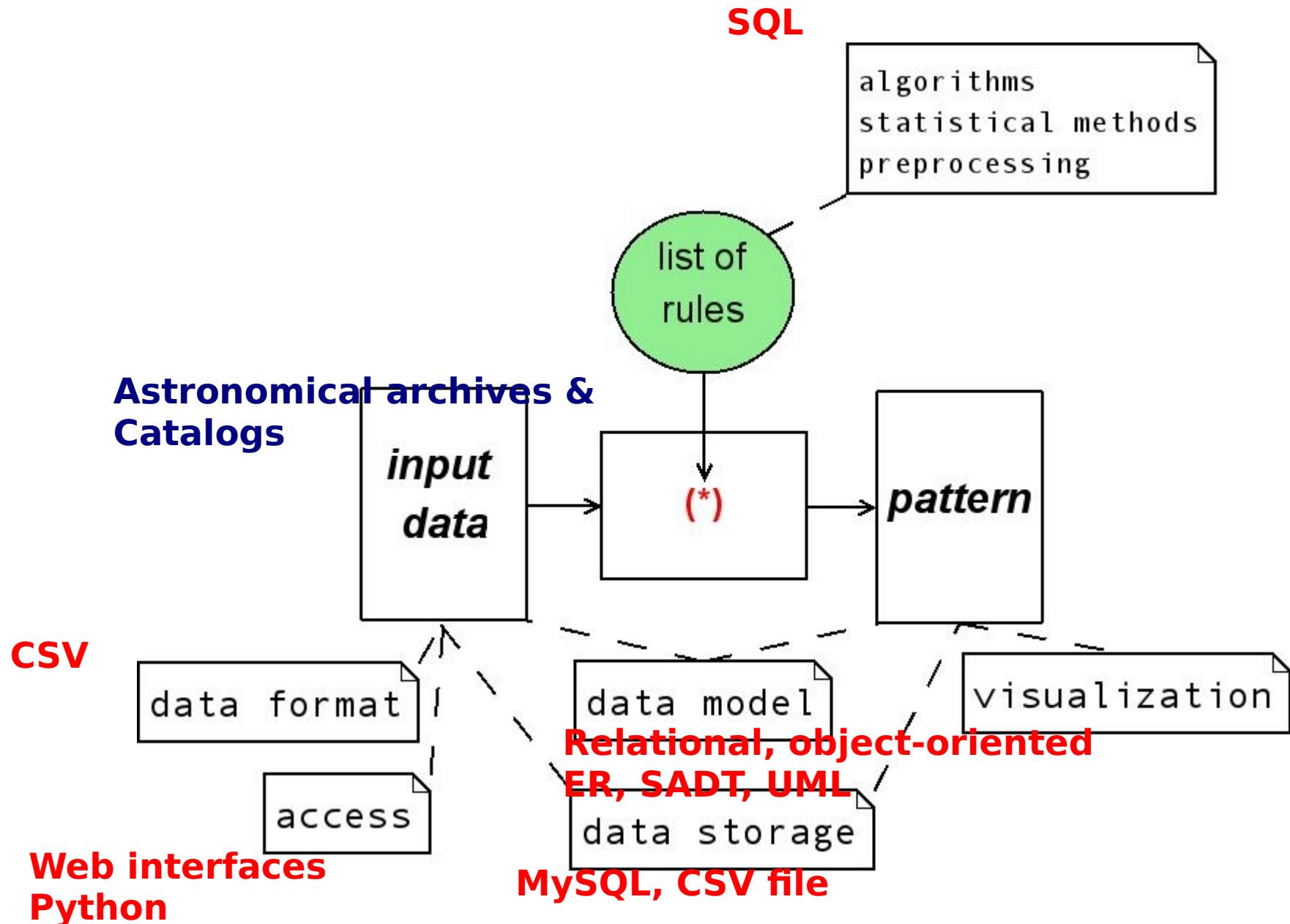


Virtual Observations 2016 Data Mining in Astronomy

Databases in Astronomy

Data Mining: lectures so far



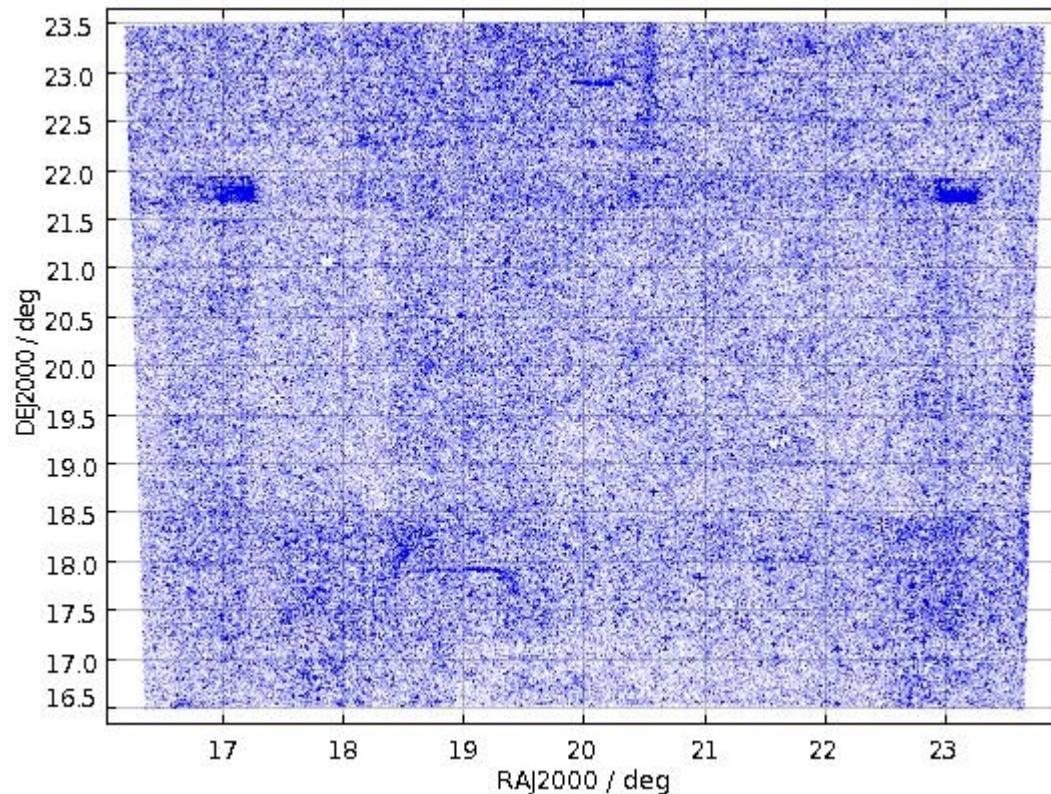
- Databases in Astronomy
- In-depth SDSS
- UKIDSS, CFHT
- Theoretical data: Millenium database
- SQL again
- Science with databases

- Catalogs, archives, surveys
- Bibliographical
- Product of optical-, infra-red-, radio-surveys
- Astrophysical parameters
- Publications

- Fundamental catalogs: astrometric, photometric & astrophysical
- Surveys
- Libraries

- Deliverables
- Selection criteria
- Spectral coverage and resolution
- Spacial coverage and resolution
- Quality
- Completeness
- Effects of selection

Effects of selection



Markarian Galaxies: First Buyrakan Survey

- List of galaxies with coordinates, magnitudes, classification, S(eyfert) or Q(uasar) flag; cross identifications to other catalogs; B1950 positions; major and minor axis (arcsec); magnitude and spectral types, including a code for the strength of the UV continuum.
- UV excess moderate to strong ultraviolet continuum as detected by an objective-prism survey

Depth: no galaxies brighter than mag 13.0 or fainter than mag 17.5 (which Band?)

- Single broad-band magnitude
- (Observatory location)
- (not applicable)
- (Observatory location)



- POSS → DSS
- 2MASS
- SDSS
- UKIRT surveys → UKIDSS
- ESO public surveys
- Space-based surveys (GAIA, Euclid)



- Bonner Durchmusterung (1859 -1968)
- Henry Draper
- Bright Star Catalog
- GCVS
- USNO -A1,2; -B1
- PPMX
- Hipparcos, Tycho, ASCC

Catalogs: BD family



BD: +90 - -2 (Bonner Durchmusterung),
324198 stars, $9^{\text{m}.5}$

SBD: -1 - -23 (Suedlicher Bonner Durchmusterung),
133659 stars, 10^{m} .

CoD or CD: -22 - -90 (Cordoba Durchmusterung),
613953 stars, 10^{m} .

CPD: -18 - -90 (Cape Photographic Durchmusterung),
454875 stars, $9^{\text{m}.2}$

Catalogs: surveys & catalogs

Carte du Ciel	1920s	refractors	-90 - +90	15(B)	AC, AC2000, AC200.2
Palomar -NGS Survey, POSS I	1958	Palomar 1.2m	-33 - +90	21(B) 20(R)	USNO-A2 APM POSS I USNO-B
Palomar Quick -V	1984	Palomar 1.2m	+3 - +90	19(V)	GSC, TIC
SERC Southern Sky Survey, SERC-J	1987	UK 1.2m	-90 - -17	23(B)	GSC, TIC, USNO-A, USNO-B
ESO Southern Sky Survey, ESO-R	1990	ESO 1.0m	-90 - -17	22(R)	USNO-A, USNO-B
AAO Survey	1998	UK 1.2m	-90 - -20	21.5(R)	USNO-B
Second Palomar Sky Survey	1999	Palomar 1.2m	-3 - +90	22(B) 21(R) 20(I)	USNO-B

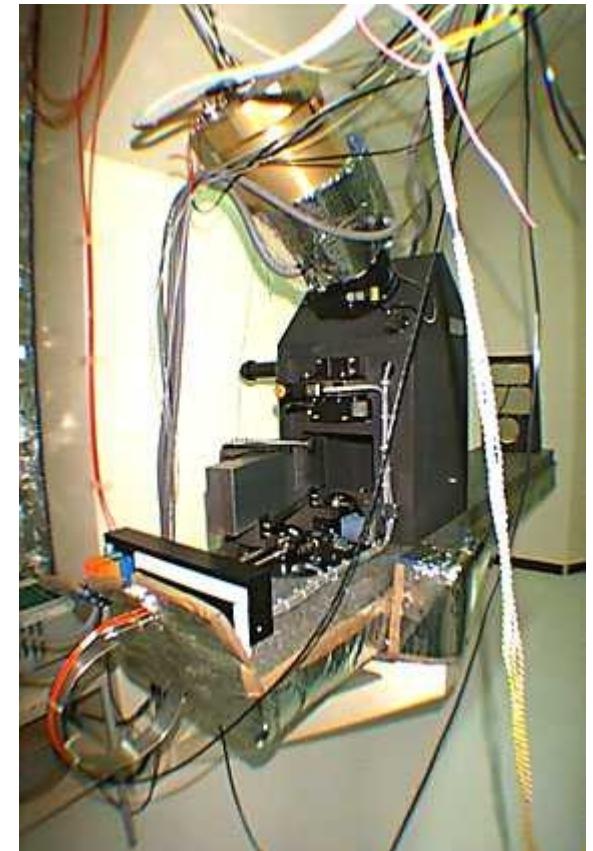
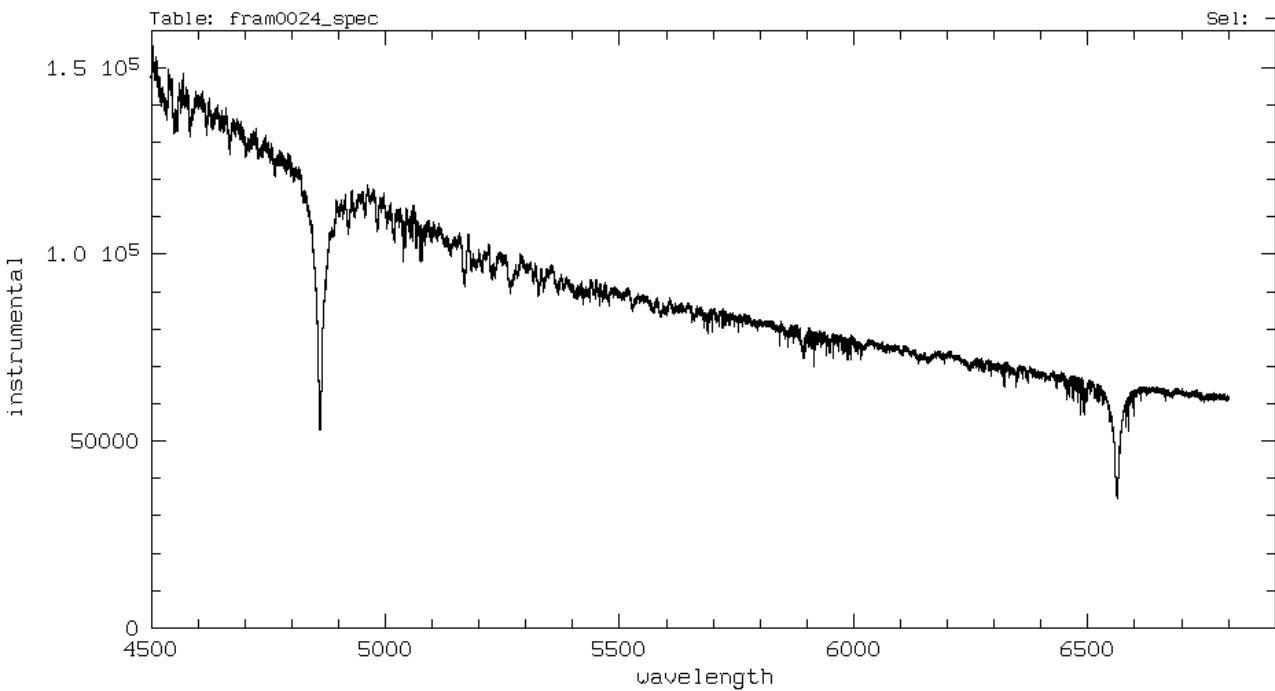
- New General Catalog
- Index Catalog
- Catalogs of interacting galaxies (Arp)
- Reference catalog of Bright Galaxies (RC3)
- ESO/Uppsala survey

- Survey-related
- Reference systems (FK4, FK5, Hipparcos, Tycho, FK6)
- Photometric standards

Libraries of stellar spectra



- Pickles spectral flux library 1998
- ELODIE
- MILES



- Iben (1966)
- Stahler & Palla (1993)
- D'Antona & Mazzitelli (1994)
- Geneve group
- Padova group <http://stev.oapd.inaf.it/>
- BASTI

Libraries of stellar atmospheres



- ATLAS9 – Kurucz <http://kurucz.harvard.edu/grids.html>
- NEXT
- MARCS

- 1999 – with the use of Hipparcos
- 878 “original” stars + 3278 stars
- Positions
- Proper Motions
- Radial velocity
- Parallaxes
- V magnitude



- New version – van Leeuwen, 2007
- 117995 stars
- Coordinates
- Proper Motions
- Radial velocity
- Parallaxes
- Hipparcos magnitude
- Additional photometry

- 2.5 mln stars
- Coordinates
- Proper Motions
- Radial velocity
- Parallaxes
- B,V magnitudes - Tycho system

Combine the data



Centaurus A – Peculiar Galaxy

Distance: 11,000,000 ly light-years (3.4 Mpc)

Image Size = 15 x 14 arcmin

Visual Magnitude = 7.0



X-Ray: Chandra



Ultraviolet: GALEX



Visible: DSS



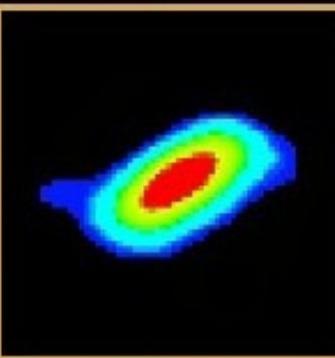
Visible: Color ©AAO



Near-Infrared: 2MASS



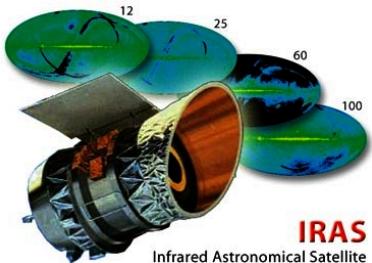
Mid-Infrared: Spitzer



Far-Infrared: IRAS



Radio: VLA



- Data centers
- Observatory archives
- Mission Data Centers
- Data Centers with specialization
- VO

- Simbad ~ 7 mln. objects
- Vizier ~ 11,000 catalogs
- Aladin images
- Catalog service

Data centers: ESO archive



ESO Archive Query Form - Mozilla Firefox

File Edit View History Bookmarks Tools Help

ESO Archive Query F... +

archive.eso.org/eso/eso_archive_main.html

star.gif esodown.gif esodata.gif esorequest.gif

If you would like to query the Archive for instrument specific parameters, please use the [dedicated query forms](#).
To search for **reduced Data Products**, please have a look at the [ESO Data Products](#) page and the [Advanced Data Products](#) query form.

The checkboxes on the right of the parameters define whether or not they will be displayed on the query result page.

SEARCH ShowAll ShowNone Reset query Help Status of Requests

ESO Archive Query Form

ESO Archive Overview **Help Page** **FAQ** **Archive Policy HOME** **ESO HOME**

Target, Program and Scheduling Information

Target Name Resolved by SIMBAD
RA **DEC** **Search Box** **Input** RA(h) DEC(deg) **Output** Sexagesimal (h, deg)
List of Targets

Night (DD MM YYYY)
OR give a query range using the following start/end dates:
Start **12 hrs [UT]** **End** **12 hrs [UT]**
Program ID **Program Type** Any
PI Col **SV** Any
Title

Observing Information

Imaging **Spectroscopy** **Interferometry** **Other**
ALL **NONE** **ALL** **NONE** **ALL** **NONE** **ALL** **NONE**
 VLT/FORS1 VLT/CRIRES VLT/VINCI APEX/HET
 VLT/FORS2 VLT/FORS1 VLT/MIDI APEX/BOL
 VLT/HAWKI VLT/FORS2 VLT/AMBER UKIRT/WFCAM
 VLT/ISAAC VLT/GIRAFFE VLT/ISAC LSOF
 VLT/NACO VLT/SINFONI VLT/VIMOS MASCOT
 VLT/VIKING VLT/VISIR VLT/VISIR
 VISTA/VIRCAM VLT/XSHOOTER VLT/XSHOOTER
 VST/OMEGACAM NTT/EMMI VLT/VIKING
 NTT/EMMI NTT/SOFI VLT/VIKING
 NTT/SOFT NTT/SUSI2 VLT/VIKING
 3.6/EFOSC2 NTT/SOFT 3.6/EFOSC2
 3.6/TIMMI2 3.6/CE5 3.6/EFOSC2
 2.2/WFI 3.6/HARPS 3.6/TIMMI2
 3.6/FEROS 3.6/EFOSC2 3.6/EFOSC2
Polarimetry **Coronagraphy**
ALL **NONE** **ALL** **NONE** **ALL** **NONE** **ALL** **NONE**
 VLT/FORS1 VLT/FORS2 VLT/ISAAC NTT/SOFT
 VLT/FORS2 VLT/VIMOS VLT/NACO 3.6/EFOSC2
 VLT/VISIR VLT/XSHOOTER VLT/VIKING 3.6/EFOSC2
 VLT/XSHOOTER NTT/EMMI VLT/VIKING 3.6/EFOSC2
 NTT/EMMI NTT/SOFT VLT/VIKING 3.6/EFOSC2
 NTT/SOFT 3.6/CE5 VLT/VIKING 3.6/EFOSC2
 3.6/EFOSC2 3.6/HARPS VLT/VIKING 3.6/EFOSC2
 3.6/TIMMI2 3.6/EFOSC2 VLT/VIKING 3.6/EFOSC2
 2.2/WFI 3.6/FEROS VLT/VIKING 3.6/EFOSC2

Instrument & Mode

Data Product Info

Type Any
User defined input:
Mode Any
User defined input:
Dataset ID
Orig Name
Release Date
OB Name
OB ID
Category
 SCIENCE
 CALIB
 ACQUISITION

Instrumental Setup

Exptime
Filter
Grism
Grating
Slit

SEARCH ShowAll ShowNone Reset query Help Status of Requests

Additional Tools and Parameters

Airmass FITS Header Display Distance Number of Frames / Instrument

Data centers: NED at IPAC



NASA/IPAC Extragalactic Database - NED - Mozilla Firefox

File Edit View History Bookmarks Tools Help

NASA/IPAC Extragala... WISE: Wide-Field I... ned.ipac.caltech.edu eso archive

NASA/IPAC EXTRAGALACTIC DATABASE

News & Featured Updates (June - August, 2012)

- 8.1 million multiwavelength cross-IDs
- 6.2 million objects within the SMC
- 43,000 objects and redshifts from the 2MASS Redshift Survey
- 11,000 spectra from the UZC Galaxy Redshift Survey
- About 500 7-year WMAP and 500 5-year WMAP radio sources
- Review article *Extragalactic Line Surveys* in LEVEL 5

OBJECTS **DATA** **LITERATURE** **TOOLS** **INFO**

By Name	Images By Object Name or By Region	References by Object Name	Coordinate Transformation & Extinction Calculator	Introduction Latest News/Updates
Near Name	Photometry & SEDs	References by Author Name	Velocity Calculator	Features FAQ
Near Position	Spectra	Text Search	Cosmology Calculators	Overview (pdf)
IAU Format	Redshifts	Knowledgebase <small>LEVEL 5</small>	Extinction-Law Calculators	Source List
By Parameters (All-Sky)	Redshift-Independent Distances	Galaxy Distance Tabulations (NED-D)	Skyplot	Web Links
By Classifications Types, Attributes	Classifications by Object Name	Abstracts	X/Y offset to RA/DEC	Glossary & Lexicon
By Refcode	Positions	Thesis Abstracts	Batch Job Submission Pick Up Results	Team
Object Notes	Diameters		<small>NEW</small> Build Data Table from Input List By Name Near Name/Position (Cross-Matching)	Contact Us or Comment

If your research benefits from the use of NED, we would appreciate the following acknowledgement in your paper: *This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.*

NASA **JPL** **California Institute of Technology** **ipac**

Data centers: Bibliographical database ADS



SAO/NASA ADS: ADS Home Page - Mozilla Firefox

File Edit View History Bookmarks Tools Help

ads SAO/NASA ADS: A... x WISE: Wide-Field I... x +

adsabs.harvard.edu

Sign on

ADS Services

- [Search](#)
- [Browse](#)
- [myADS](#)
- [Mirrors](#)
- [Feedback](#)
- [FAQ](#)
- [what's new](#)
- [Site Map](#)
- [Help](#)

Other NASA Centers

- [CXC](#)
- [HEASARC](#)
- [IRSA](#)
- [MAST](#)
- [NED](#)
- [NSSDC](#)
- [PDS](#)
- [SPITZER](#)

Related Sites

- [AAS](#)
- [ADEC](#)
- [arXiv](#)
- [CDS](#)
- [IAU](#)

The SAO/NASA Astrophysics Data System

Welcome to the Digital Library for Physics and Astronomy

This site is hosted by the  [High Energy Astrophysics Division](#) at the  [Harvard-Smithsonian Center for Astrophysics](#)

The SAO/NASA Astrophysics Data System (ADS) is a Digital Library portal for researchers in Astronomy and Physics, operated by the Smithsonian Astrophysical Observatory (SAO) under a NASA grant. The ADS maintains three bibliographic databases containing more than 9.7 million records: Astronomy and Astrophysics, Physics, and arXiv e-prints. The main body of data in the ADS consists of bibliographic records, which are searchable through [highly customizable query forms](#), and full-text scans of much of the astronomical literature which can be [browsed or searched](#) via our full-text search interface. Integrated in its databases, the ADS provides access and pointers to a wealth of external resources, including electronic articles, data catalogs and archives. We currently have links to over 9.4 million records maintained by our collaborators.

Please note that all abstracts and articles in the ADS are copyrighted by the publisher, and their use is free for personal use only. For more information, please read our page detailing the [Terms and Conditions](#) regulating the use of our resources.

In addition to its databases, the ADS provides the [myADS Update Service](#), a free custom notification service promoting current awareness of the recent technical literature in astronomy and physics based on each individual subscriber's queries. Every week the myADS Update Service will scan the literature added to the ADS since the last update, and will create custom lists of recent papers for each subscriber, formatted to allow quick reading and access. Subscribers are notified by e-mail in html format. As an option, users can elect to receive updates on preprints published on the [arXiv](#) e-print archive via daily emails or by subscribing to a custom RSS feed.

The importance of ADS's role in supporting the scientific community [has been recognized by societies and individuals](#). If you wish to acknowledge us in a publication, kindly use a phrase such as the following:
``This research has made use of NASA's Astrophysics Data System.''

Thanks!

The ADS personnel are:

Theoretical data: Millenium



A very large N-body simulation of the Universe (Springel et al 2005).

To make predictions for galaxy evolution semi-analytic models can be run “on-top”. The Millennium database has two such models - from Bower et al (2006) and de Lucia et al (2006)/de Lucia & Blaizot (2007)

Contents:

Stellar masses, star formation rates, metal content, star formation histories, assembly histories, dark matter halos, predictions on observables such as absolute magnitudes.

<http://gavo.mpa-garching.mpg.de/Millennium/>

```
SELECT *
  FROM millimil..DeLucia2006a
 WHERE snapnum=63
   AND mag_b BETWEEN -26 AND -18
   AND stellarMass > 10
```

Some issues:

This table contains only 1/512 of the full Millennium simulation

- you can get access to the full one by registering.

There is a limit on the max number of rows and execution time

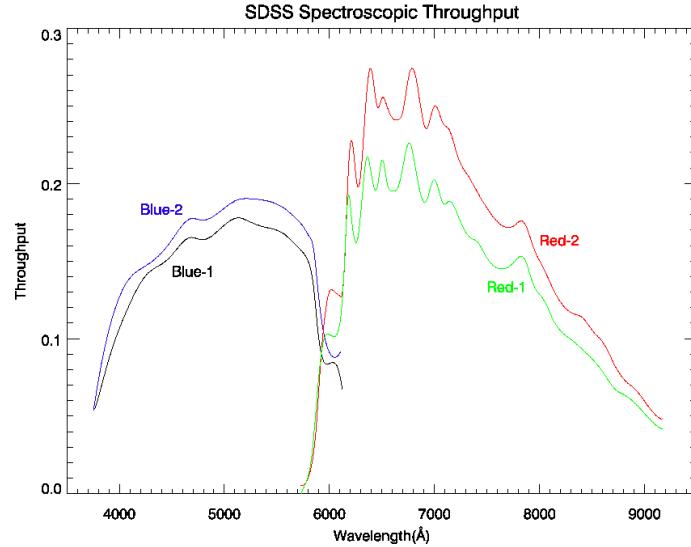
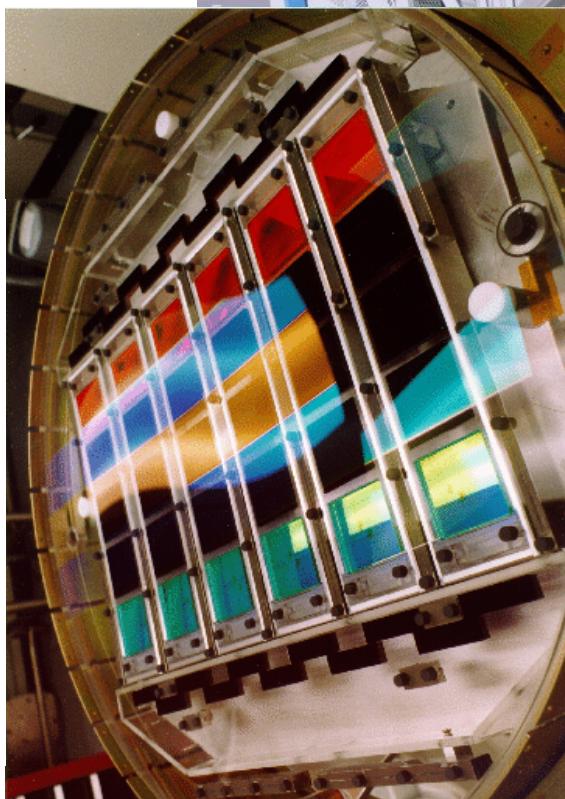
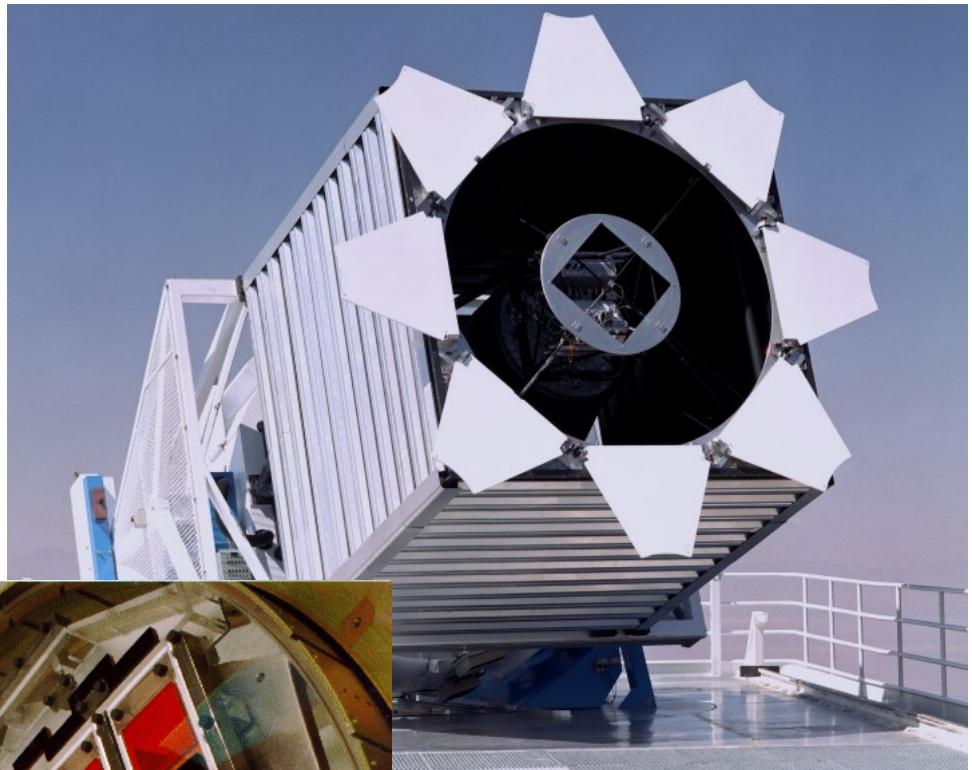
Download is a bit awkward - but there are examples on how to download straight to R/IDL and you can access it directly from Topcat which might be the best option!

SDSS: Initial goals

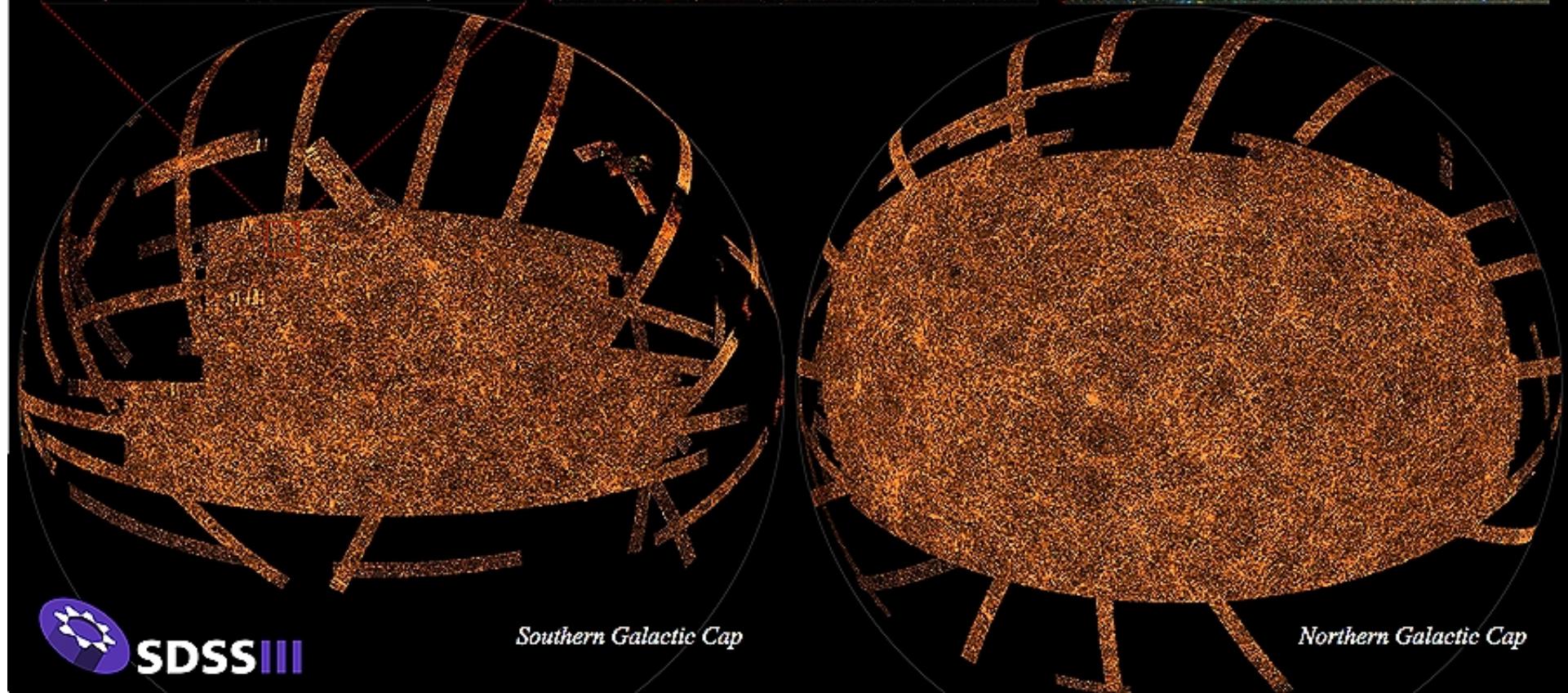
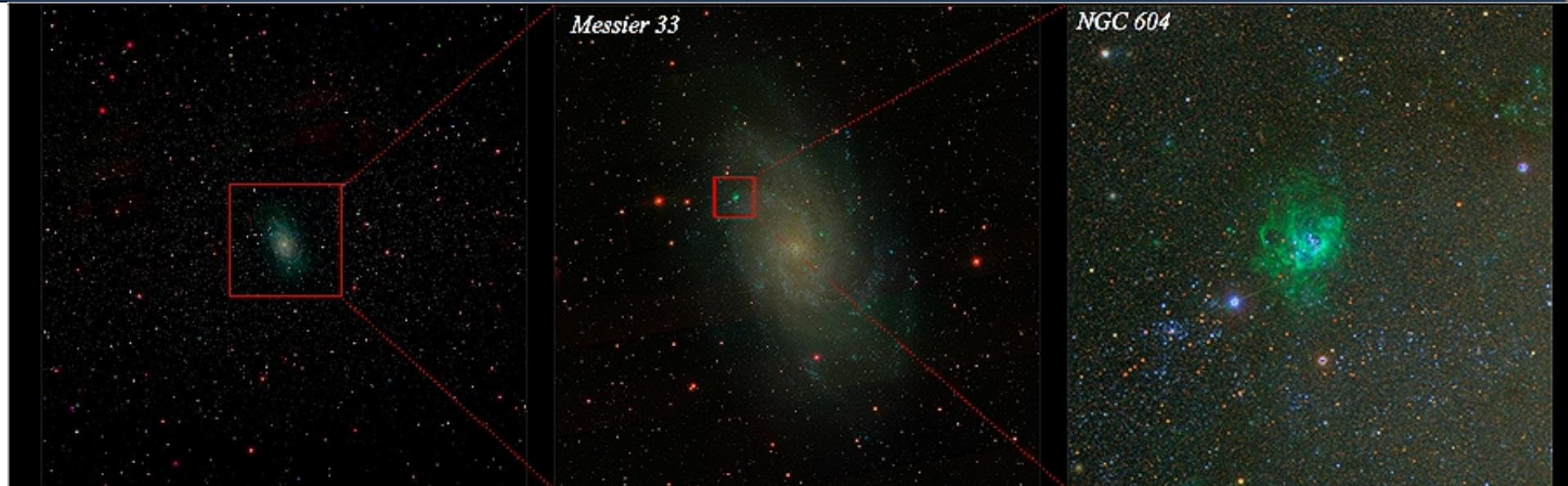


- Broad-band combined photometric & spectroscopic survey for extragalactic studies
- Extended to cover galactic plane (SEGUE)
- 2.5m + 0.5m PT
- 5-band photometry
- 3600-10000 Å
- Started 2000
- Periodical data releases

SDSS: Instruments



SDSS: Science



SDSSIII

Southern Galactic Cap

Northern Galactic Cap



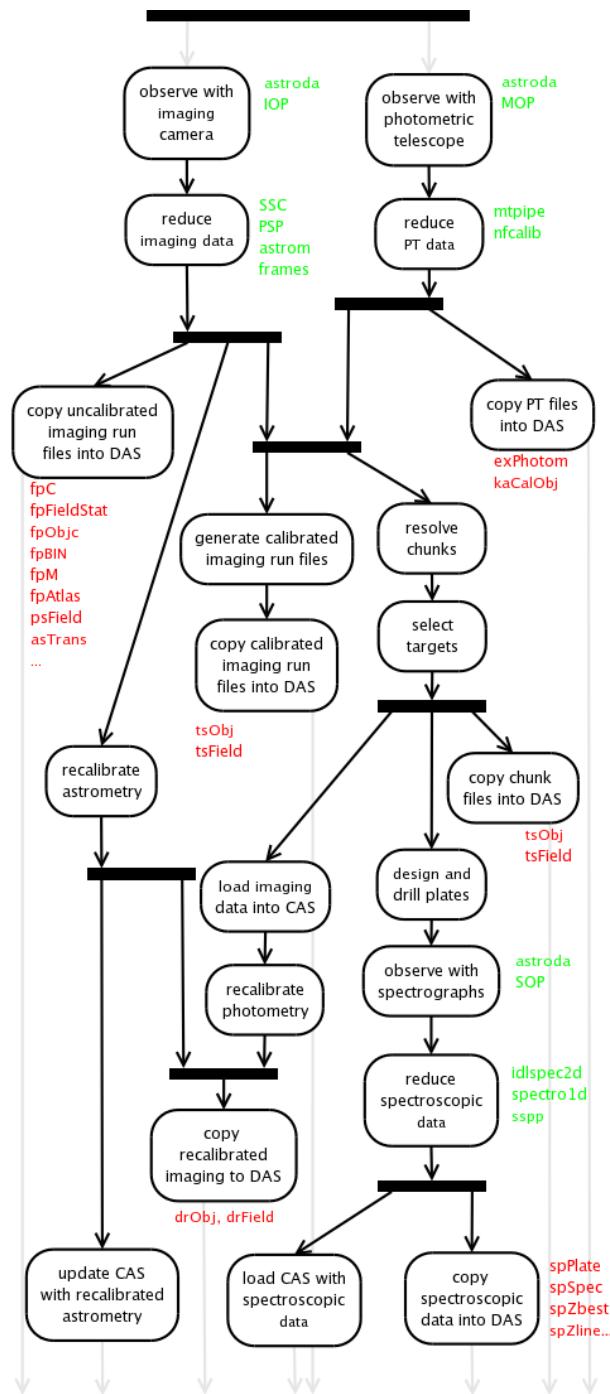
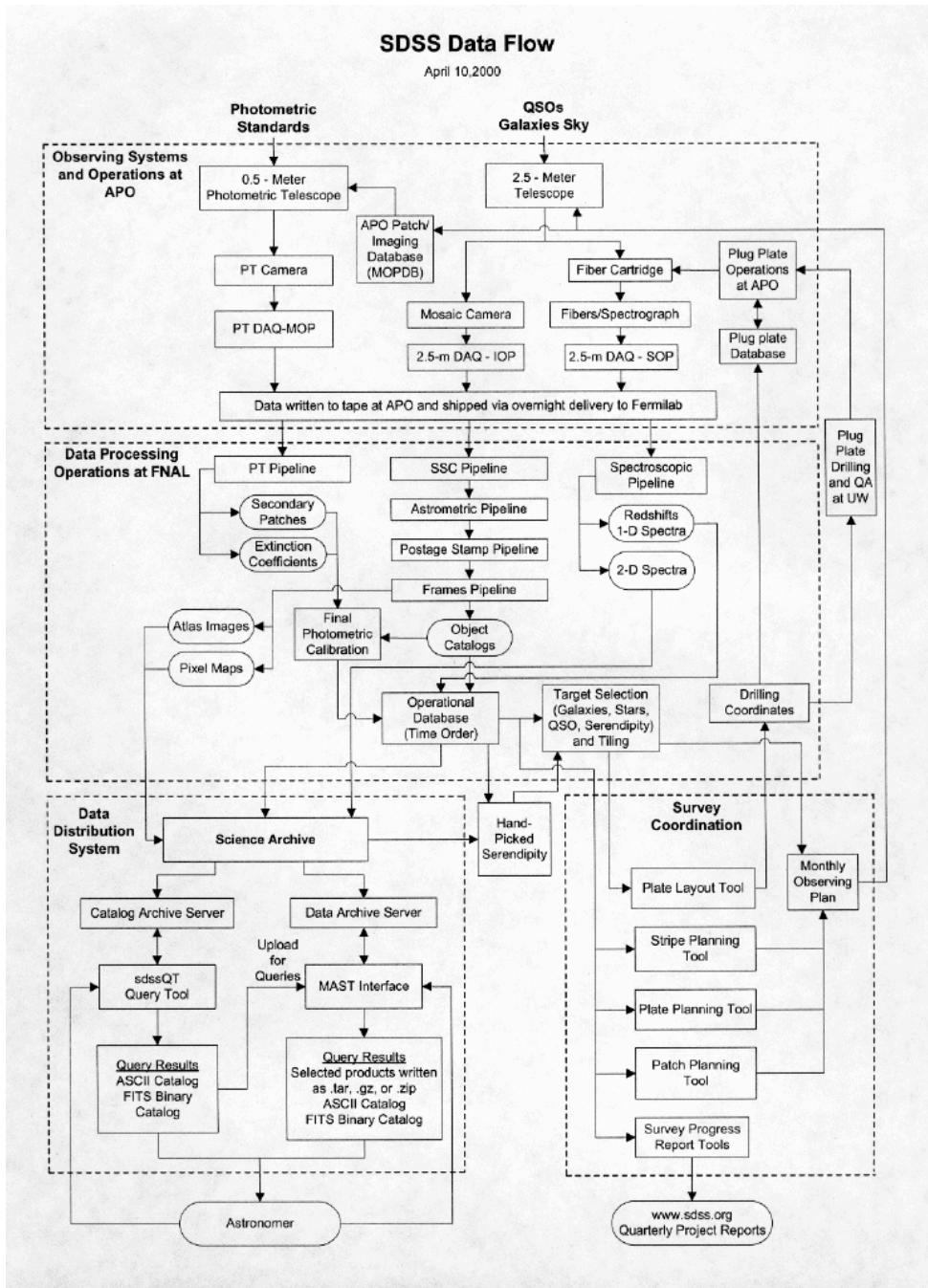
- 14000 sq. deg
- Galaxies and quasars survey
- SN survey
- Halo survey
- BD search
- DR9

SDSS: Data Processing



- Images, spectra → Fermilab
- Data analysis → Instrument calibration
- Data analysis → reduced images, spectra, catalog

SDSS: Data Processing



SDSS: Data Processing



cas.sdss.org/astrodr7/en/help/browser/browser.asp

Sloan Digital Sky Survey / SkyServer

Home Tools SQL Search Schema Finding Chart Download Projects DR7 DAS Site Search Help

Schema Browser

TABLE PsObjAll

Pan-STARRS photometry for selected SDSS objects.

name	type	length	unit	ucd	description
objID	bigint	8		ID_MAIN	Unique SDSS identifier composed from [skyVersion=1,run,run,camcol,field,obj].
run	smallint	2		OBS_RUN	Run number
rerun	smallint	2		CODE_MISC	Rerun number
camcol	tinyint	1		INST_ID	Camera column
field	smallint	2		ID_FIELD	Field number
obj	smallint	2		ID_NUMBER	The object id within a field. Usually changes between reruns of the same field.
catId	int	4			Pan-STARRS id matching this run-rerun-camcol-field-obj
rowc	real	4	pix	POS_CCD_Y	Row center position ('r' coordinates)
rowcErr	real	4	pix	POS_CCD_Y_ERROR	
colc	real	4	pix	POS_CCD_X	
colcErr	real	4	pix	POS_CCD_X_ERROR	
ra	float	8	deg	POS_EQ_RA_MAI	
dec	float	8	deg	POS_EQ_DEC_MAI	
offsetRa_u	real	4	arcsec	POS_OFFSET	
offsetRa_g	real	4	arcsec	POS_OFFSET	
offsetRa_r	real	4	arcsec	POS_OFFSET	
offsetRa_i	real	4	arcsec	POS_OFFSET	
offsetRa_z	real	4	arcsec	POS_OFFSET	
offsetDec_u	real	4	arcsec	POS_OFFSET	
offsetDec_g	real	4	arcsec	POS_OFFSET	

SDSS: Observing mode



- Images are taken in a drift-scan using a mosaic camera with five filters. Each scan is called a strip, but to cover gaps a second scan is made and the combination is called a stripe. The exposure time is ~54s with ~73s between each filter
- The images are then analysed using a code called Photo and based on the measurements objects are selected for spectroscopy
- Spectroscopy is carried out the following season using a fibre spectrograph with 3" fibres covering 3800Å-10000Å
- Spectroscopic data are analysed using two pipelines and imaging & spectroscopy is released to the public

SDSS: Results - DR7



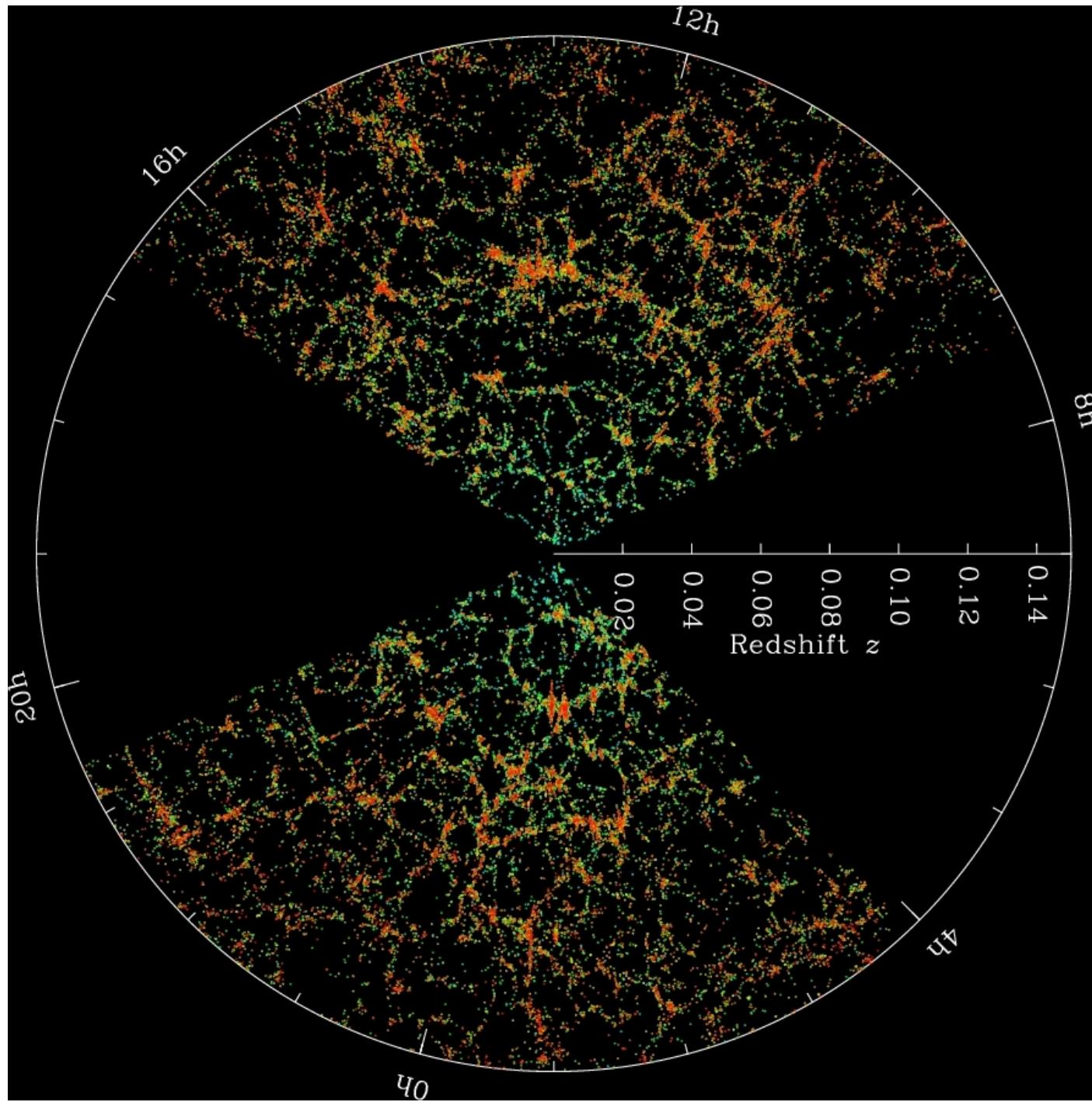
- 5 band (ugriz) imaging of 11,663 sq. deg deg
- 357 mln objects
- R~1800 spectroscopy covering 3800-9200Å for 1,640,960 objects [929,555 galaxies, 121,373 QSOs, 380,214 stars including 84,047 M and later]
- Stars with spectroscopy: log Teff, log g, [Fe/H]
- Main galaxy sample: $r < 17.77$ - also a QSO sample, Luminous Red Galaxies sample + stars & special objects
- Median seeing for images: 1.43" (50%: [1.3", 1.58"])

SDSS: Results - DR9



- 5 band (ugriz) imaging of 14,555 sq. deg deg
- 932,891,133 objects
- Spectra: galaxies (1,457,002), quasars (222,468), stars (668,054)
- MS SQL Server
- Stars with spectroscopy: log Teff, log g, [Fe/H], [α /Fe]

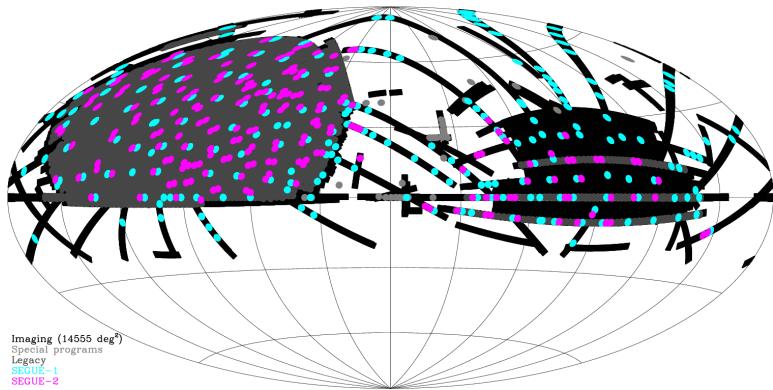
SDSS: Results - DR9



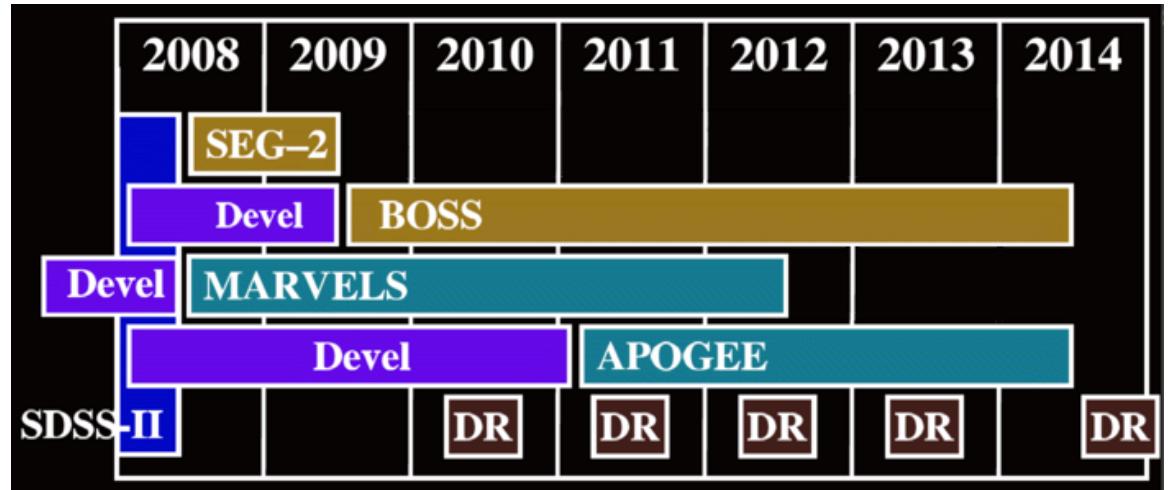
SDSS: Results - DR12



- 5 band (ugriz) imaging of 14,555 sq. deg deg
- 932,891,133 objects
- Spectra: galaxies (2,401,952), quasars (477,161), stars (851,968)
- MS SQL Server
- Stars with spectroscopy: log Teff, log g, [Fe/H], [α/Fe]
- Radial velocities



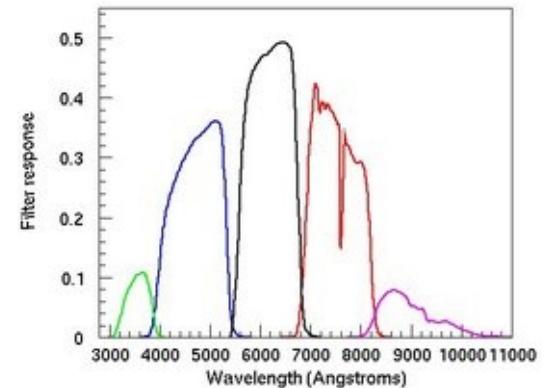
Imaging (14555 deg²)
Special programs
Legacy
SEGUE-1
SEGUE-2



SDSS: Photometric system



- 5 band (ugriz) – broad band
- Atmosphere 1.3
- AB system
- PT + SDSS main telescope



Pogson

$$\text{mag} = -2.5 * \log_{10}(f/f_0)$$

$$\text{error(mag)} = 2.5 / \ln(10) * \text{error(counts)} / \text{counts}$$

asinh

$$\text{mag} = -(2.5/\ln(10)) * [\text{asinh}((f/f_0)/2b) + \ln(b)]$$

$$\text{error(mag)} = 2.5 / \ln(10) * \text{error(counts)}/\text{exptime} * 1/2b * 100.4 * (\text{aa} + \text{kk} * \text{airmass}) / \sqrt{1 + [(f/f_0)/2b]^2}$$

SDSS: Magnitudes



Relationship to other photometric calibrations

The SDSS is a near AB magnitude system.

$$m_{AB} = 2.5 \times (23 - \log_{10} F [\text{Jy}]) - 48.6$$

Offsets:

$$u_{AB} = u_{SDSS} - 0.04$$

$$z_{AB} = z_{SDSS} + 0.02$$

<http://www.sdss.org/dr7/algorithms/sdssUBVRITransform.html>

Example:

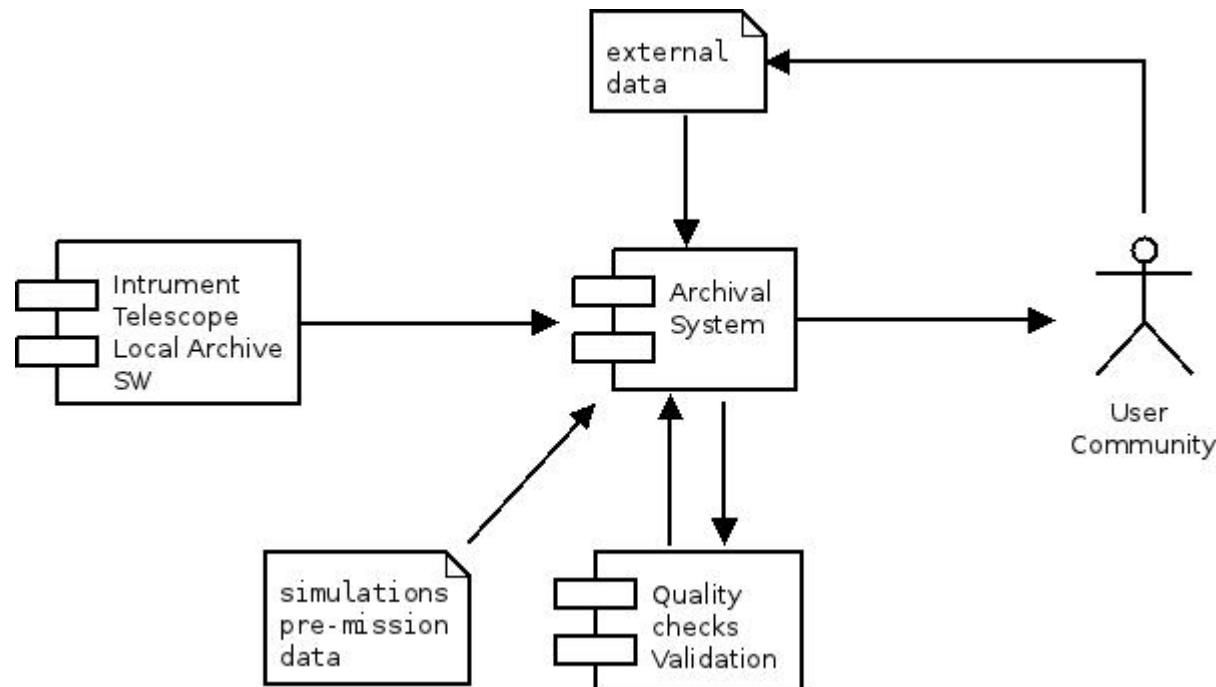
$$B = g + 0.3130*(g - r) + 0.2271$$

$$V = g - 0.5784*(g - r) - 0.0038$$

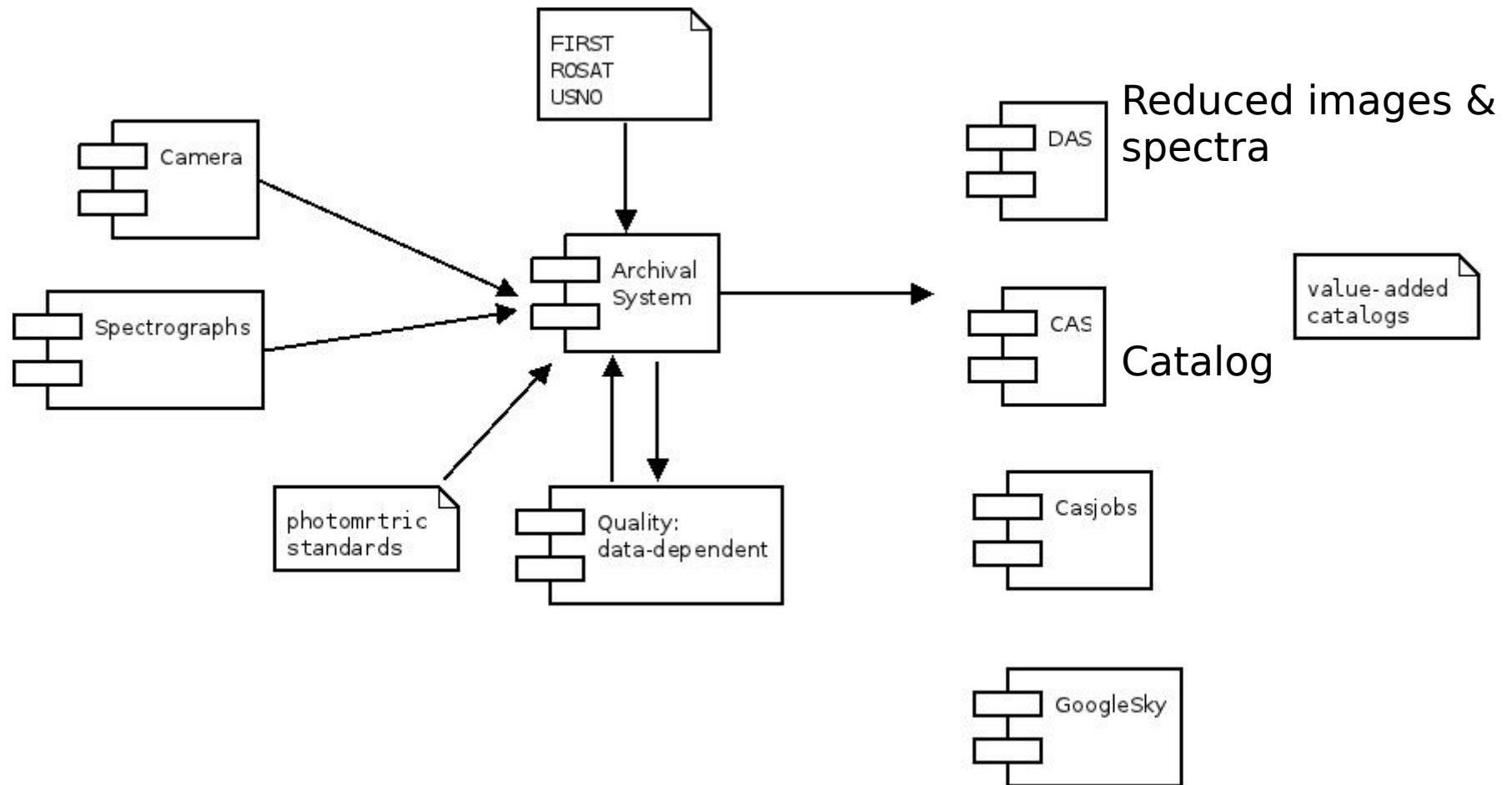
$$R = r - 0.2936*(r - i) - 0.1439$$

$$I = i - 0.3780*(i - z) - 0.3974$$

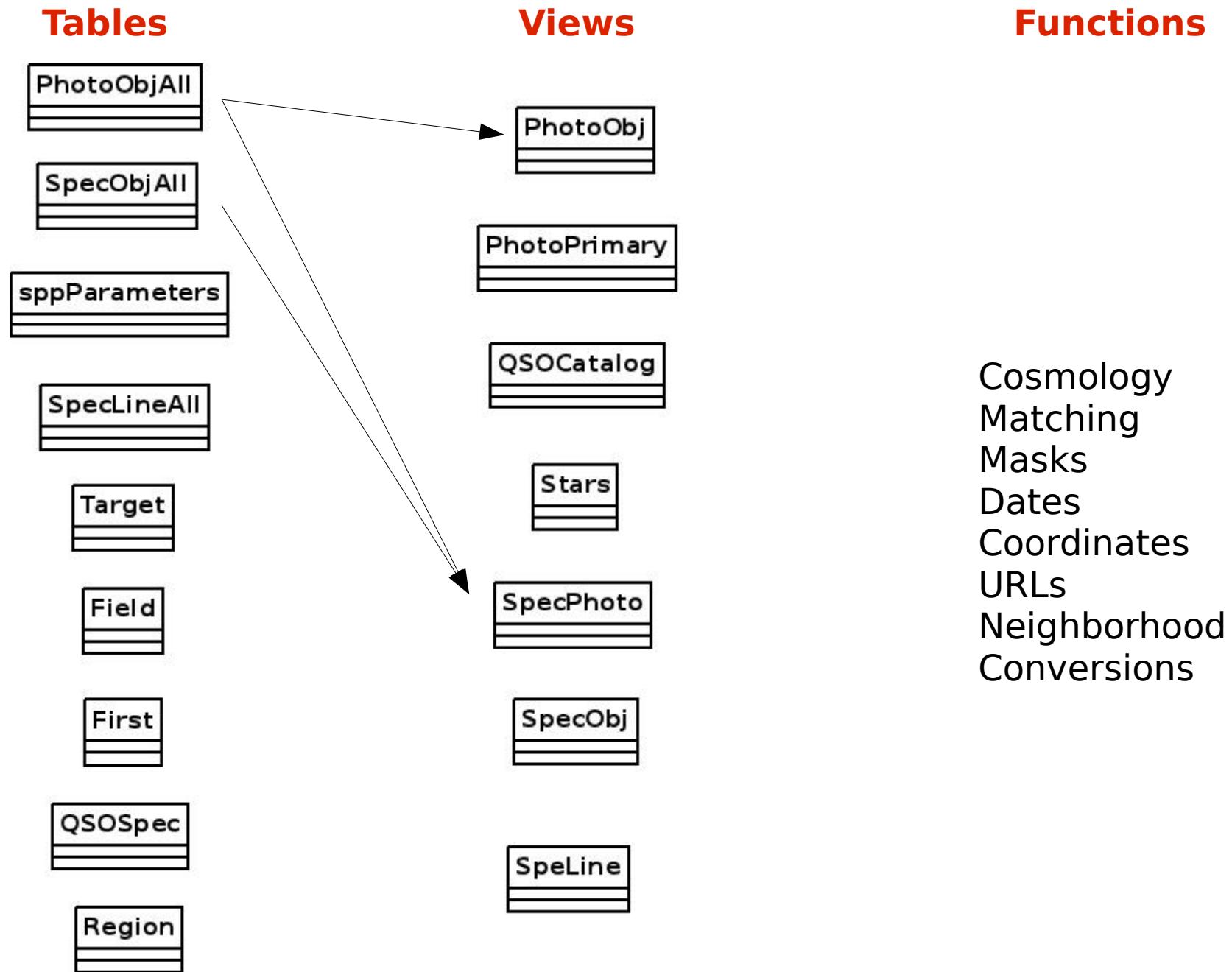
SDSS: Data Flows



SDSS: Data Flows



SDSS: Database Structure



SDSS: Identification



- ObjID – photometric detection
- TargetObjID – specific objID for target selection
- SpecObjID – spectroscopic identification
- BestObjID - ID for the current best data reduction
- Indexes

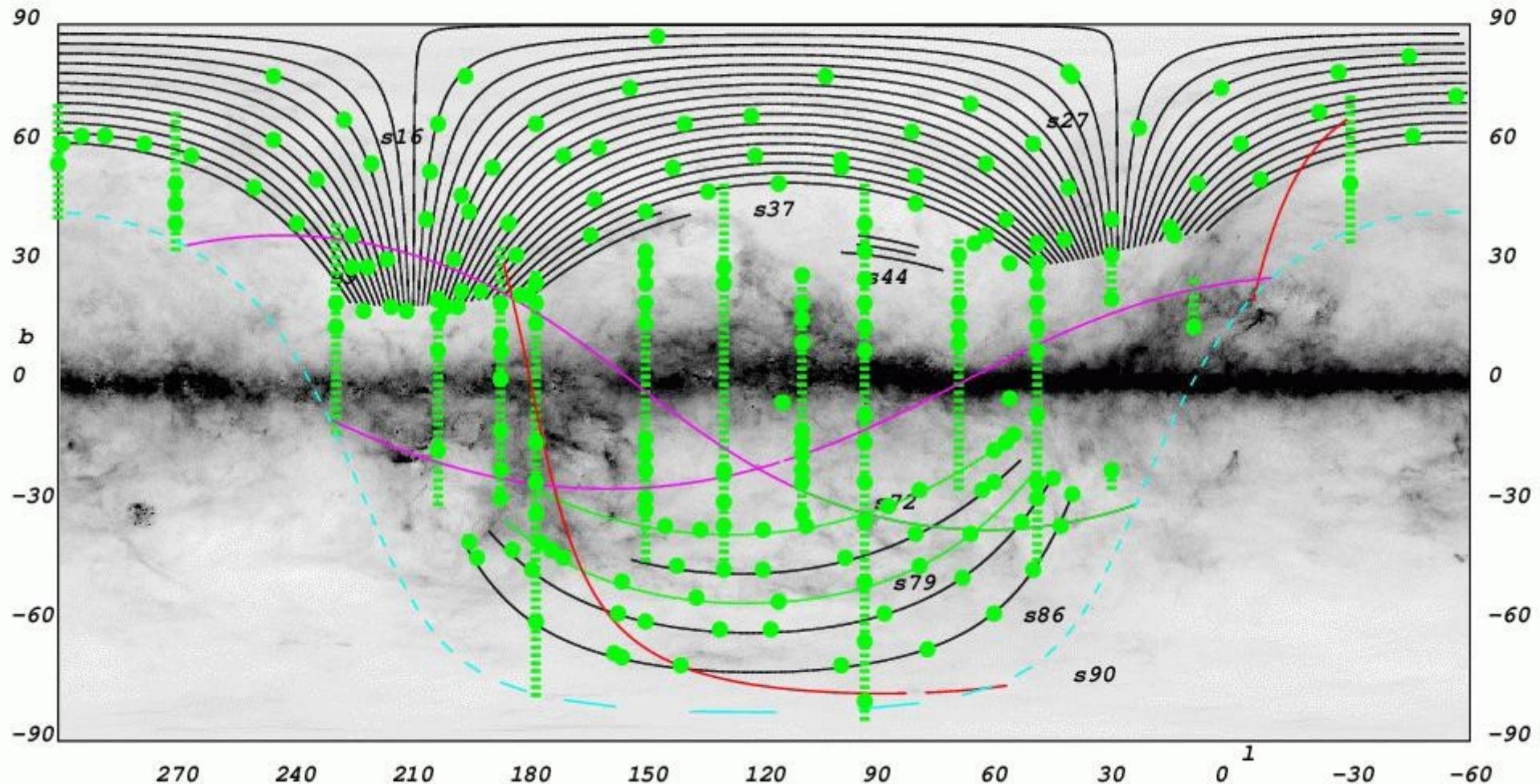
SDSS: PhotoObjAll



- Magnitudes (Petrosian, Model, Aperture, PSF)
- Sizes (radii & moments)
- Positions
- Shapes, orientations
- Images (links)
- Image quality

http://www.sdss.org/dr12/algorithms/flags_detail/

SEGUE imaging sky coverage and plate layout July 15, 2008 (Complete)



SDSS: added-value & classification data



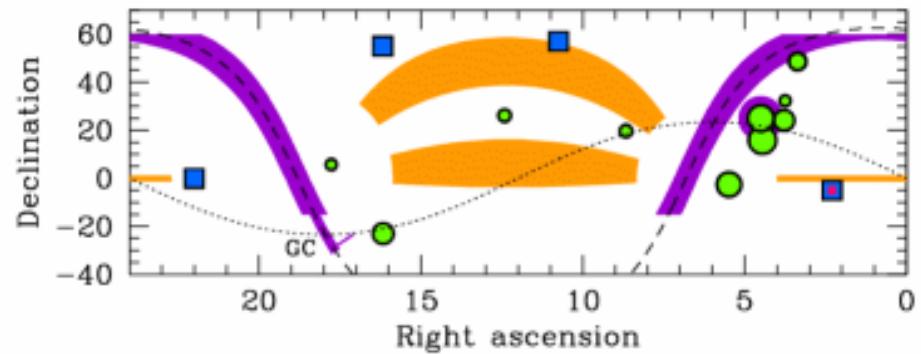
- QSO catalogues (Schneider et al) & BAL-QSOs.
- White Dwarf catalogues (Eisenstein et al)
- Variable star catalogues from Stripe 82 & CV catalogues
- Moving objects catalogues



UKIRT surveys

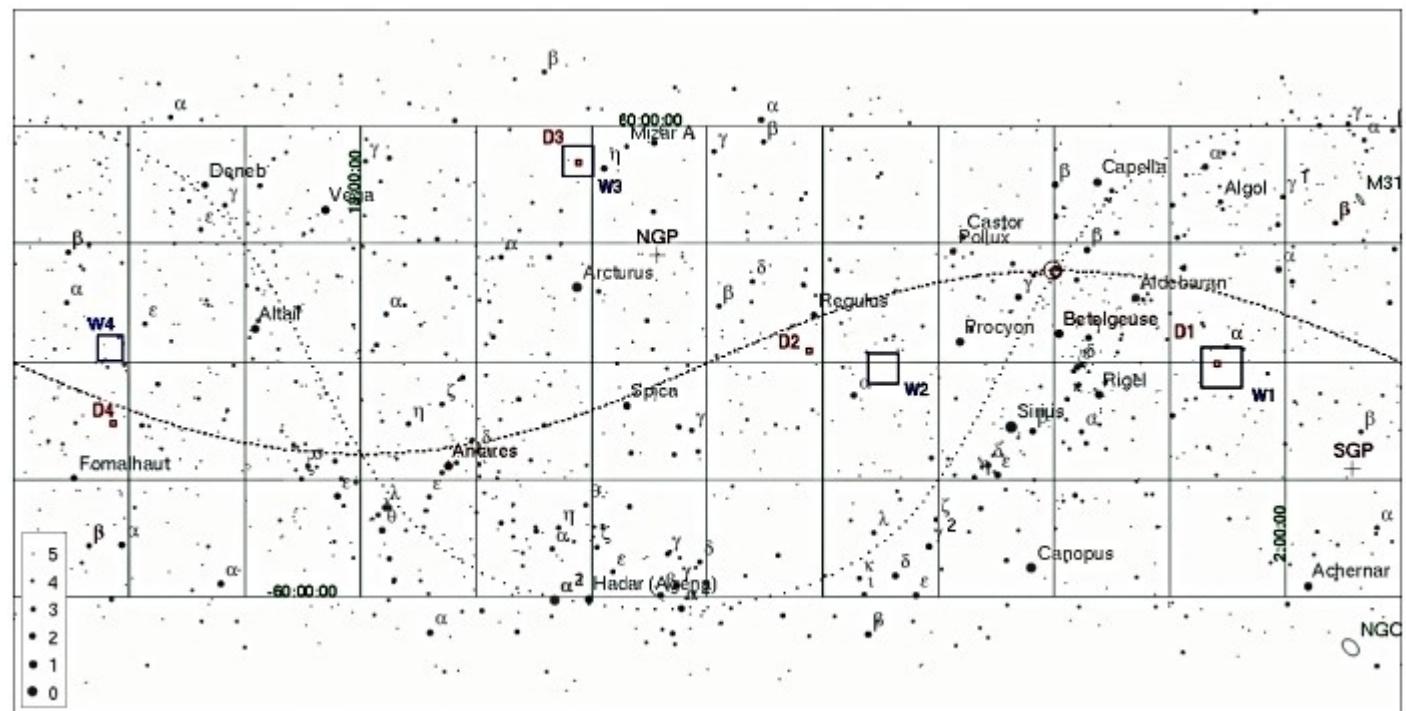


- UK Infra-red telescope
- WFCAM
- Large Area Survey, Galactic plane survey, Galactic clusters survey, deep extragalactic survey, ultra deep survey
- ZYJHK



<http://surveys.roe.ac.uk/wsa/index.html>

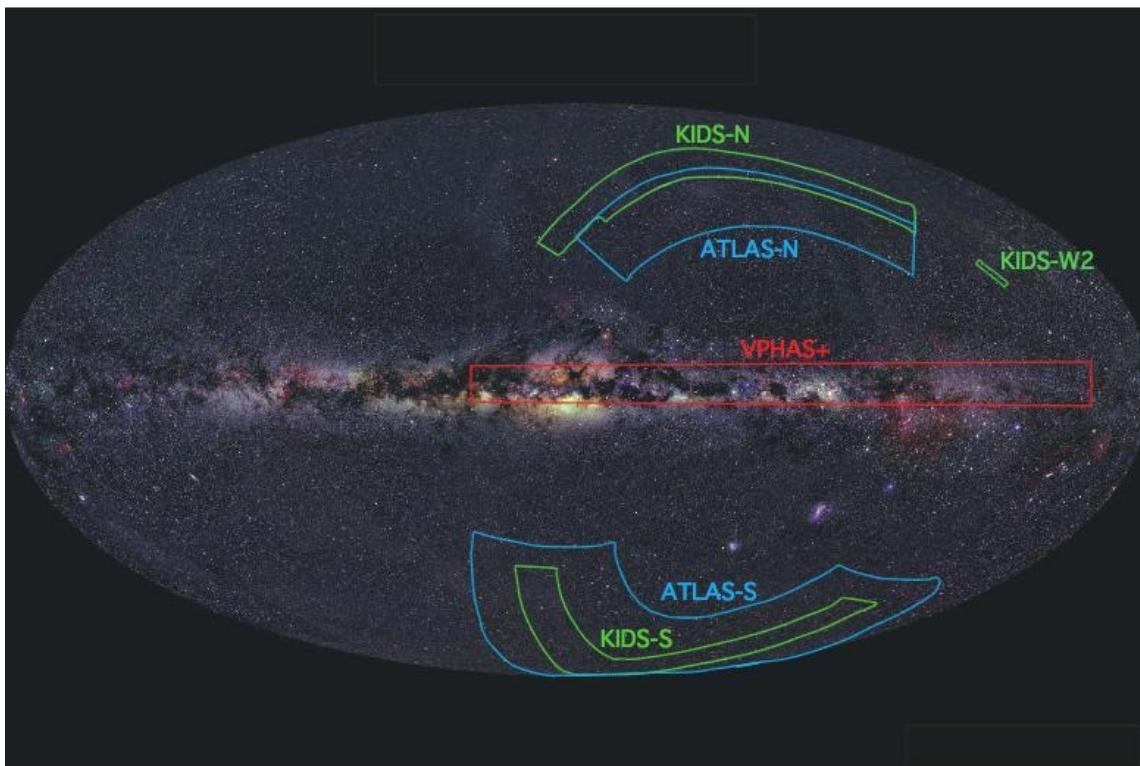
- Canada-Frence-Hawaii Telescope
- Mauna Kea observatory
- MegaCam: 1 sq. deg
- Extragalactic surveys
- ugriz



ESO Public Surveys (1)



- KIDS (1500 sq. deg., 25 g)
- ATLAS (4500 sq. deg, 22 g)
- VPHAS+ (1800 sq. deg, 22g)
- VST
- ugri



ESO Public Surveys (2)



- VIKING (1500 sq. deg., 21 K)
- UltraVISTA, VHS, VVV, VMC, VIDEO
- VISTA
- ZYJHK

