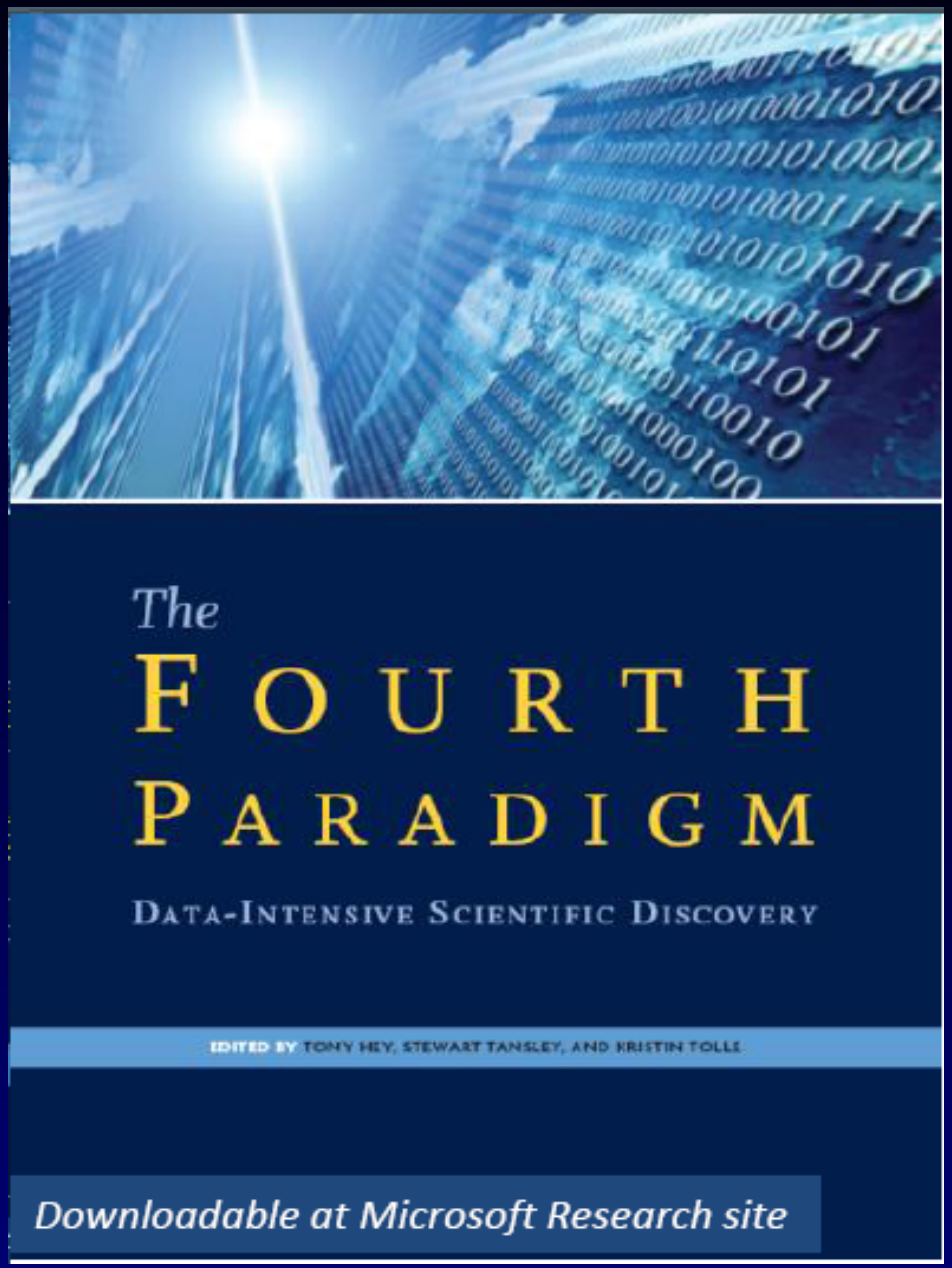


(With thanks to Jim Gray)

From T. Hey, AI2010

T. S. Kuhn – Structure of Scientific Revolutions

X-informatics

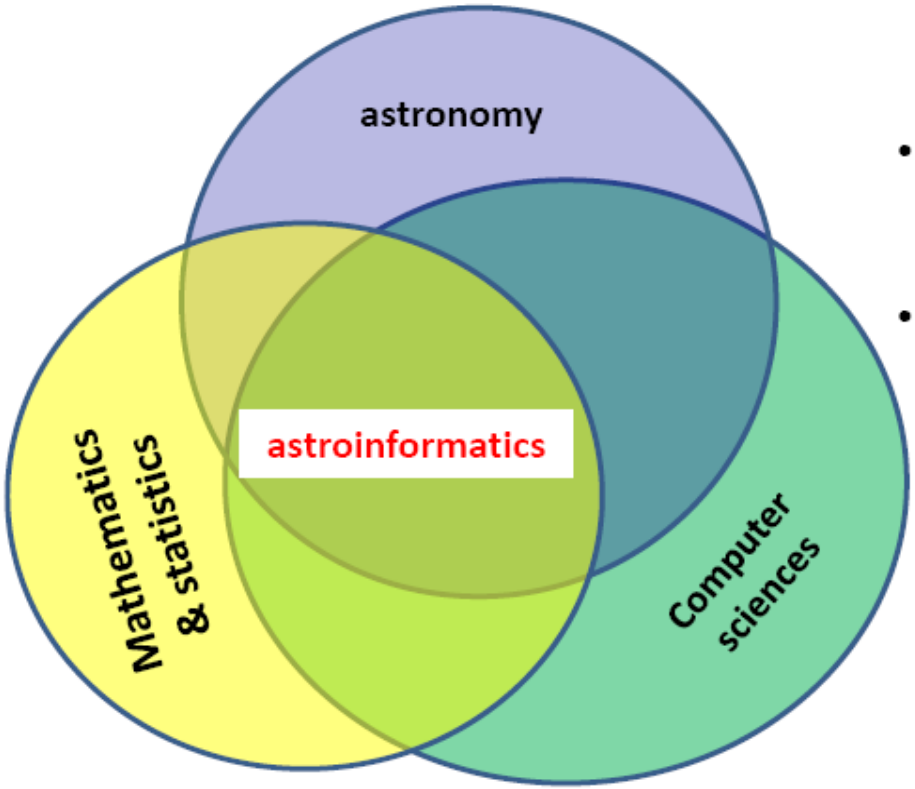


Changing methodology of the Science

Synergy between different worlds

Sociological aspects (net-based research communities)

Experimental astronomy has become a three players game



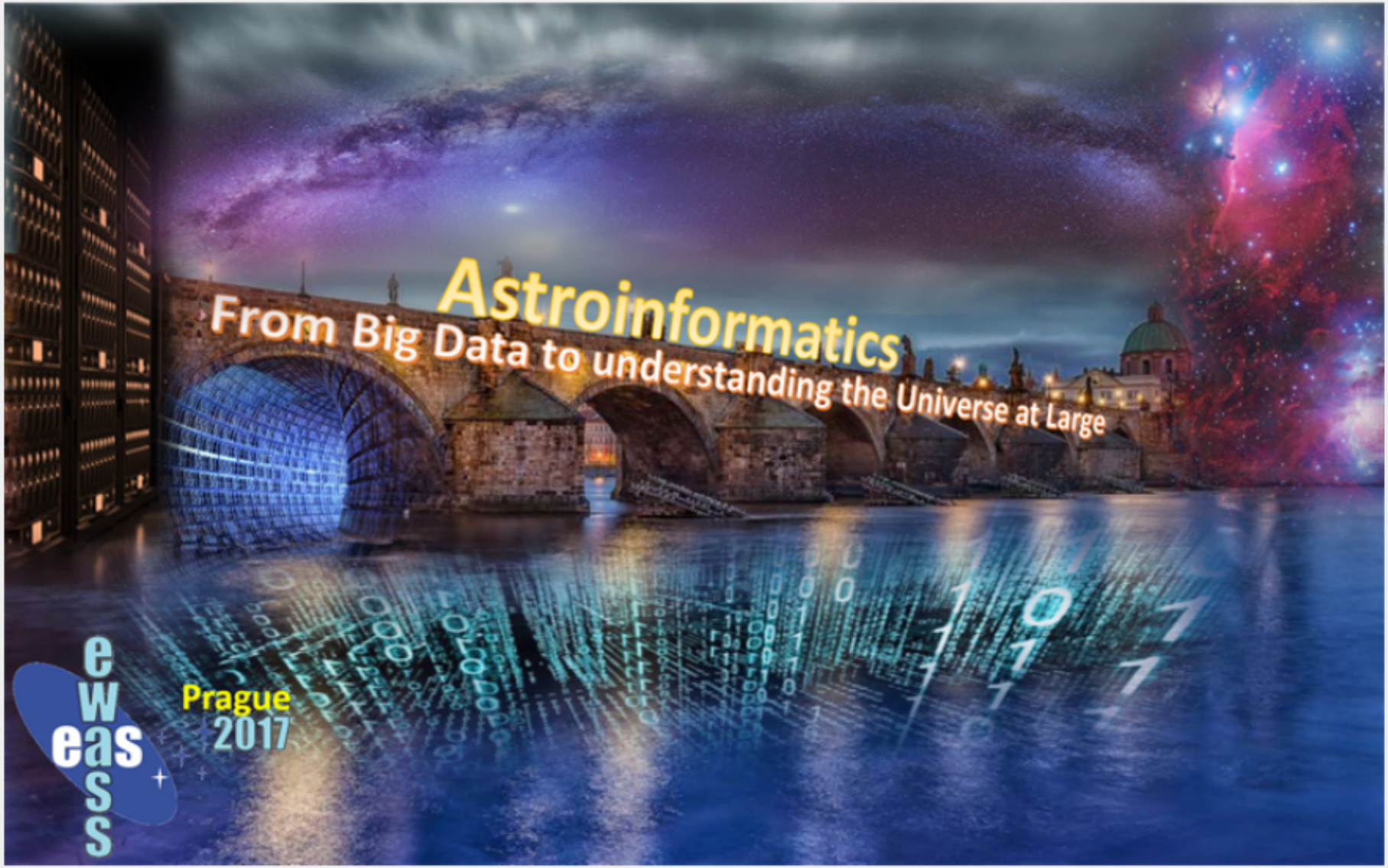
- **astronomy**: problems, data, understanding of the data structure and biases
- **mathematics**: evaluation of the data, falsification/validation of theories/models, etc
- **computer science**: implementation of infrastructures, databases, middleware, scalable tools, etc

- **Astroinformatics: AAS n. 215, Washington, December 2009,** chairperson: K. Borne
- **Astroinformatics 2010: Caltech (USA) June 16-19 2010;**co-chairpersons: S.G. Djorgovski, G. Longo
- **Astroinformatics 2011: UNINA – Sorrento,** co-chairpersons: S.G. Djorgovski, G. Longo

Longo 2010

Symposium S14

29 – 30 June 2017



http://eas.unige.ch/EWASS2017/session.jsp?id=S14

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***

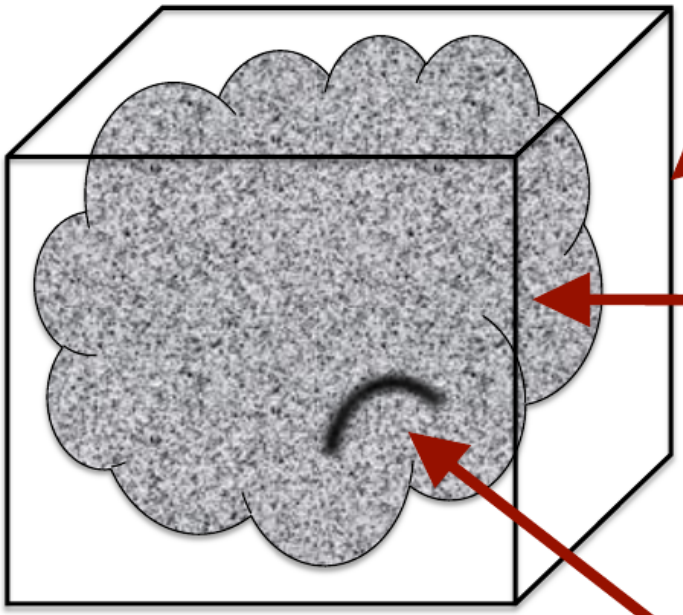
➡ ***Data driven vs. hypothesis driven science***

- 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 >> 3 parameter space hypercube

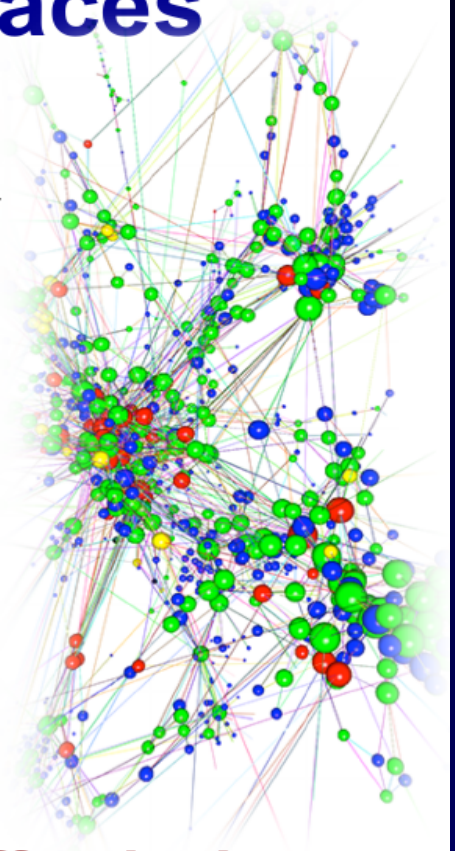
High-D data cloud: mostly noise, of an arbitrary distribution

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

Visualization in Machine Learning

A Key Challenge: Visualisating Multidimensional Data Spaces

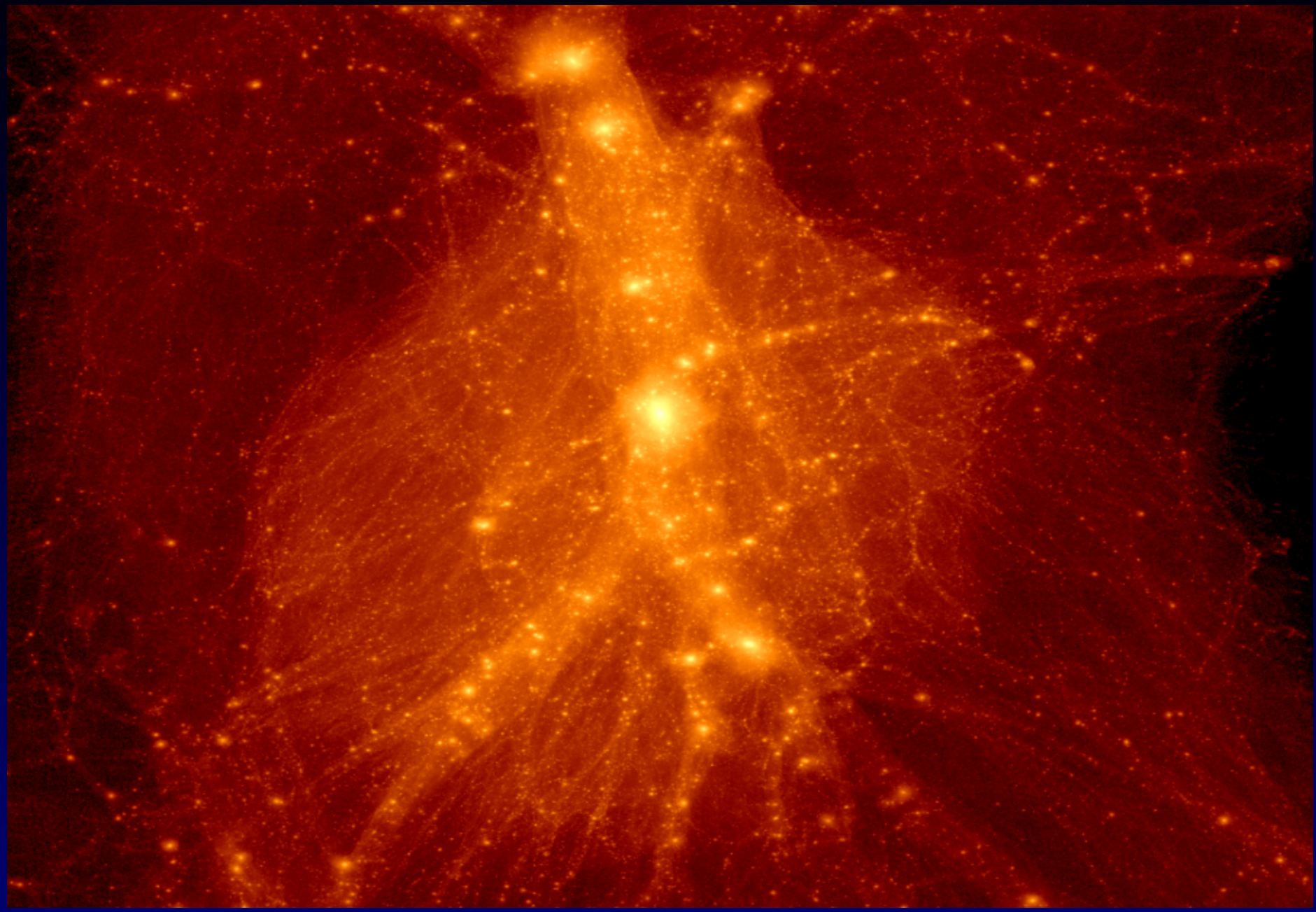
- Hyperdimensional structures (clusters, correlations, etc.) may be present in many complex data sets, whose dimensionality may be D ~ 10^2 – 10^4, or higher
- It is a matter of ***data understanding***, choosing the right data mining algorithms, and interpreting the results
- We are biologically limited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



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

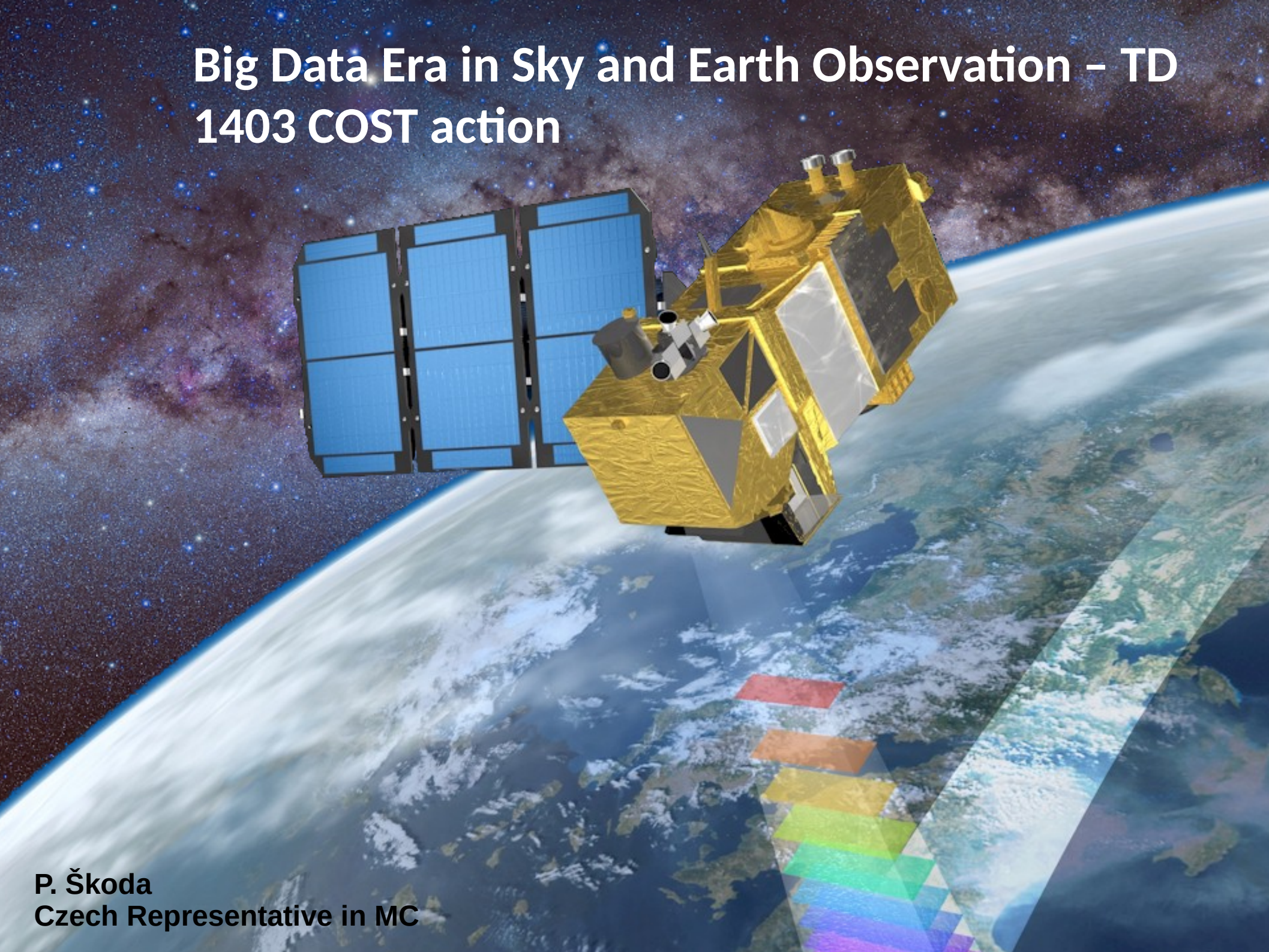
Djorgovski

Visualization of Big Data



Breddels 2016

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



P. Škoda
Czech Representative in MC

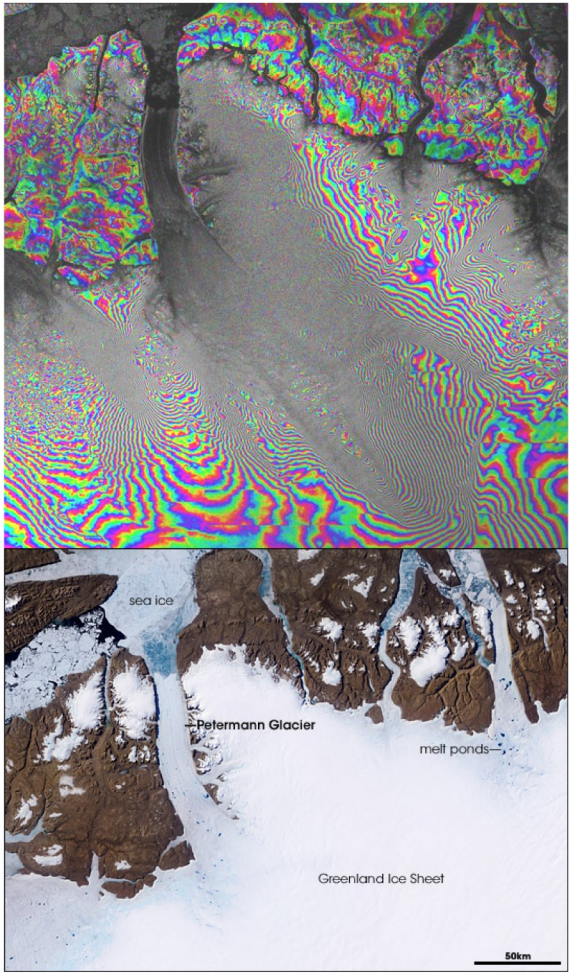
sentinel-1A

Launch date: 3. April 2014
up to **2.4 TB/day** of imaging radar data for **7 years** (fully open and free data access policy)

Applications: Oceans and ice, Changing lands, Emergency response

The era of Big Data has arrived!

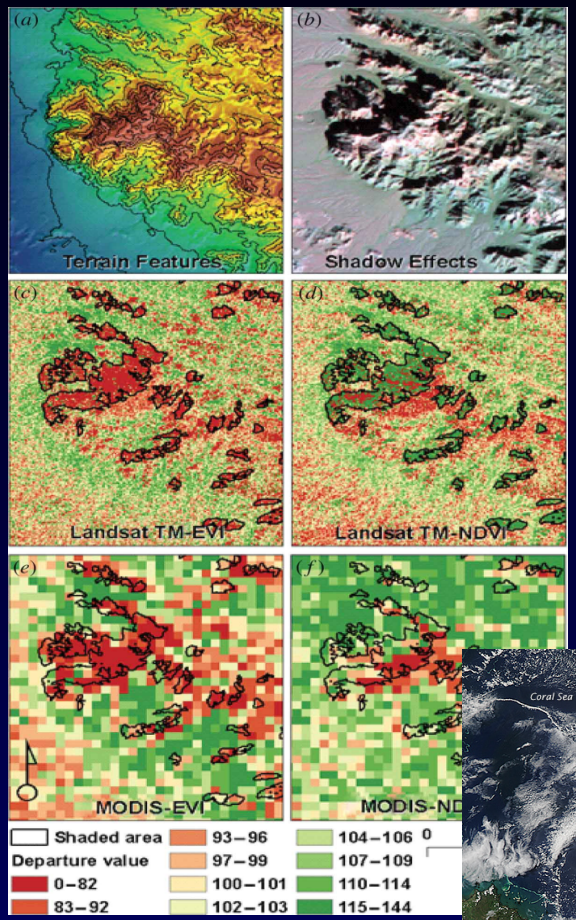
Example: images + time = surface movements

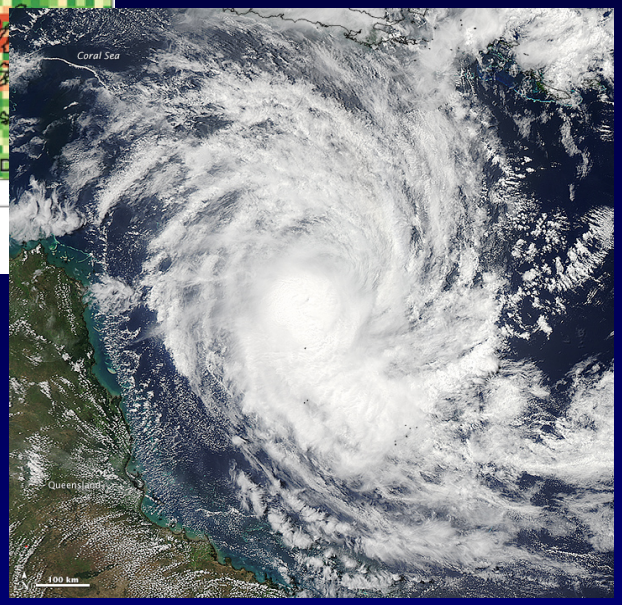


Part of the **European Earth Observation Programme Copernicus:** the most ambitious Earth observation programme to date: 30 satellites: peta-bytes now: zetta-bytes in a decade

opernicus
The European Earth Observation Programme

Remote sensing – Big Data
Machine Learning





Precise farming

Forestry

Ore mining

Water resources monitoring

Automatic classification of terrain

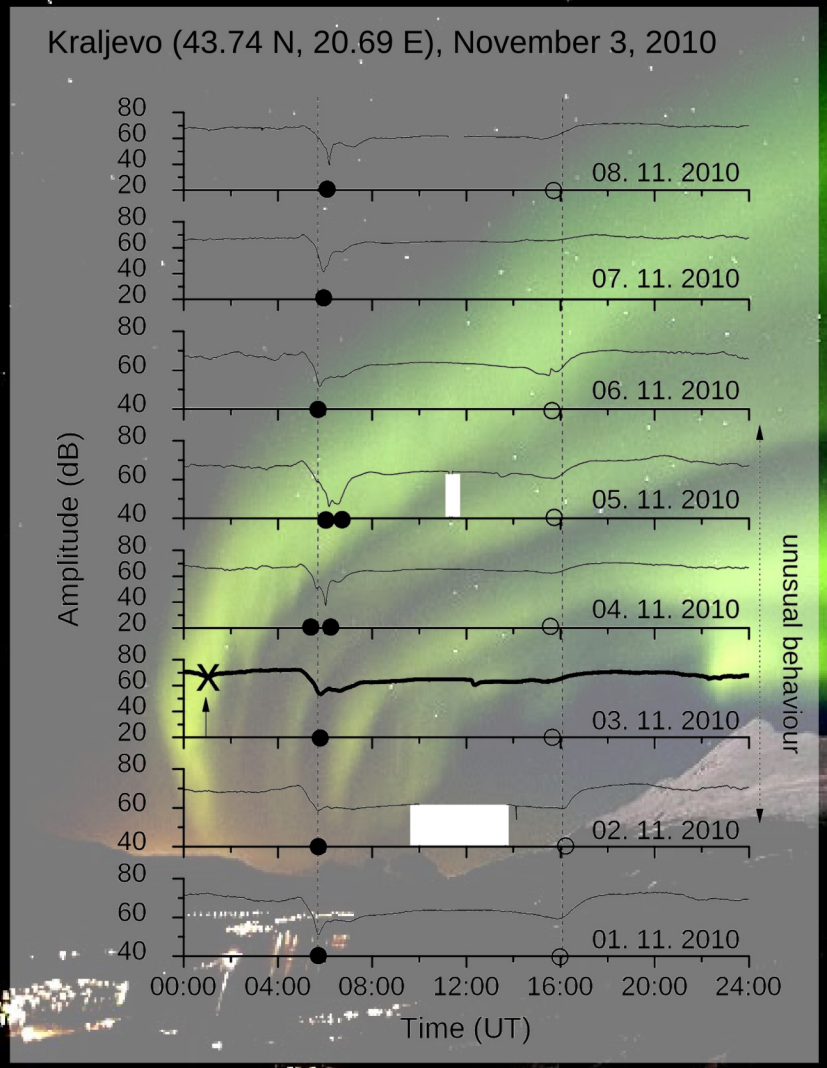
Resistence of buildings (Aquilla)

Mobile network structure – disasters

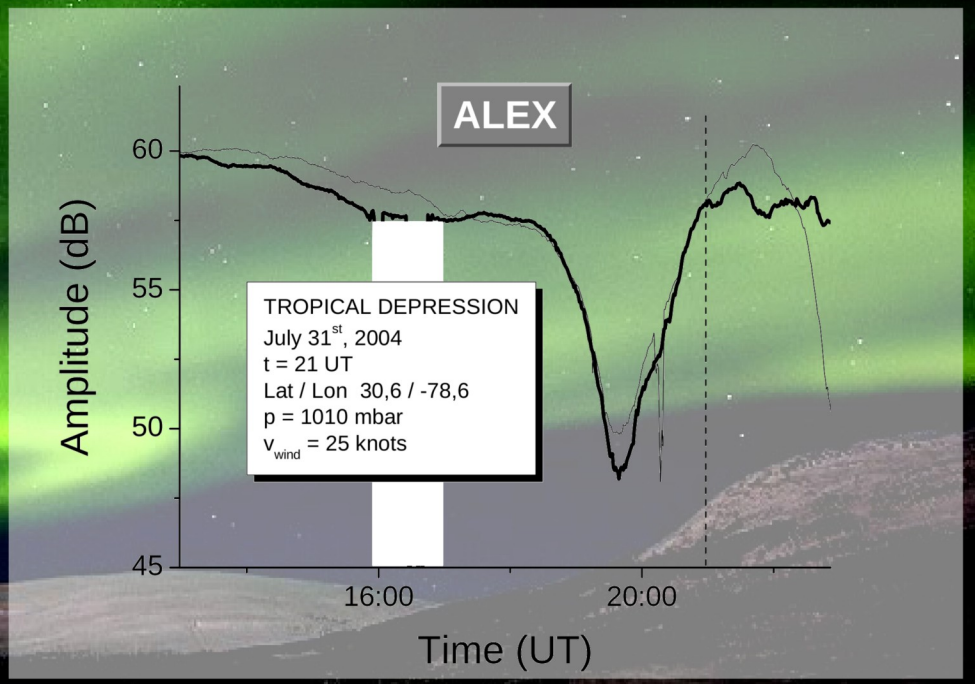
Customized product for every user

Ionospheric variations and natural disasters predictions

Earthquake

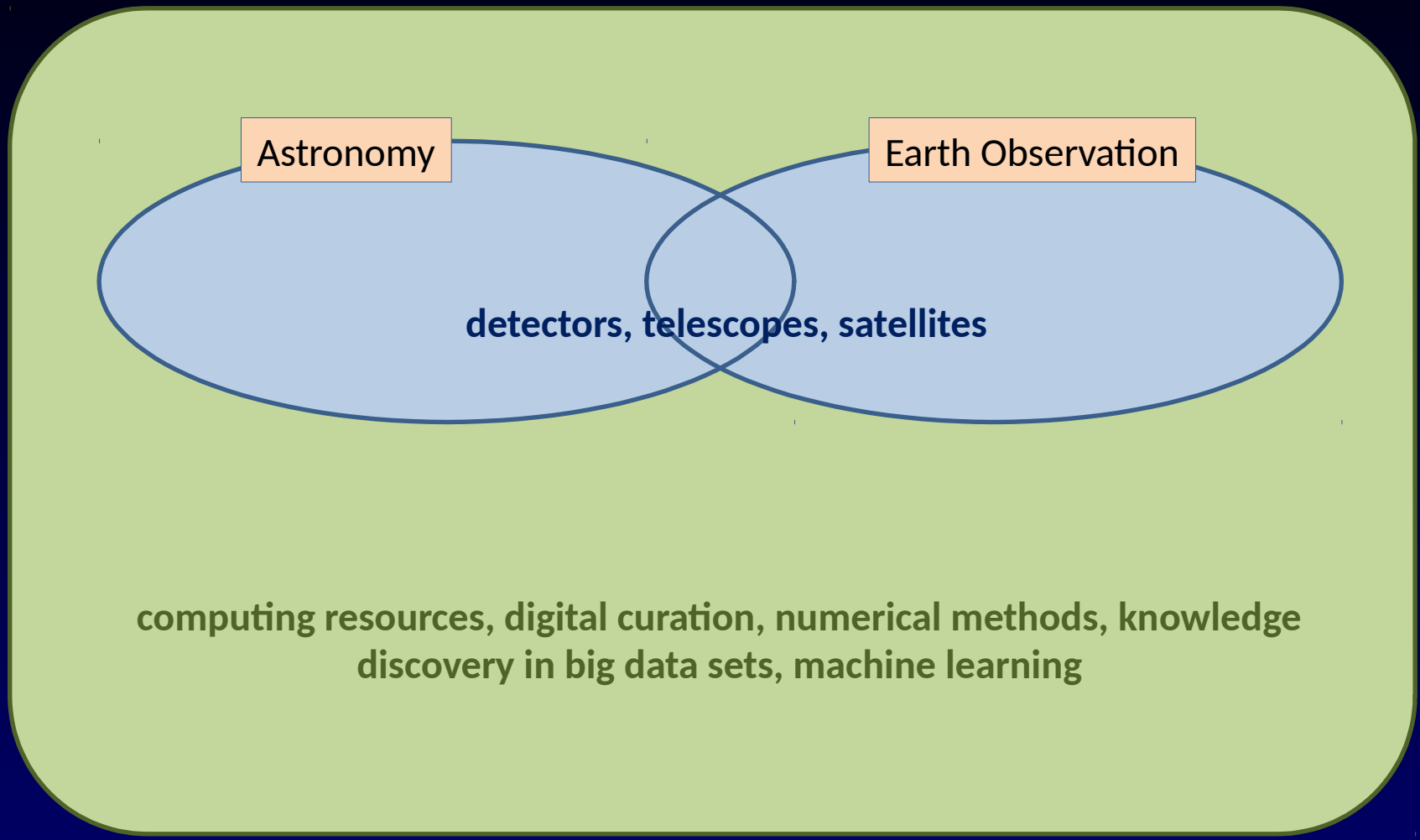


Cyclon



Nina 2016

Big Sky Earth Commons



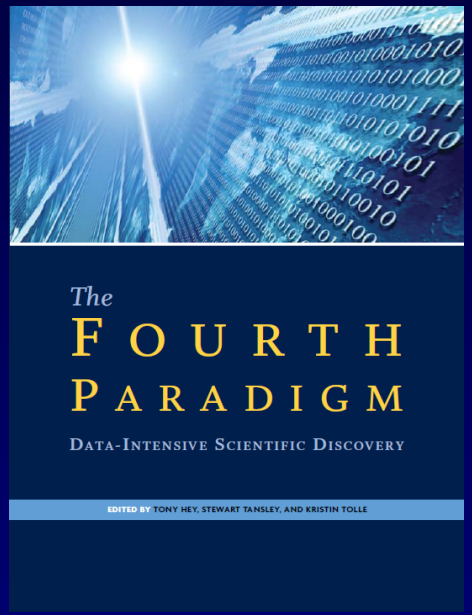
Vinkovic 2016

Impact

Big Data is not just bigger, it is different!

Success in research will depend on the ability to mine knowledge from that data.

And some of the most interesting science probably hasn't even been imagined!

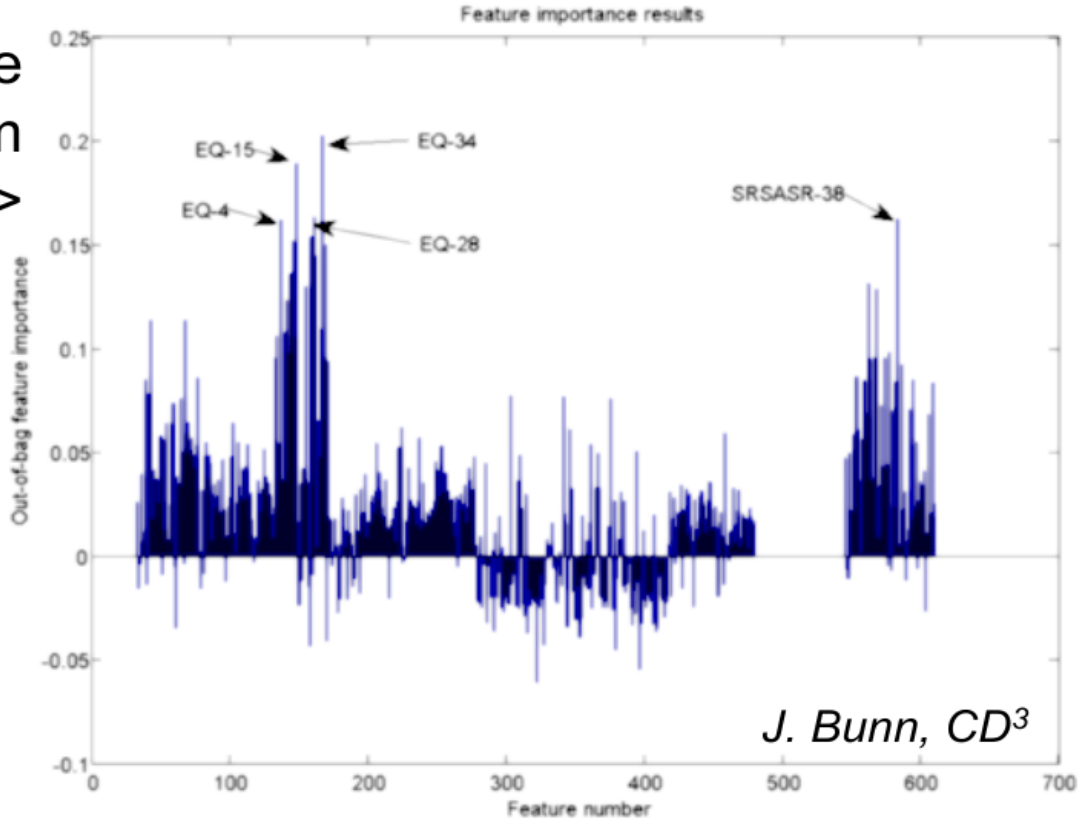


Astro-Neurology

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

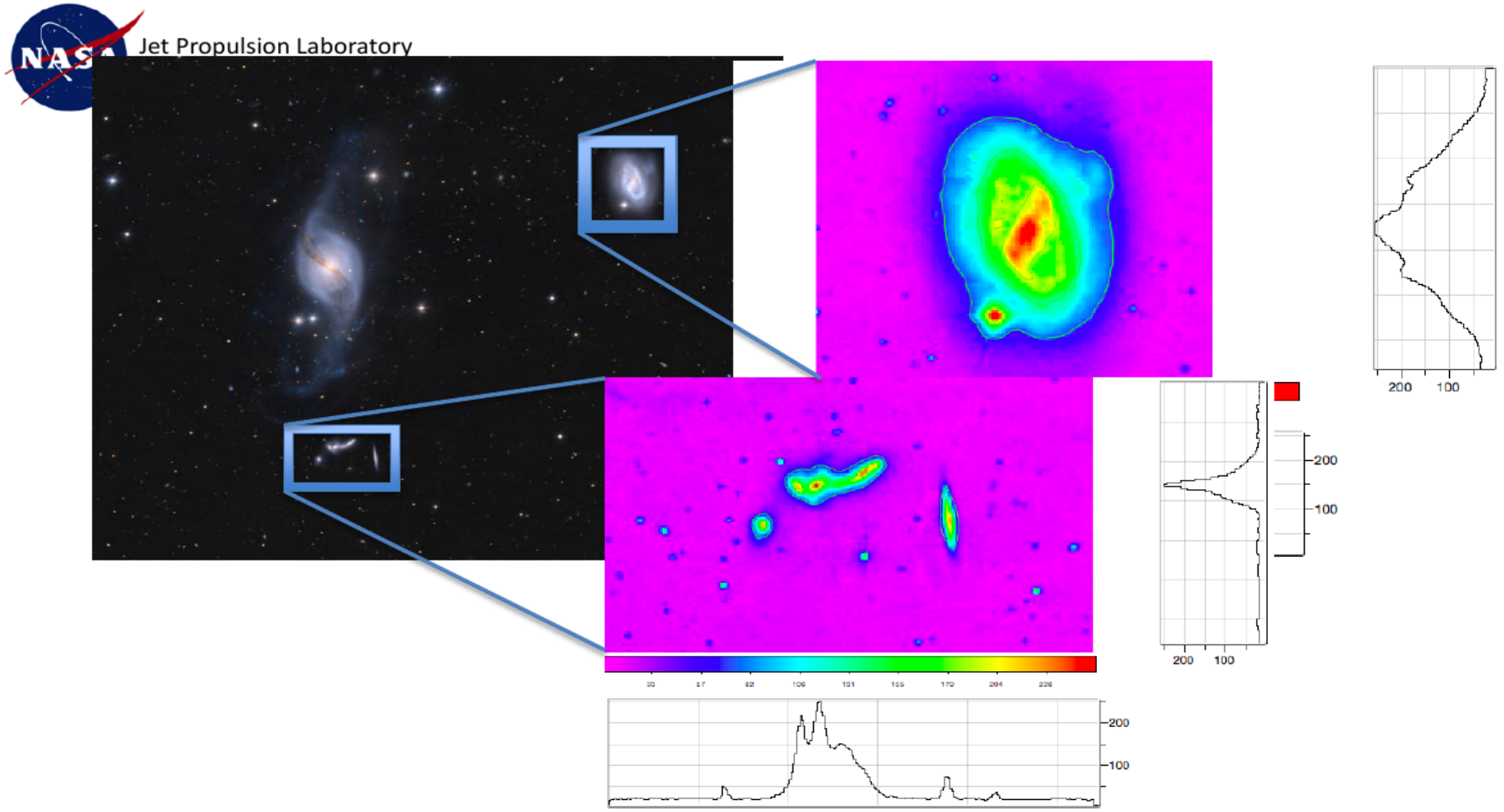
- Feature importance using random forests =>



- Next: correlate with MRI scans

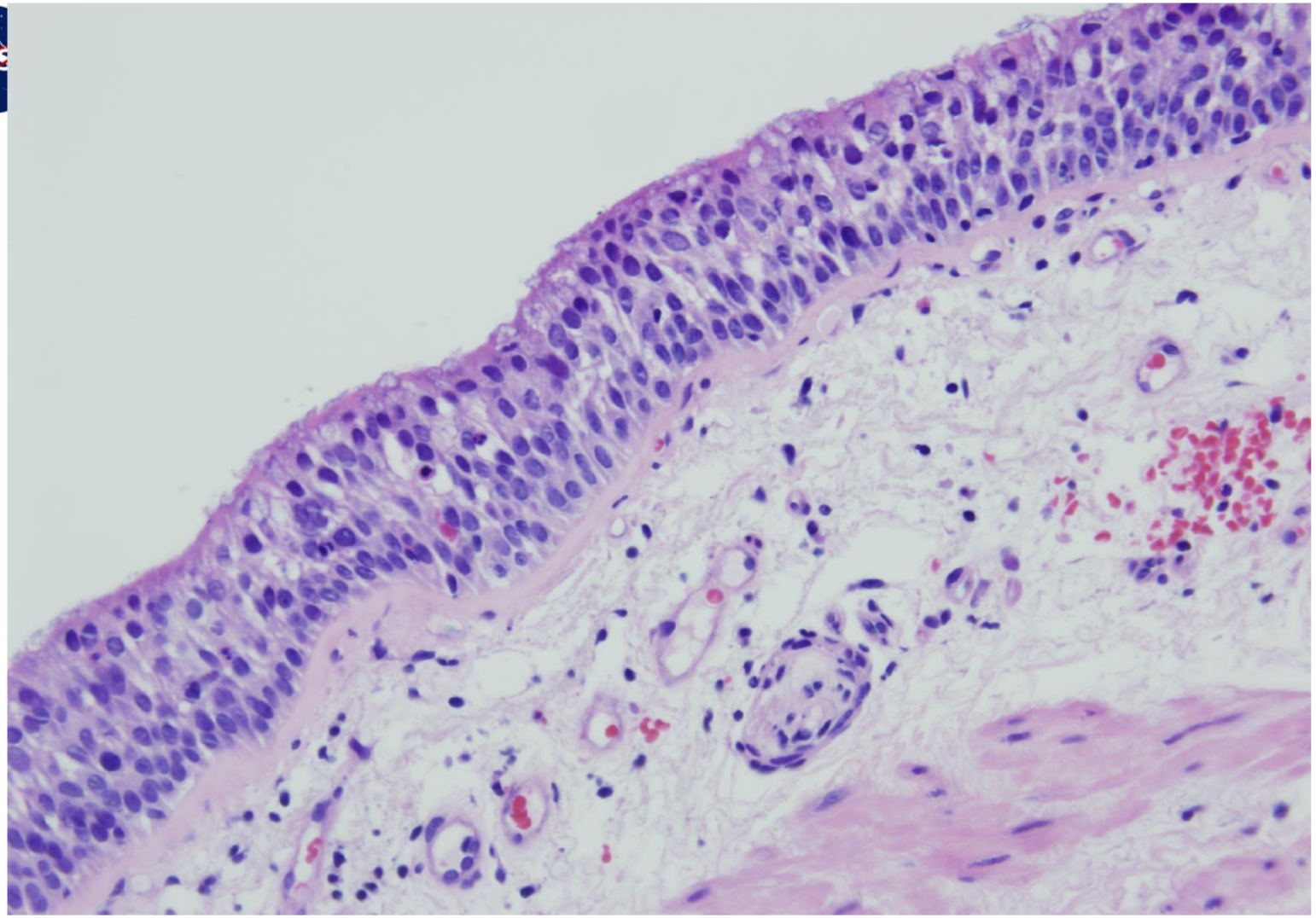
(with R. Adolphs et al.)

Finding Galaxies by Shape NASA



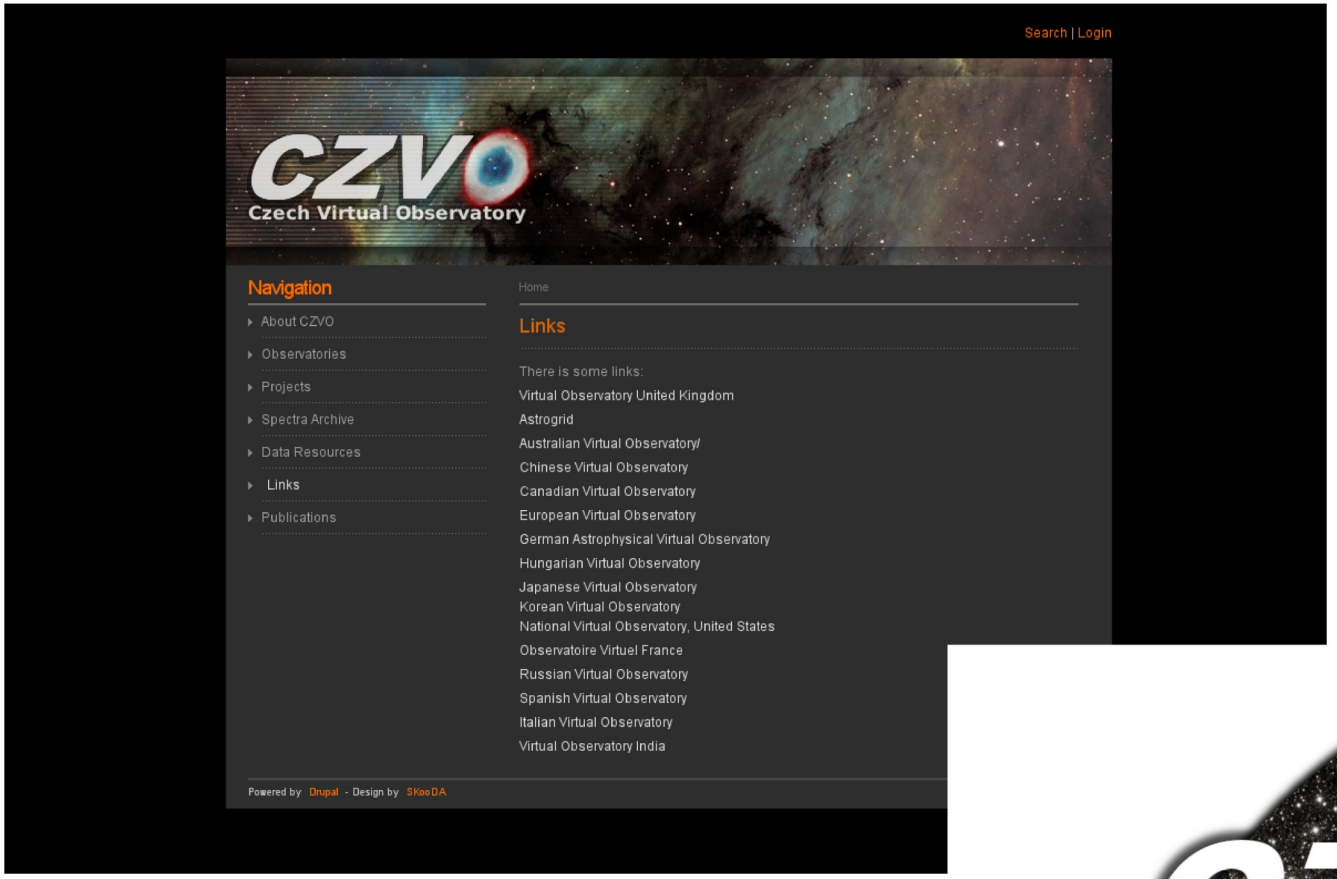
Description: Detecting objects from astronomical measurements by evaluating light measurements in pixels using intelligent software algorithms.
Image Credit: Catalina Sky Survey (CSS), of the Lunar and Planetary Laboratory, University of Arizona, and Catalina Realtime Transient Survey (CRTS), Center for Data-Driven Discovery, Caltech.

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

Czech VO – CZVO (not in IVOA)



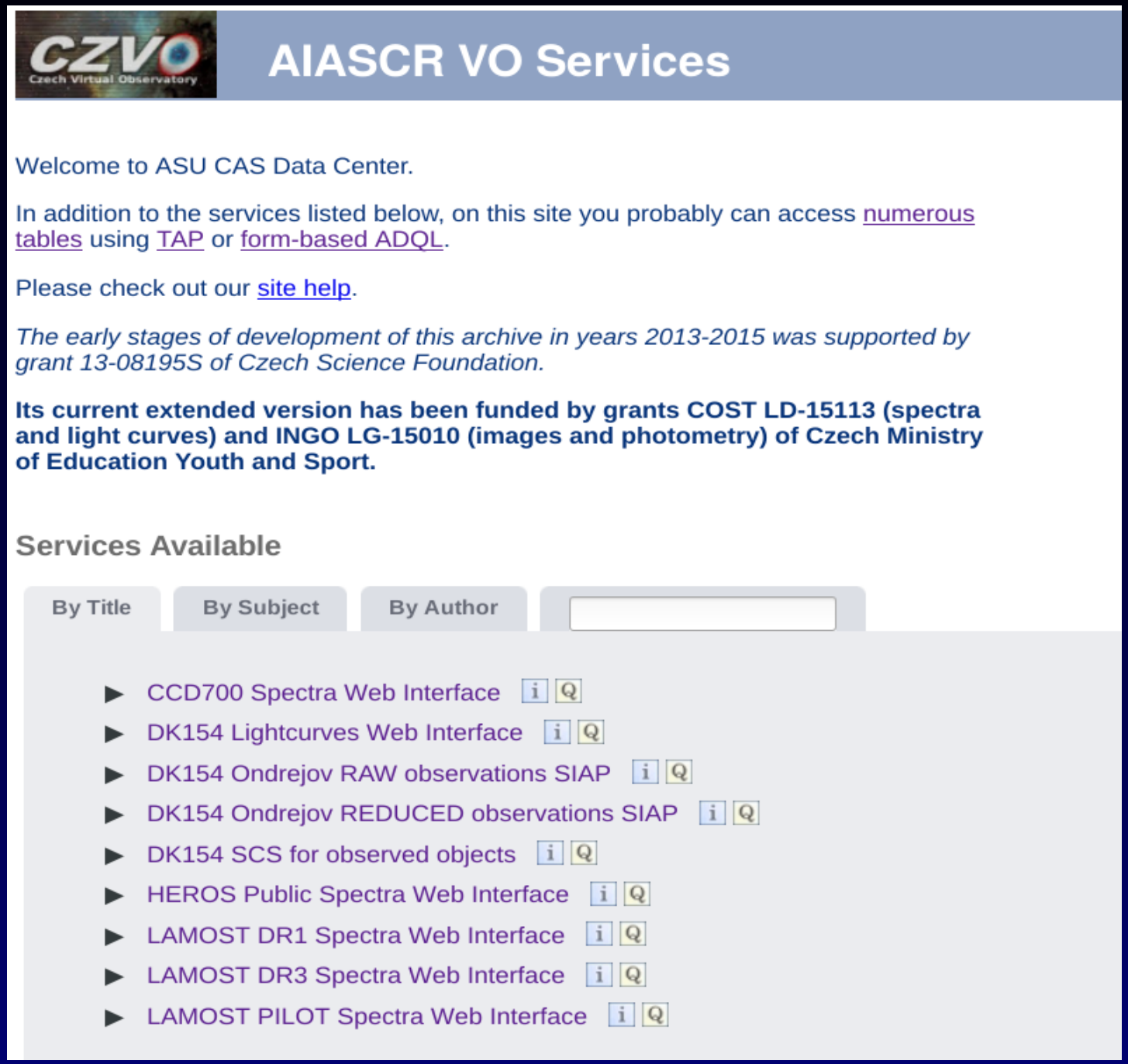
BT and MT at Faculties of IT astroinformatics and VO

- FIT VUT Brno
  - 2011 1 BT (Random Forests in Astronomy)
  - 1 PhD – Wavelets Dimensionality Reduction (pending)

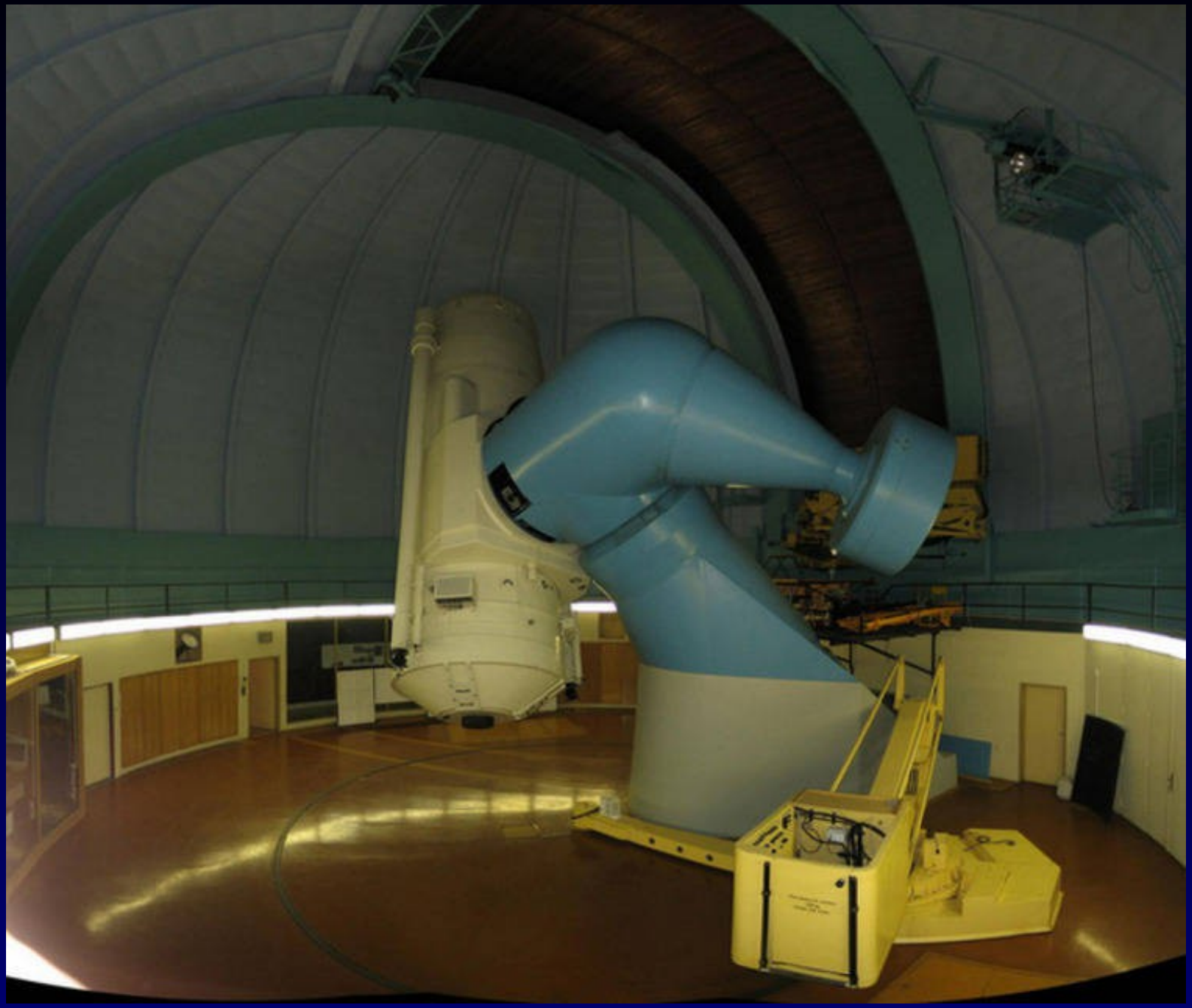
- VŠB-TU Ostrava
  - 2013+2015 1 BT + 1 MT - SPLAT-VO

- FIT ČVUT (8 BT+5 MT)
  - 2012 2 BT (VO-Korel+SSA proxy)
  - 2013 2 BT (OSPS Image + Catalogue Server)
  - 2014 2 BT (Random Forests + SOM)
  - 2015 1 BT (VO-Cloud)
    2 MT (Clustering OSPS + Deep Learning)
  - 2016 2 MT (Semisupervised learning + Outlier finding)
  - 2017 1 MT (VO Cloud) + 1 BT (deep learning)+ 1 PhD (VO light curve)

VO Services of CZVO

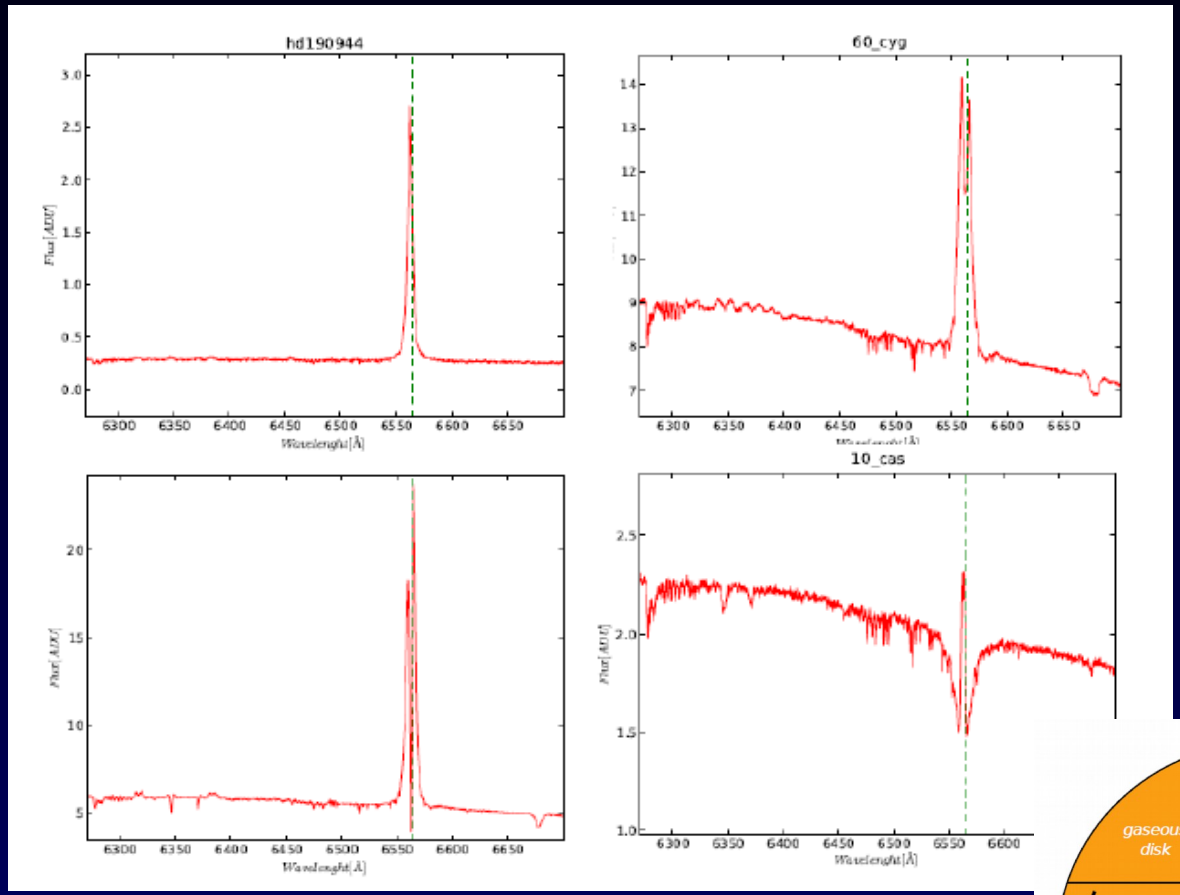


2m Perek telescope (1967)



Machine Learning of Spectra

Use case: ML of spectra profile of Halpha line (Be stars)

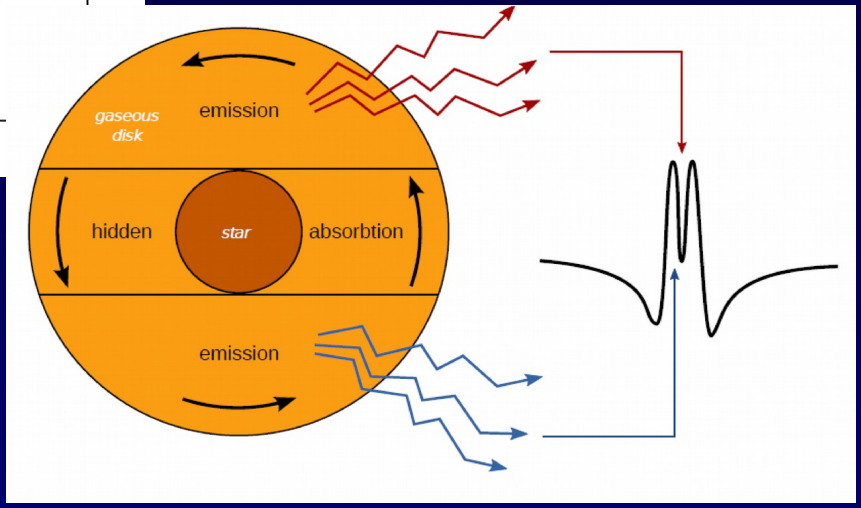


Be stars

Disk or envelope

Rotates, Hot

Origin ?????



Spectra - query output and previews

CZVO Czech Virtual Observatory

CCD700 Spectra Web Interface

Help | Service info

Related
CCD700 SSAP

Metadata
Identifier
ivo://asu.cas.cz/ccd700/q/
Description
CCD700 public web interfa
Keywords
Optical spectroscopy
Creator
[Logo]
Created
2014-05-09T10:57:00
Data updated
2016-06-14
Reference URL
Service info

Try ADQL to query our data.

Please report errors and problems to the site operators. Thanks.
Privacy | Disclaimer
Log in

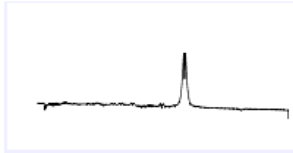
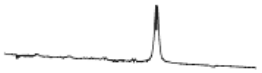
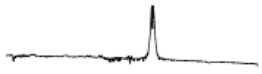
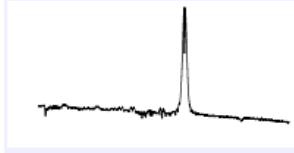
Parameters

- Object standard name: ['psiper']

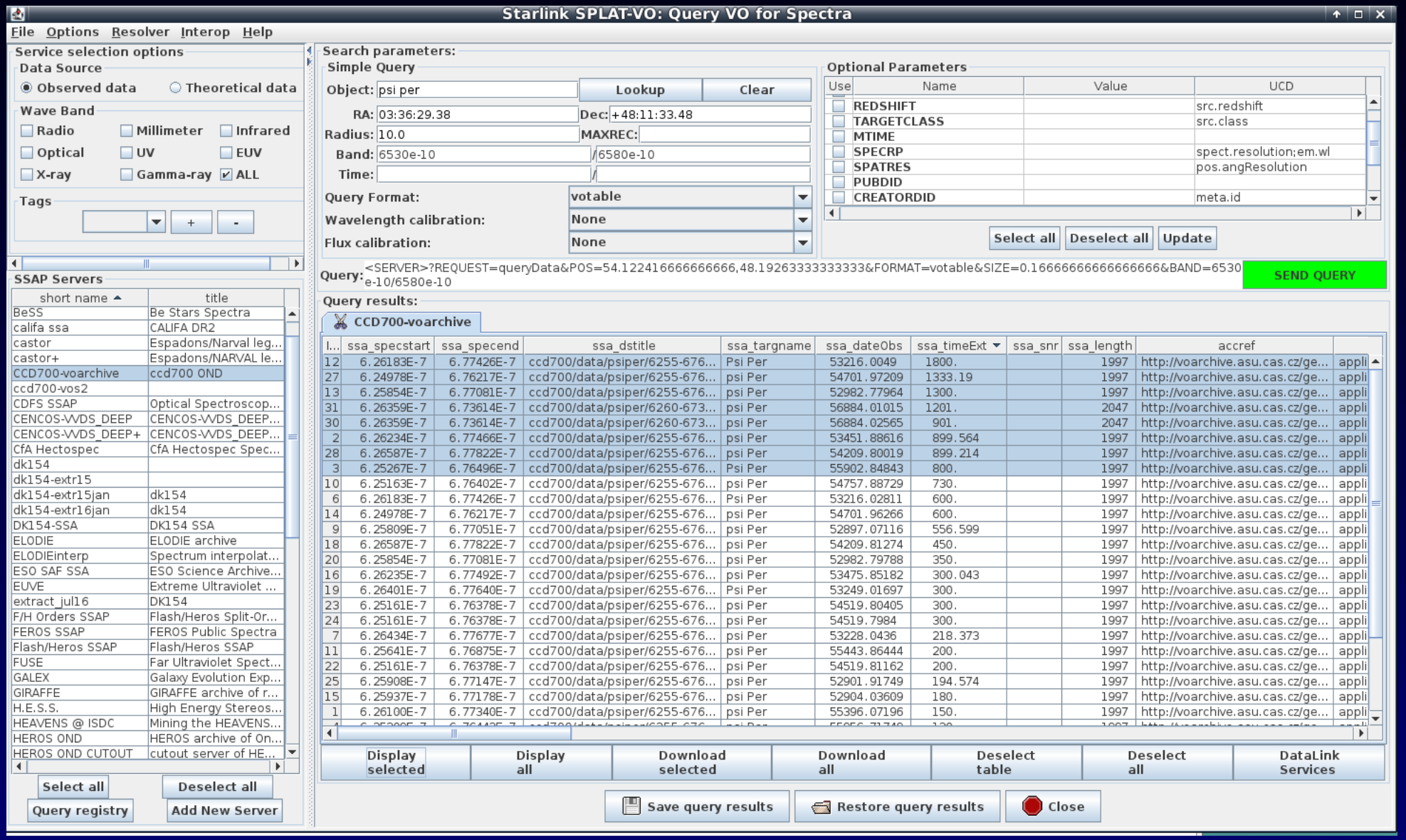
Result

Matched: 44

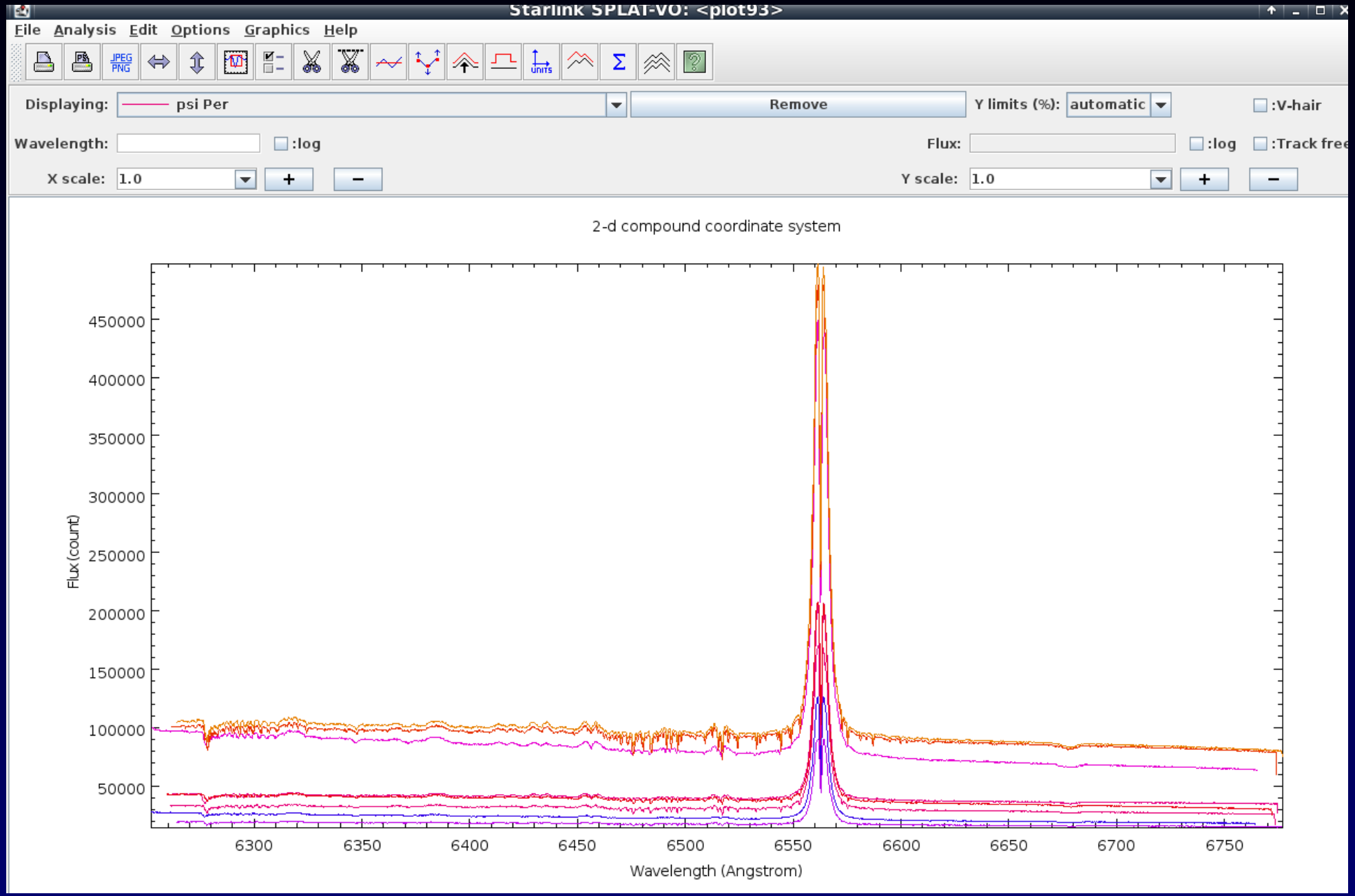
Send via SAMP | Quick Plot

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
 oc220022.fit	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
 ul070006.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
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
 nh100015.fit	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
 nh310030.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

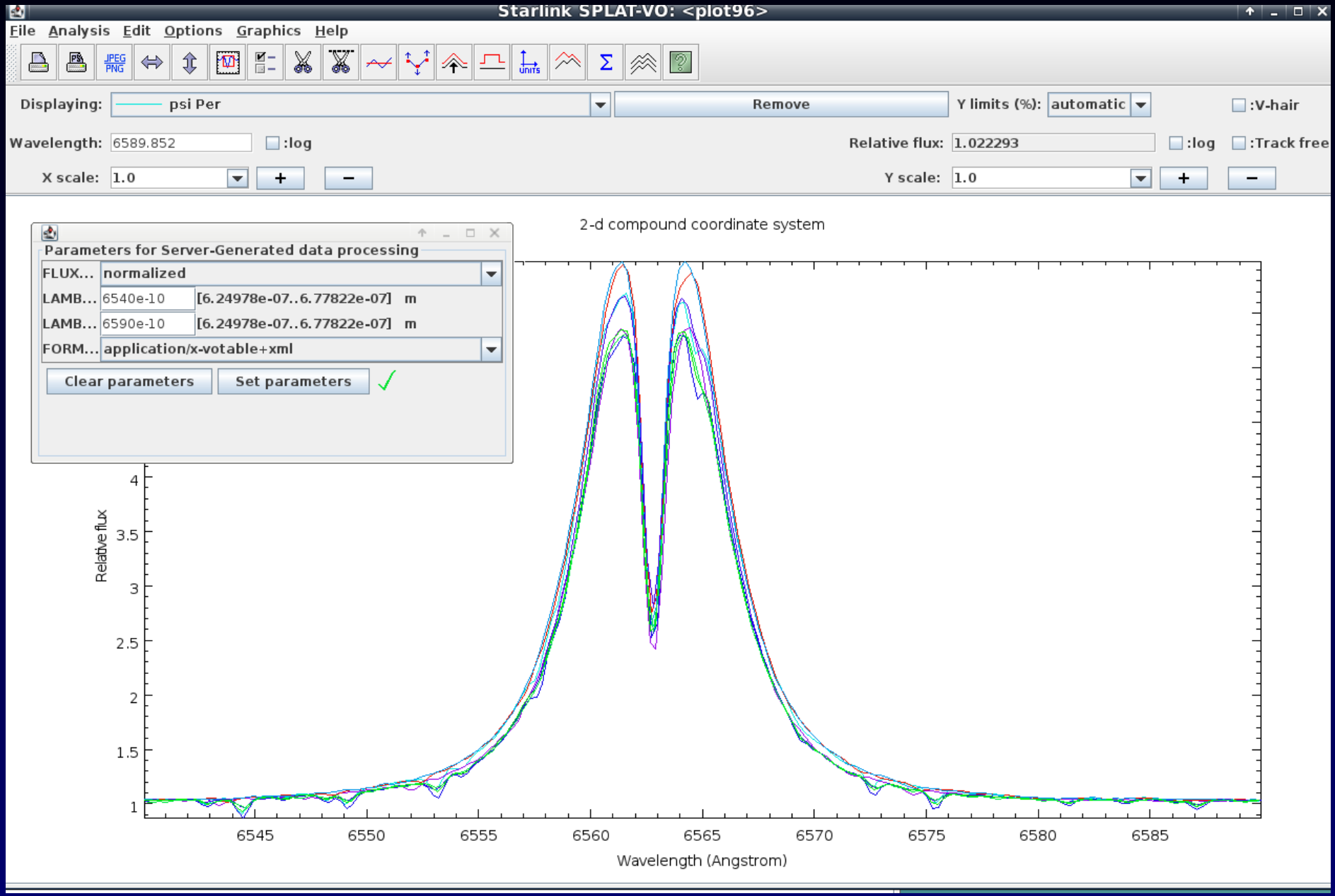
Spectra in SPLAT-VO - query



Spectra in VO – direct access plot

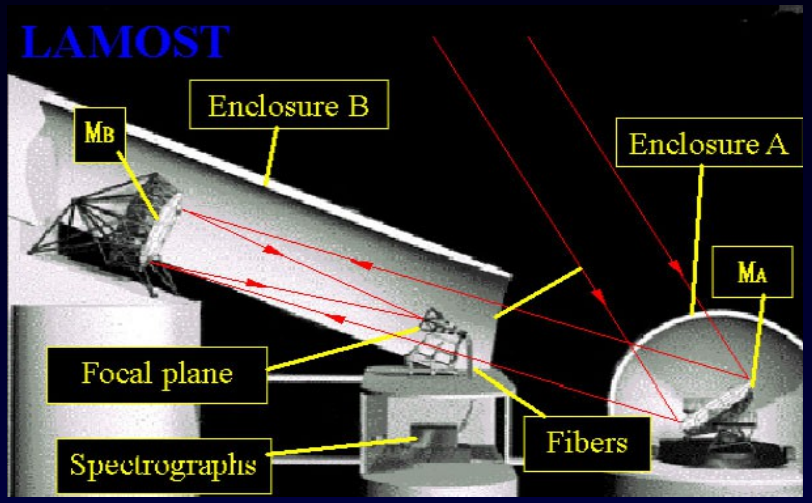


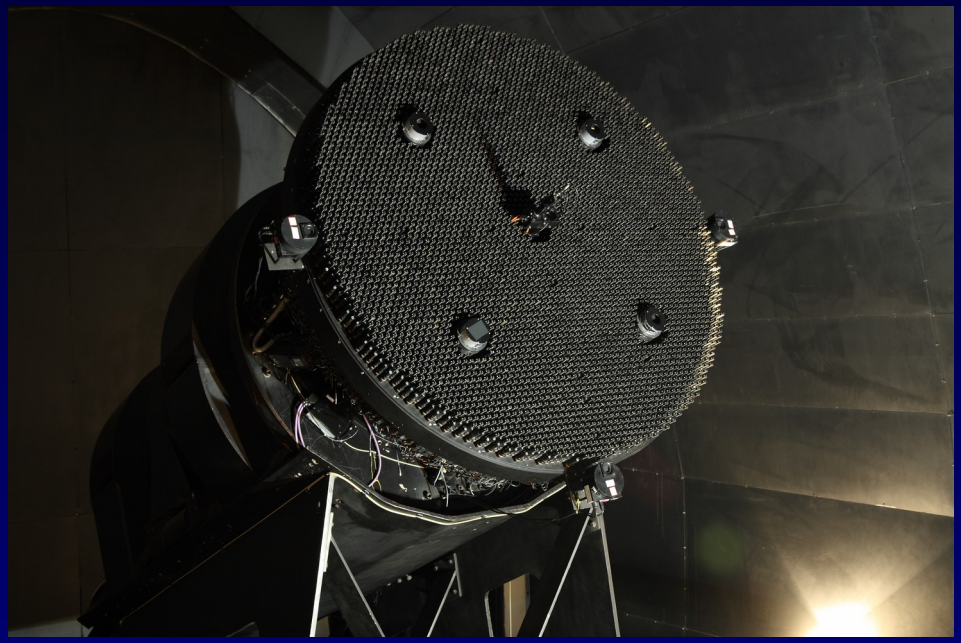
Spectra in SPLAT-VO - DataLink

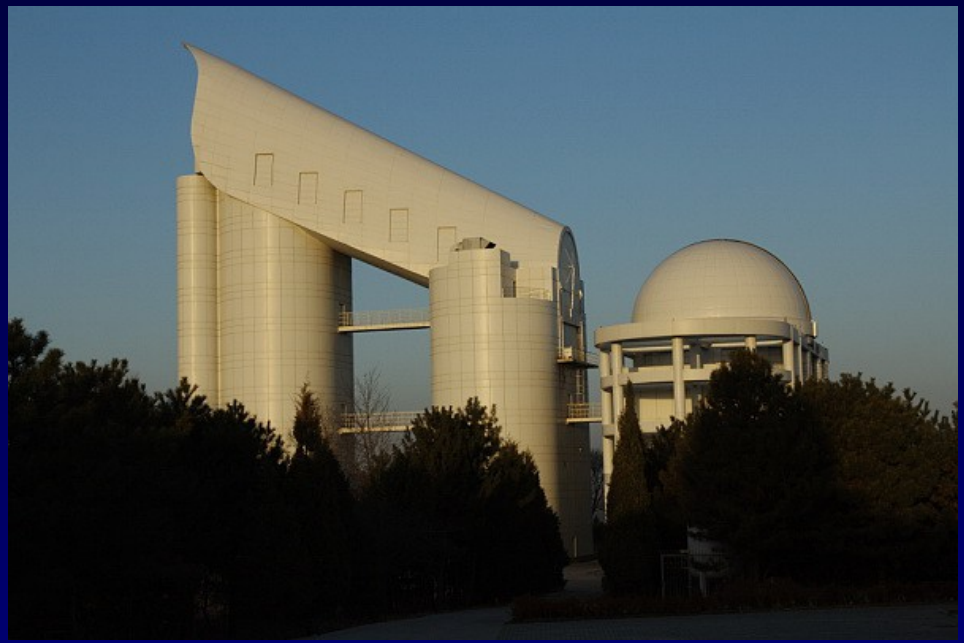


LAMOST (Guoshoujing)

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

SOM Worker example

VO-CLOUD DETAILS OF JOB

Home | Manage filesystem | Jobs | Download history | Create job | Jupyter | Settings | Admin | Help | Logout (admin)

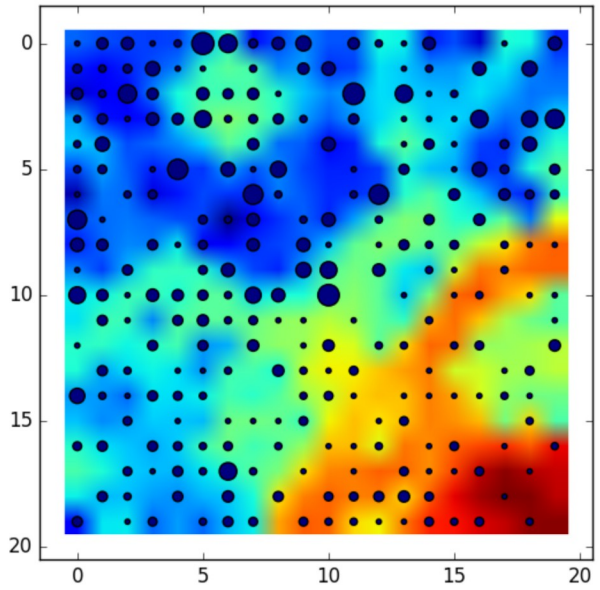
AllSpecOndSOM

Type	Id	Phase	Worker	Created	Started	Finished	Executing time
SOM	1-92	COMPLETED	local worker	2/14/17 11:32:47 PM	2/14/17 11:32:47 PM	2/14/17 11:32:59 PM	11 sec

Run again | Delete

Preview

index.html - Fullscreen



Files

VO-CLOUD DETAILS OF JOB

Home | Manage filesystem | Jobs | Download history | Create job | Jupyter | Settings | Admin | Help | Logout (admin)

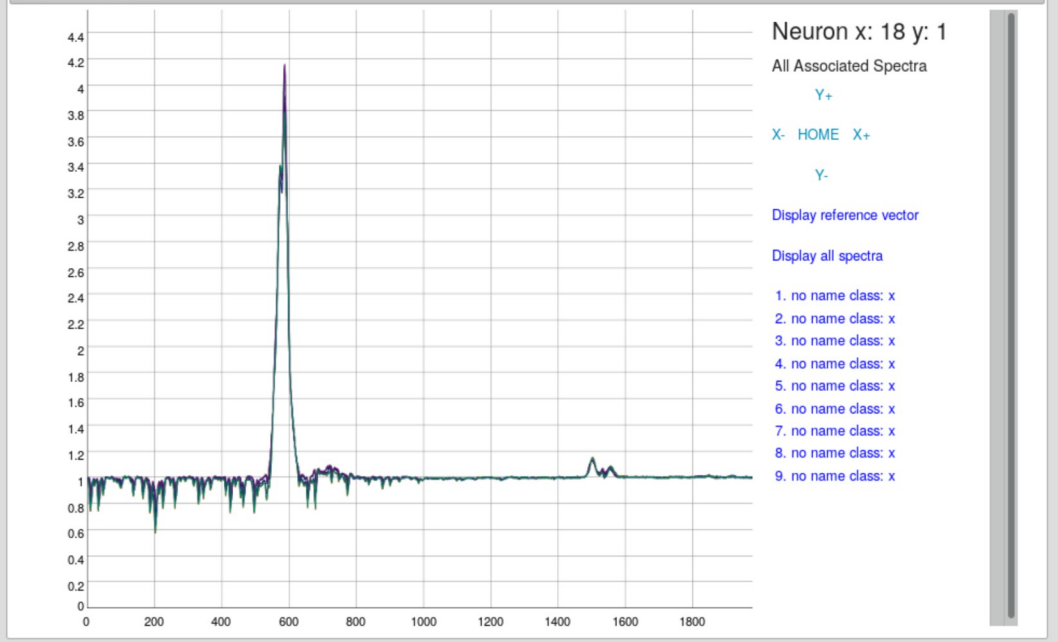
AllSpecOndSOM

Type	Id	Phase	Worker	Created	Started	Finished	Executing time
SOM	1-92	COMPLETED	local worker	2/14/17 11:32:47 PM	2/14/17 11:32:47 PM	2/14/17 11:32:59 PM	11 sec

Run again | Delete

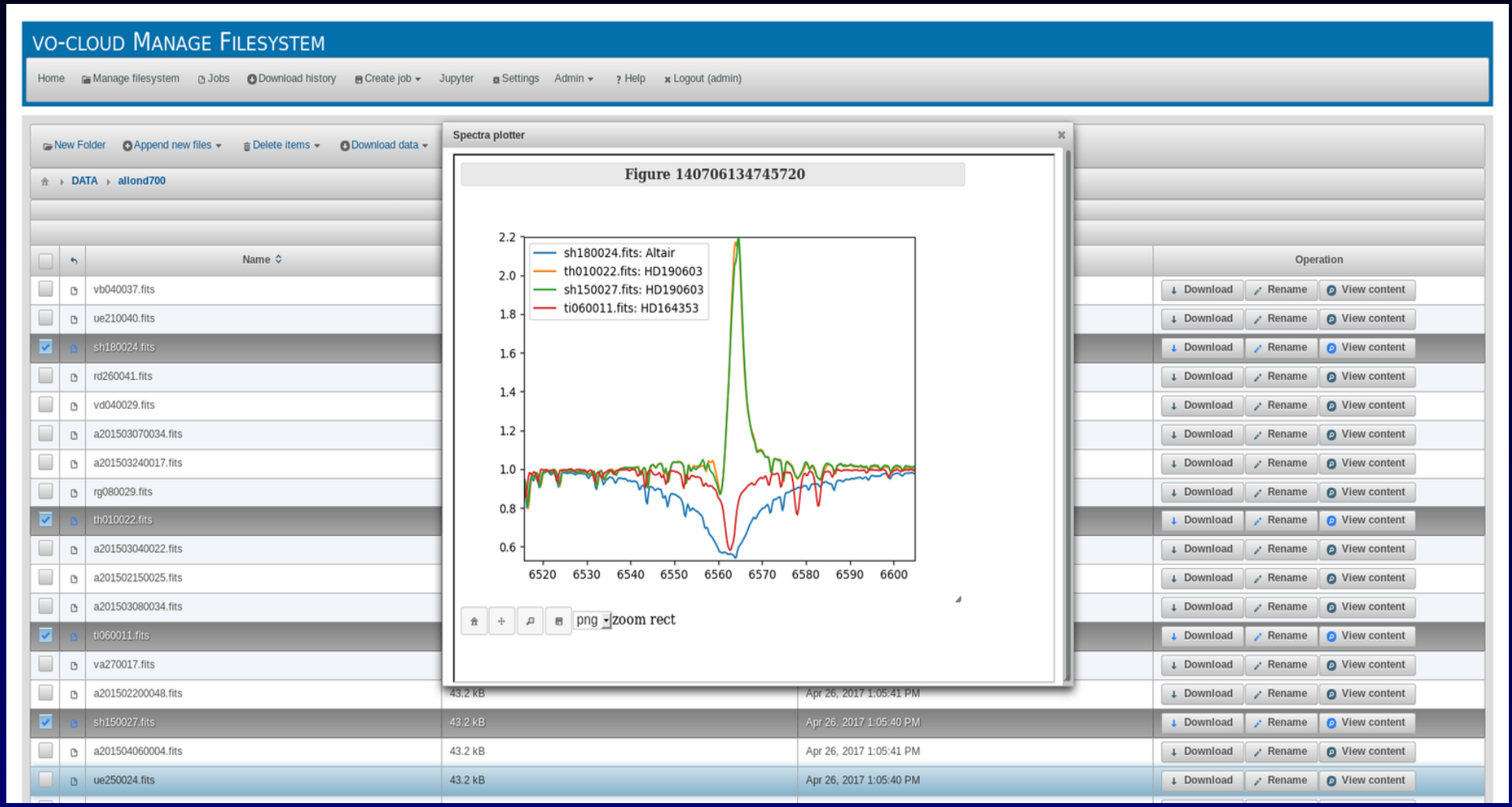
Preview

index.html - Fullscreen

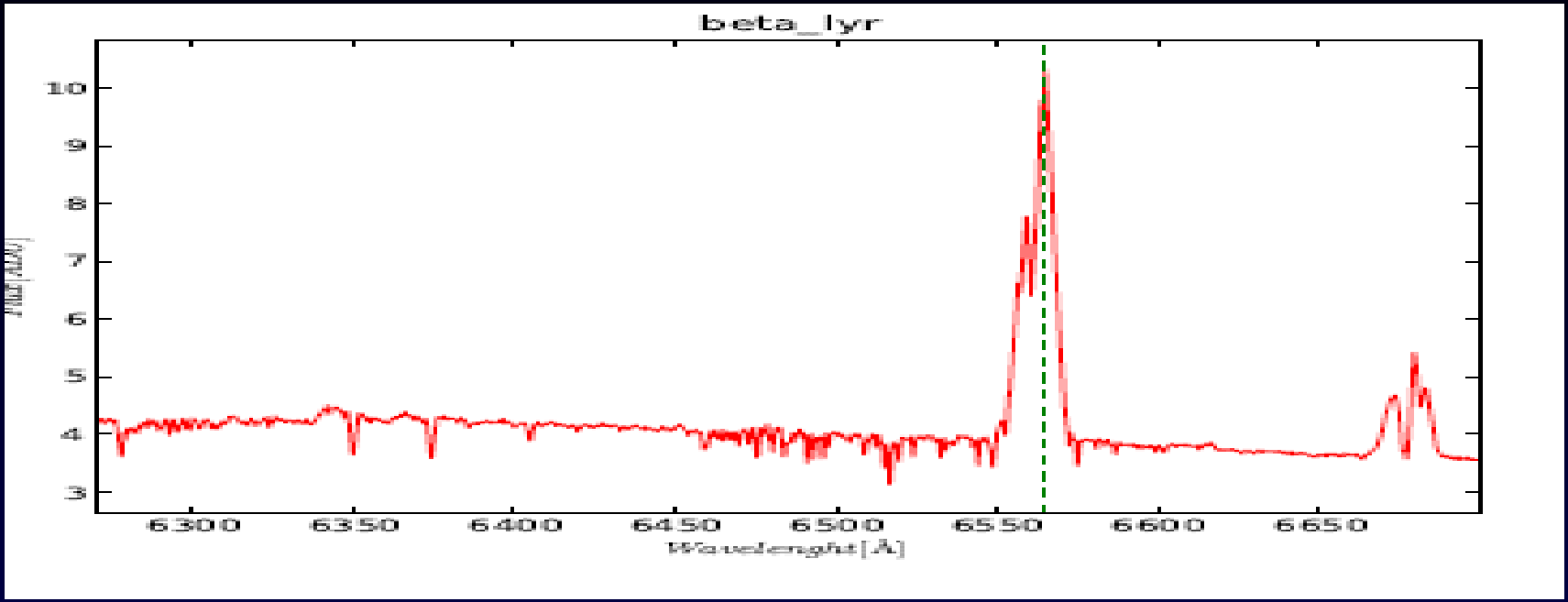


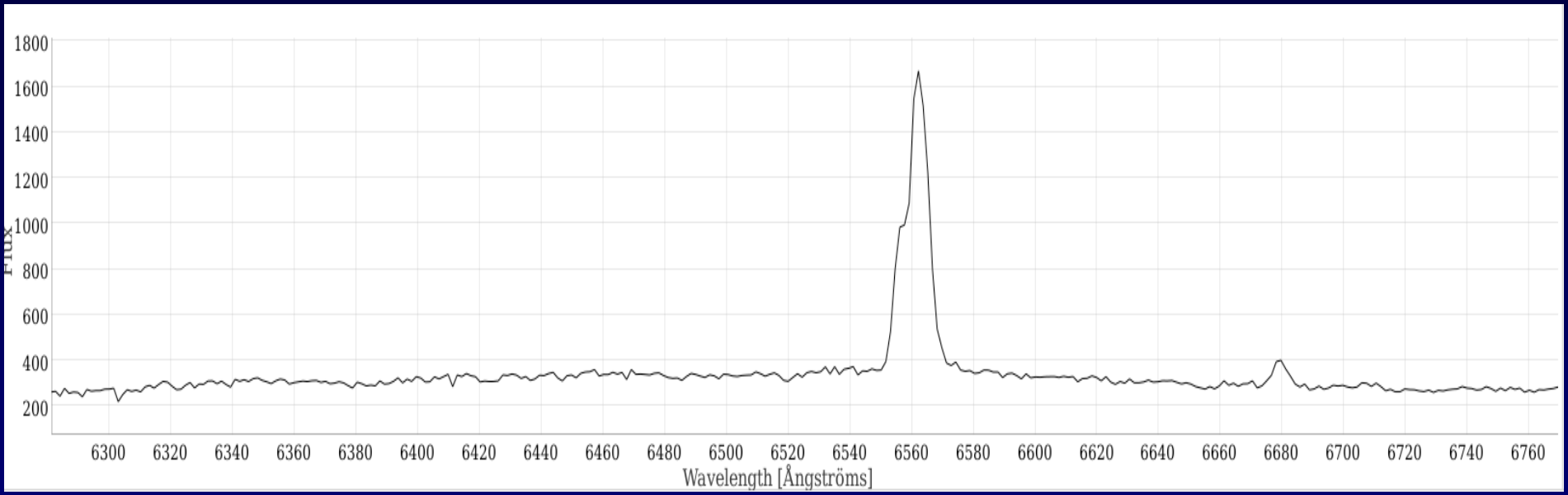
Files

VO-CLOUD spectra visualisation

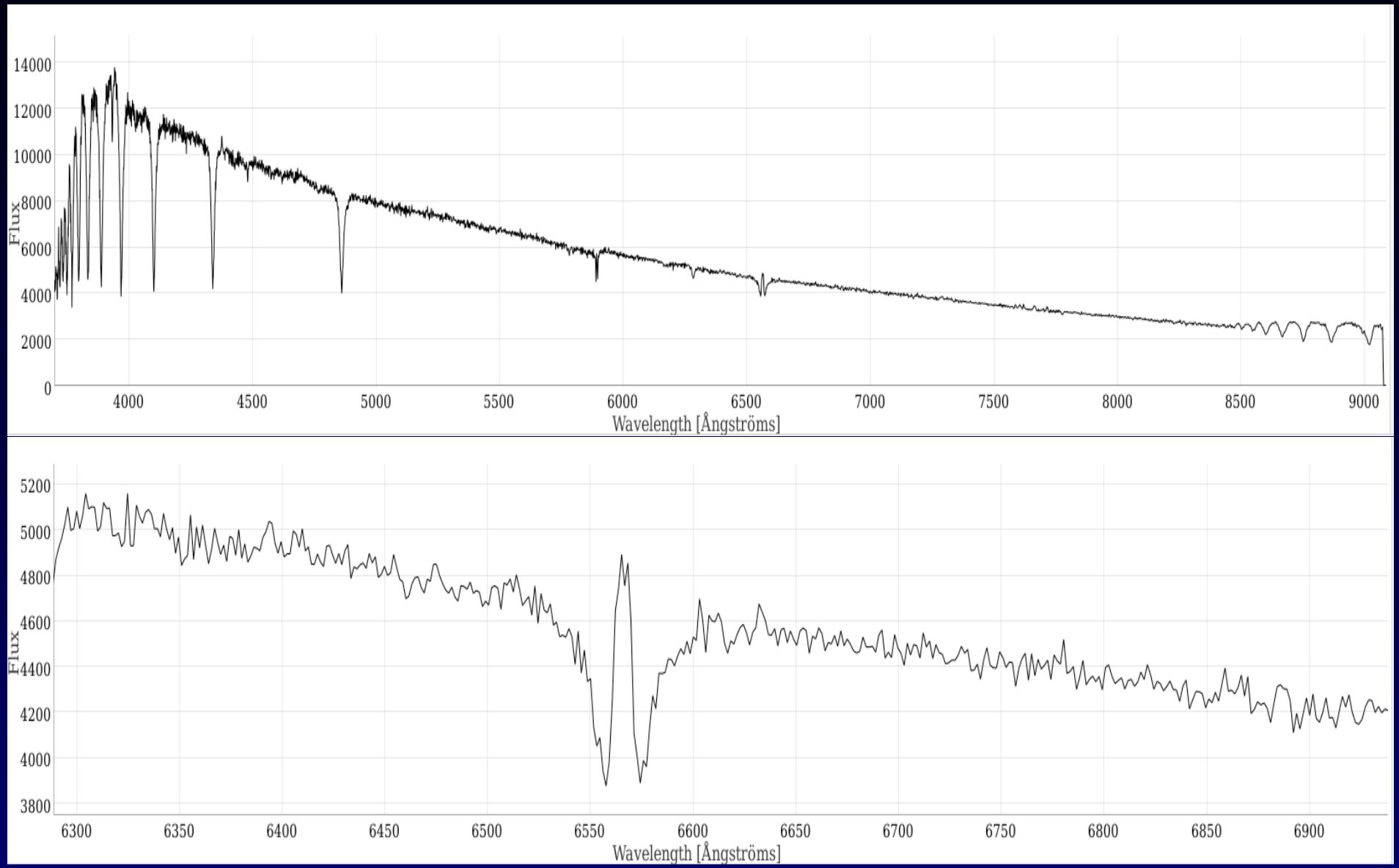


Be Candidates Found





Yet Unknown Be Star (UCAC ...)



Virtual Observatory inside

- OND 2m archive on SSAP protocol (spectra access)
- LAMOST DR1 on SSAP (using DaCHS)
- Preprocessing (rectify, cutout) – DataLink on server
- SAMP (send spectra to SPLAT-VO - view details)
- Visualization on sky ALADIN, X-MATCH CDS

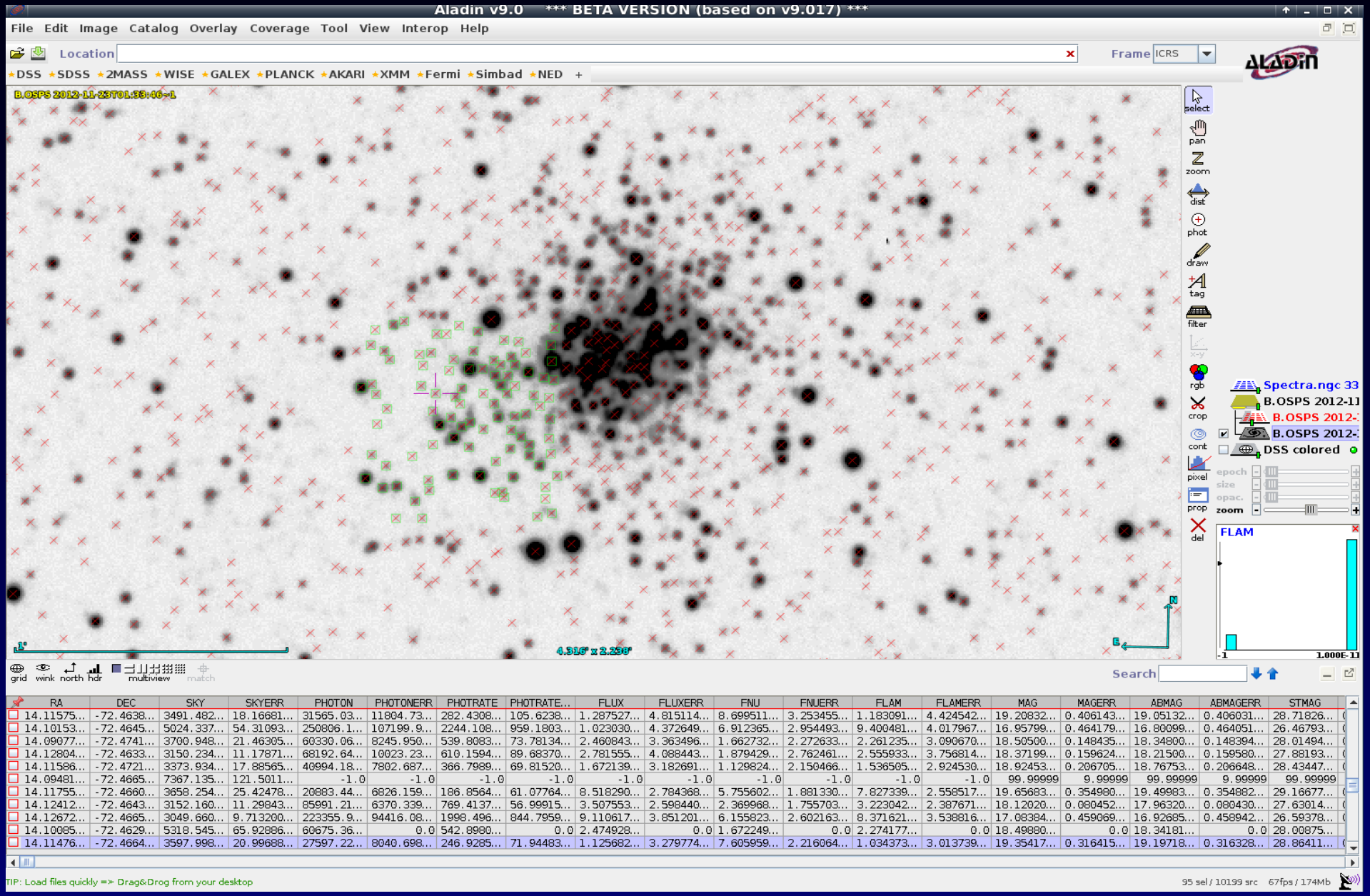
- VO-CLOUD – cloud engine based on UWS REST jobs
- Cross-matching (ADQL, TAP, TOPCAT, TAPhandle, pyVO, Vizier)

- Very similar methodology in e.g. Mass spectrography
- DNA analysis – search for patterns

Danish 1.54m at La Silla robotized in Summer 2012

Danish 1.54m Telescope

Reduced OSPS image + bintable photometry in 2nd extension



SIAP – Raw images query

Raw images results

DK154 Ondrejov RAW observations SIAP

Parameters

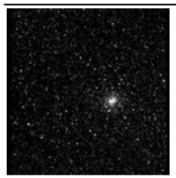
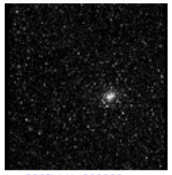
- Band: ALL
- Field size: 0.05
- File format: image/png
- Img_type: ALL
- Position: ngc 330

Result

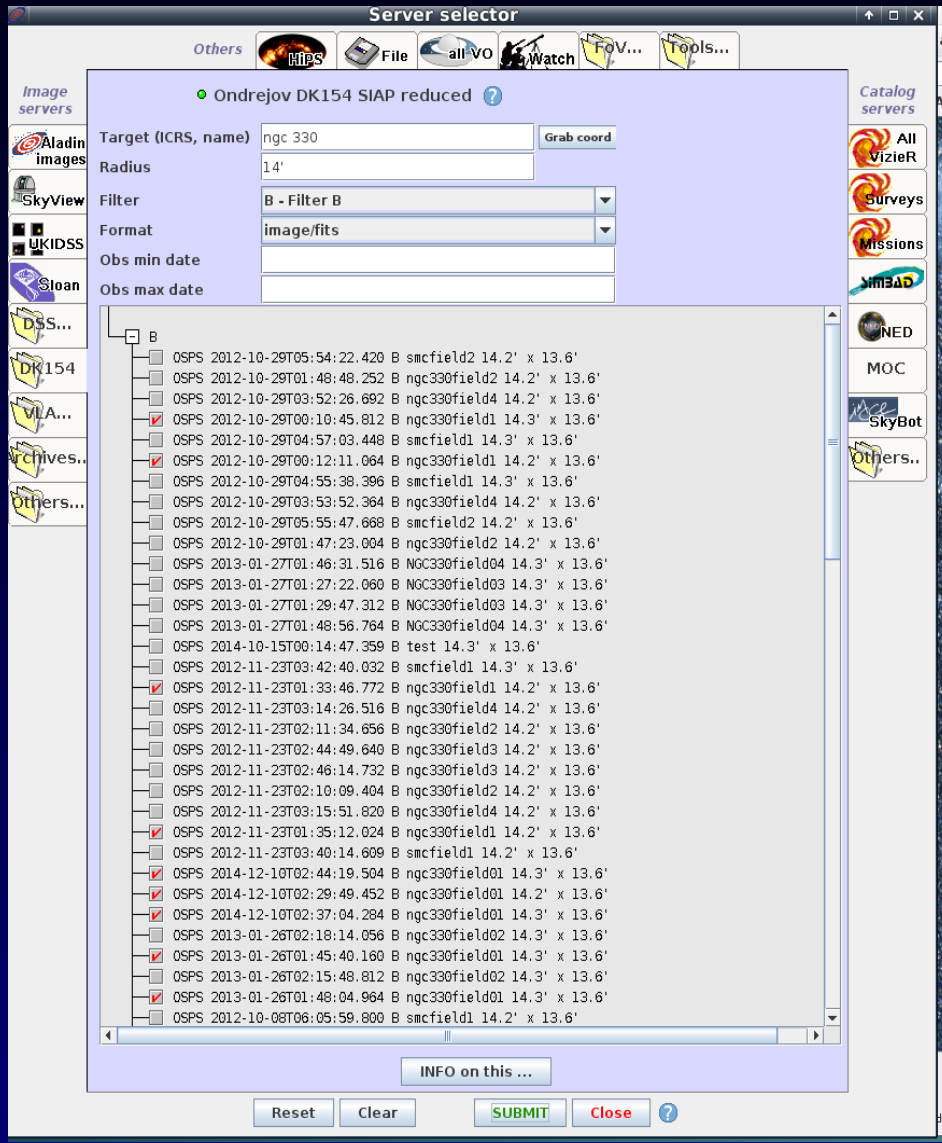
Matched: 100

Send via SAMP | Quick Plot

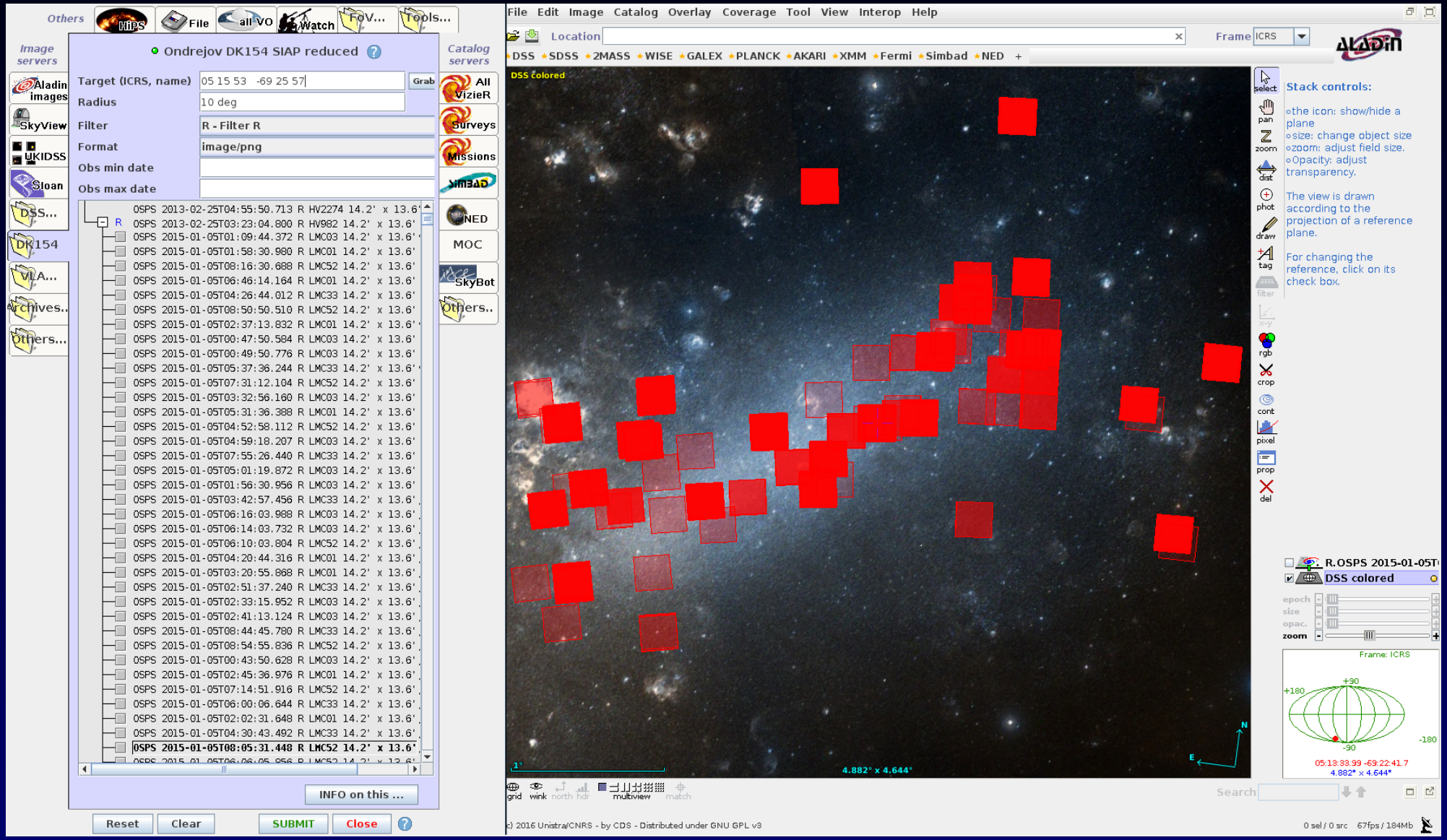
The query limit was reached. Increase it to retrieve more matches. Note that unsorted truncated queries are not reproducible (i.e., might return a different result set at a later time).

Accref	Owner	Embargo ends	Type	File size [byte]	Ctr. RA [deg]	Ctr. Dec [deg]	Title	Instrument	Obs. date	#axes	Axes Lengths [pix]	Scales [deg/pix]	Ref. Frame	Equinox [yr]	Proj.	Ref. pixel [pix]	Ref. values [deg]	CD matrix [deg/pix]	Bandpass	Bandpass unit	Band Ref. [m]	Band upper [m]	Band lower [m]	P. Flags	Coverage [deg]	Exp. time [s]	Telescope	Img_type
 ngc330field1_000003.png	beusers	2010-12-31 00:00:00	image/png	247.6kiB	14.19	-72.45	OSPS 2012-10-08T03:32:47.212 V ngc330field1	DFOSC_FASU	2012-10-08T03:32:46Z	2	[2148, 2048]	[0.000109996, 0.000109996]	ICRS	2000.0	TAN	[1074.0, 1024.0]	[14.193875, -72.4533333333333]	[-0.000109996, -9.82956e-07, -9.82956e-07, 0.000109996]	V	m	5.4e-07	6.7e-07	4.85e-07	N/A	Polygon ICRS 14.5844654144 -72.5665328579 14.5862737491 -72.3413642854 13.807794488 -72.3392662543 13.7962514963 -72.564408589	120.0	DK-1.54	LIGHT
ngc330field1_000002.png	beusers	2010-12-31 00:00:00	image/png	234.7kiB	14.19	-72.45	OSPS 2012-10-08T03:29:15.728 B ngc330field1	DFOSC_FASU	2012-10-08T03:29:15Z	2	[2148, 2048]	[0.000109996, 0.000109996]	ICRS	2000.0	TAN	[1074.0, 1024.0]	[14.193875, -72.4533333333333]	[-0.000109996, -9.82956e-07, -9.82956e-07, 0.000109996]	B	m	4.4e-07	5.6e-07	3.6e-07	N/A	Polygon ICRS 14.5844654144 -72.5665328579 14.5862737491 -72.3413642854 13.807794488 -72.3392662543 13.7962514963 -72.564408589	180.0	DK-1.54	LIGHT
ngc330field1_000005.png	beusers	2010-12-31 00:00:00	image/png	266.2kiB	14.19	-72.45	OSPS 2012-10-08T03:37:43.924 R ngc330field1	DFOSC_FASU	2012-10-08T03:37:42Z	2	[2148, 2048]	[0.000109996, 0.000109996]	ICRS	2000.0	TAN	[1074.0, 1024.0]	[14.193875, -72.4533333333333]	[-0.000109996, -9.82956e-07, -9.82956e-07, 0.000109996]	R	m	6.2e-07	8.5e-07	5.8e-07	N/A	Polygon ICRS 14.5844654144 -72.5665328579 14.5862737491 -72.3413642854 13.807794488 -72.3392662543 13.7962514963 -72.564408589	120.0	DK-1.54	LIGHT
ngc330field1_000004.png	beusers	2010-12-31 00:00:00	image/png	246.5kiB	14.19	-72.45	OSPS 2012-10-08T03:35:12.456 V ngc330field1	DFOSC_FASU	2012-10-08T03:35:12Z	2	[2148, 2048]	[0.000109996, 0.000109996]	ICRS	2000.0	TAN	[1074.0, 1024.0]	[14.193875, -72.4533333333333]	[-0.000109996, -9.82956e-07, -9.82956e-07, 0.000109996]	V	m	5.4e-07	6.7e-07	4.85e-07	N/A	Polygon ICRS 14.5844654144 -72.5665328579 14.5862737491 -72.3413642854 13.807794488 -72.3392662543 13.7962514963 -72.564408589	120.0	DK-1.54	LIGHT
 ngc330field1_000006.png	beusers	2010-12-31 00:00:00	image/png	260.8kiB	14.19	-72.45	OSPS 2012-10-08T03:40:09.216 R ngc330field1	DFOSC_FASU	2012-10-08T03:40:09Z	2	[2148, 2048]	[0.000109996, 0.000109996]	ICRS	2000.0	TAN	[1074.0, 1024.0]	[14.193875, -72.4533333333333]	[-0.000109996, -9.82956e-07, -9.82956e-07, 0.000109996]	R	m	6.2e-07	8.5e-07	5.8e-07	N/A	Polygon ICRS 14.5844654144 -72.5665328579 14.5862737491 -72.3413642854 13.807794488 -72.3392662543 13.7962514963 -72.564408589	120.0	DK-1.54	LIGHT
ngc330field1_000001.png	beusers	2010-12-31 00:00:00	image/png	235.3kiB	14.19	-72.45	OSPS 2012-10-08T03:25:50.480 B ngc330field1	DFOSC_FASU	2012-10-08T03:25:50Z	2	[2148, 2048]	[0.000109996, 0.000109996]	ICRS	2000.0	TAN	[1074.0, 1024.0]	[14.193875, -72.4533333333333]	[-0.000109996, -9.82956e-07, -9.82956e-07, 0.000109996]	B	m	4.4e-07	5.6e-07	3.6e-07	N/A	Polygon ICRS 14.5844654144 -72.5665328579 14.5862737491 -72.3413642854 13.807794488 -72.3392662543 13.7962514963	180.0	DK-1.54	LIGHT

OSPS SIAP in Aladin (DSS in back)



OSPS Image coverage (footprints)



Data Cubes = Scientific Data Structure

International Virtual Observatory Alliance

IVOA N-Dimensional Cube Model

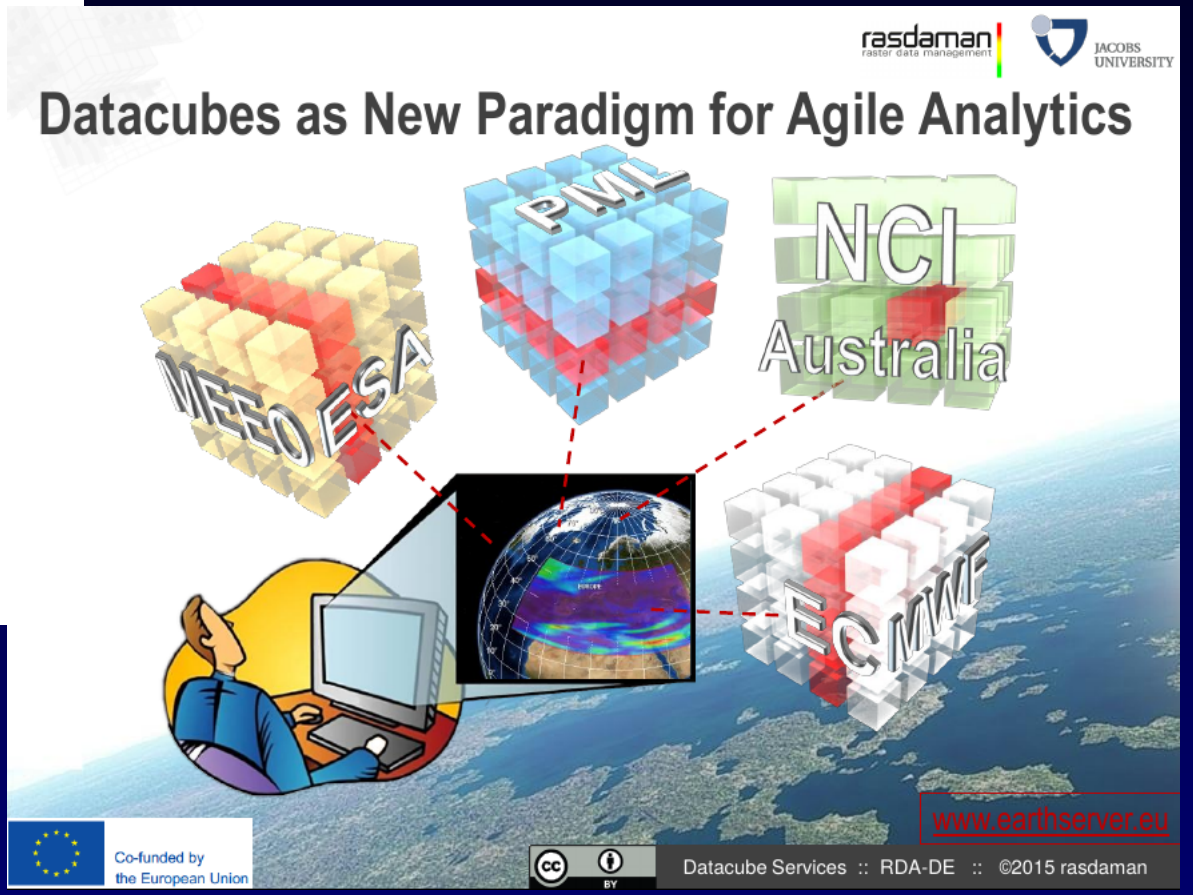
Version 1.0

IVOA Working Draft 20150320

This version:
WD-CubeDM-1.0-20140930
Previous version(s):

Editor(s):
Mark Cresitello-Dittmar

Authors:
Doug Tody, Francois Bonnarel, Omar Laurino, Mireille Louys, Arnold Rots, Jose Enrique Ruiz, Jesus Salgado, and the IVOA Data Model Working Group.



OSPS Light curves – plot (customized)

International Virtual Observatory Alliance

Time Series Cube Data Model

Version 1.1

IVOA Note 2017-02-05

Working group
Time domain interest group

This version
http://www.ivoa.net/documents/cubeDM/20170205

Latest version
http://www.ivoa.net/documents/cubeDM

Previous versions

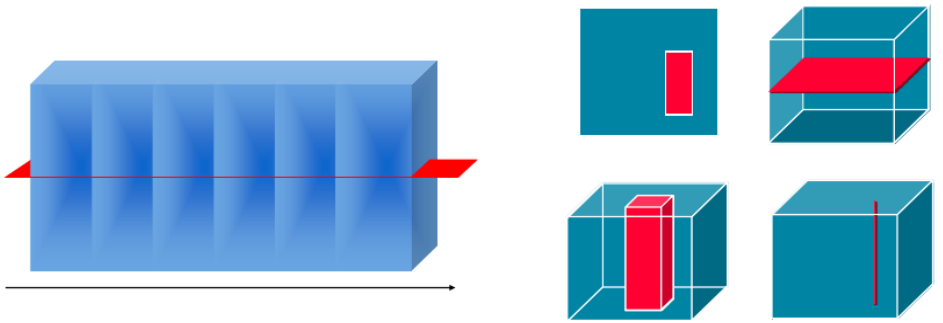
Author(s)
Jiří Nádvorník, Petr Škoda, Dave Morris, Pavel Tvrdík

Editor(s)
Jiří Nádvorník

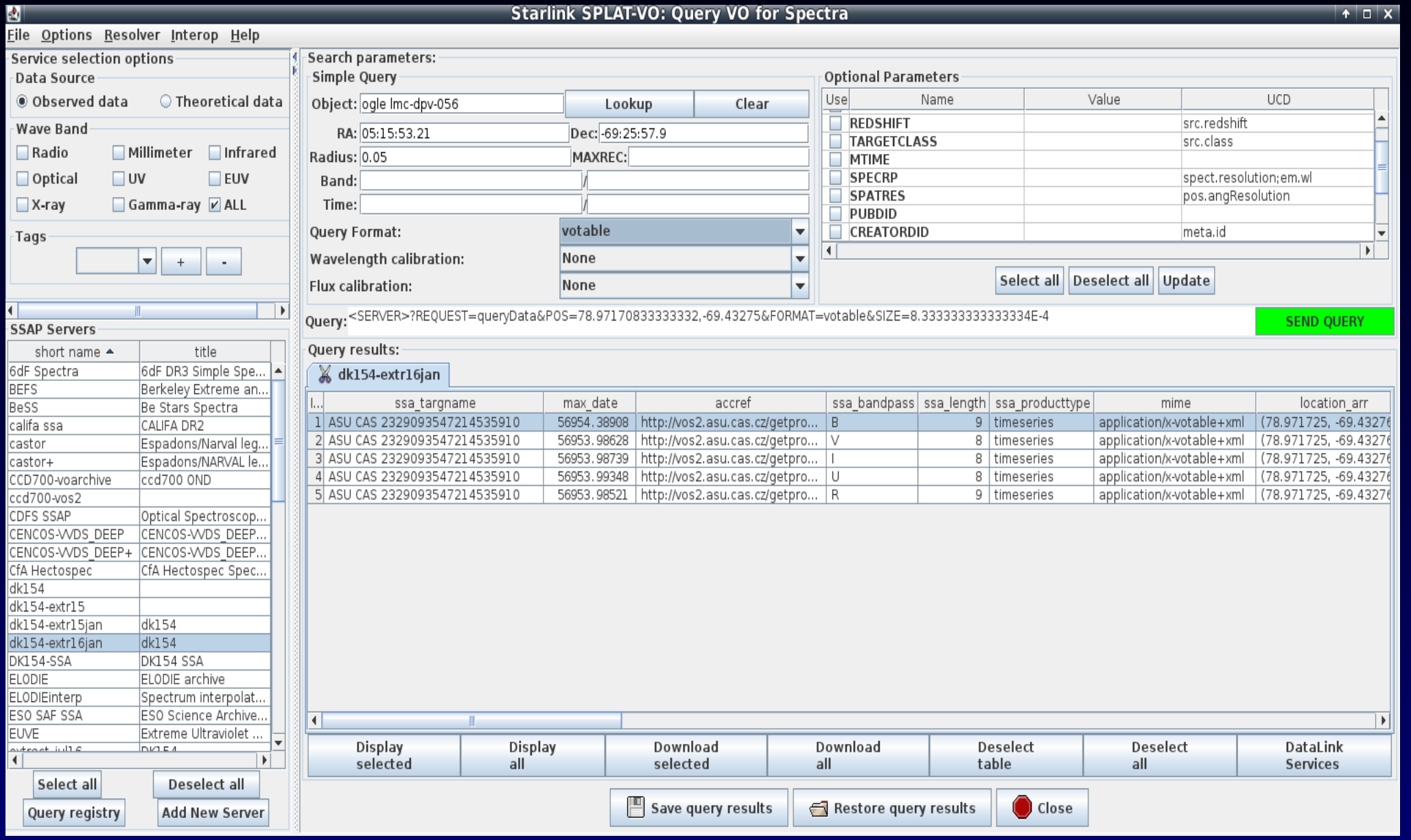
rasdaman raster data management

JACOBS UNIVERSITY

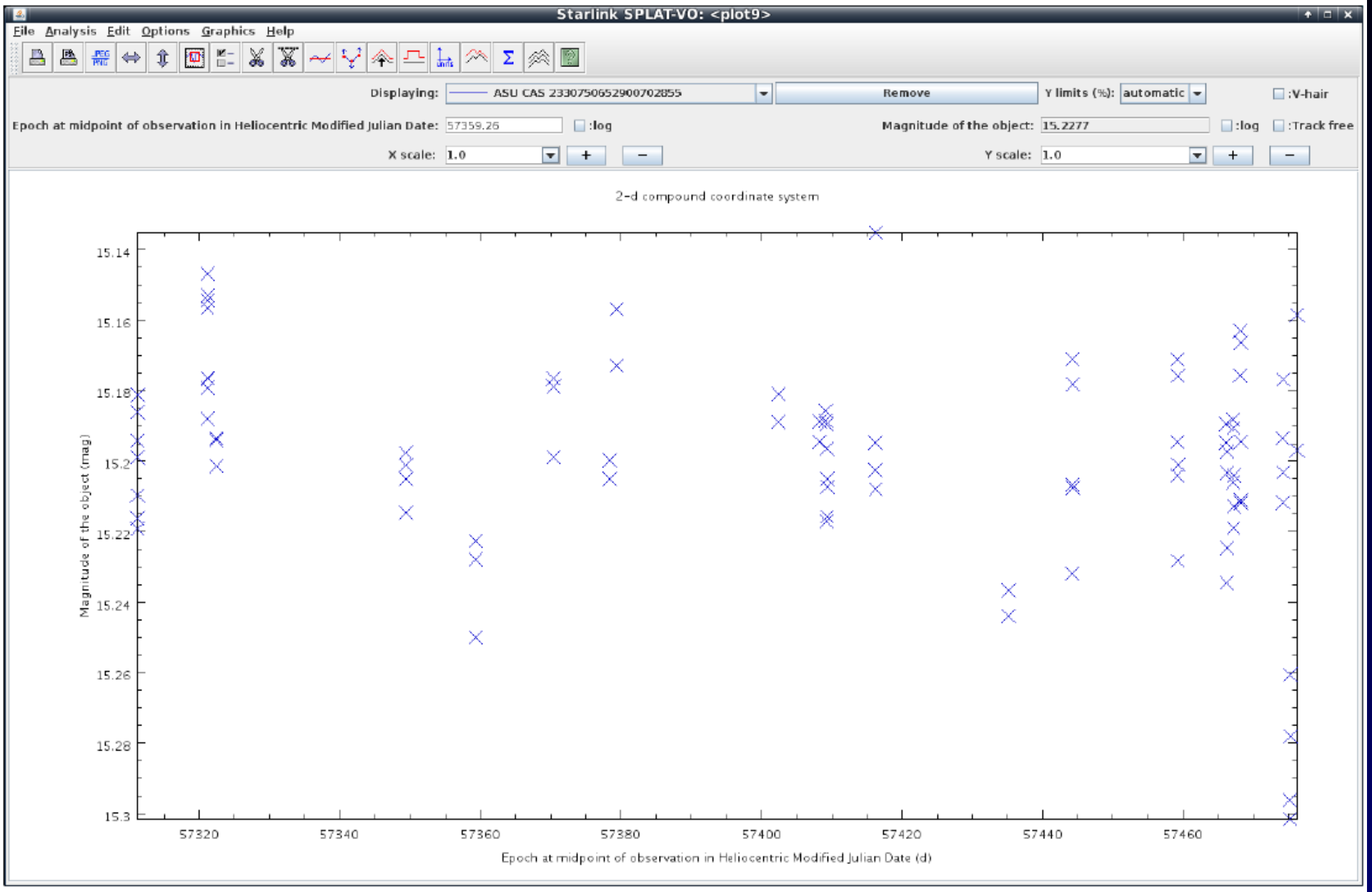
Datacube Services



OSPS Light curves in SPLAT-VO using Sparse Cube DM



OSPS Light curves – plot (customized)



Conclusions

- Machine learning on Big Data archives may identify new interesting objects yet unknown
- Global Data Federation from Multiple fields
- Crucial is interactive visualization

- Future science will be multidisciplinary
- Wide collaboration of experts and informaticians
- Education of new expert – Data Scientist
- Transfer of technology – commercial interest
- X-Informatics ← Philosophia (love of wisdom)

- VO-like technology helps in every step