



VSA

The VISTA Science Archive

Eckhard Sutorius
Wide Field Astronomy Unit, IfA, Edinburgh



R. Blake, R. Collins, N. Cross, N. Hambly, M. Read

Overview

- * Data Transfer and Ingest
- * Recalibration
- * Post-nightly pipeline data products:
 - stacks, mosaics, difference images
- * Provenance
- * Quality Control & QC Flags
- * Source Merging & Reseaming
- * Neighbour Table Creation
 - Release Database



Data Transfer & Ingest

- * Data transfer from  to  via UKLight
 - * Compressed Images (JPEGs)
 - * Extraction, process & ingest of multiframe metadata
 - Flat File Access (images, confidence maps, single band catalogues)!
 - * Extraction, process & ingest of catalogue data
-
- Parallelised processing, run-time mainly determined by ingest.

21% / 12%

- [WSA Home](#)
- [Start Here](#)
- [Data Overview](#)
- [Known Issues](#)
- [the Surveys](#)
- [Schema browser](#)
- [Data access](#)
- [Login](#)
- [Archive Listing](#)
- [GetImage](#)
- [MultiGetImage](#)
- [Region](#)
- [Menu query](#)
- [Freeform SQL](#)
- [CrossID](#)
- [Analysis services](#)
- [SQL Cookbook](#)
- [Q&A](#)
- [Glossary](#)
- [Release History](#)
- [non-Survey](#)
- [Gallery](#)
- [Publications](#)
- [Monitor](#)**
- [Downtime](#)
- [Links](#)



Monitor Page

WFCAM Science Archive Curation Statistics - Microsoft Internet Explorer

generated: 2007-12-07 07:22:20.12 UT

Legend

- This column shows whether the data are ready for inspection, ie. data are transferred, JPEGs exist, and image metadata are ingested.
- No tickmark means no metadata are ingested yet so flat file access is not available for this date.
- No JPEGs have been generated yet, but image metadata are ingested so flat file access is available for this date.
- Not all JPEGs have been generated/ingested, but image metadata are ingested so flat file access is available for this date.
- All JPEGs are calculated and ingested, image metadata are ingested.
- No WFCAM data taken.
- CU1** Number of files transferred from CASU. Subdivided by types: science frames, catalogue files, all files (incl. calibration frames).
- CU2** Number of calculated JPEGs. Subdivided by number of JPEGs and number of FITS files. Normally each FITS file has 4 JPEGs associated, one per extension.
- CU3** Number of FITS files that have image metadata ingested. Subdivided by types: pixel data files, catalogue files, all files.
- CU4** Number of catalogue data objects that have been ingested. Subdivided by survey: DXS, GCS, GPS, LAS, UDS, CALNS (CAL & Non-Surveys), all catalogue data objects.

April 2007 (2007A)

Date	🔍	CU1 (transfer from CASU) (version/#sci/#cat/#all)			CU2 (JPEGs calculated) (version/#jpgs/#files)			CU3 (Image metadata ingested) (version/#pix/#cat/#all)			CU4 (Detections ingested) (version/#dxs/#gcs/#gps/#las/#uds/#cal#ns/#all)									
		1	2	3	1	2	3	1	2	3	1	0	0	0	0	0	0	0	0	0
1	🟢	1	682	70	791	1	2884	721.0	1	721	70	791	1	0	0	0	69038	0	909274	978312
2	🟢	1	1602	152	1811	1	6636	1659.0	1	1659	152	1811	1	0	0	8608905	336194	0	243745	9188844
3	🟢	1	2176	216	2462	1	8984	2246.0	1	2246	216	2462	1	0	0	195234	443394	0	247537	886185
4	🟢	1	1774	123	1944	1	7284	1821.0	1	1821	123	1944	1	0	0	177826	56610	0	974835	1209271
5	🟢	1	3660	377	4348	1	15884	3971.0	1	3971	377	4348	1	0	0	12974794	637577	0	636778	14249149
6	🟢	1	3360	56	3448	1	13568	3392.0	1	3392	56	3448	1	0	0	0	0	0	1038105	1038105
7	🟢	1	3619	66	3722	1	14624	3656.0	1	3656	66	3722	1	0	0	2938971	0	0	936066	3875037
8	🟢	1	3450	54	3534	1	13920	3480.0	1	3480	54	3534	1	0	0	0	0	0	778109	778109
9	🟢	1	3450	54	3534	1	13920	3480.0	1	3480	54	3534	1	0	0	0	0	0	963339	963339
10	🟢	1	2918	266	3256	1	11960	2990.0	1	2990	266	3256	1	0	0	470381	303221	0	771798	1545400
11	🟢	1	4096	367	4572	1	16820	4205.0	1	4205	367	4572	1	17236	0	14400287	432594	0	819173	15669290
12	🟢	1	4933	312	5341	1	20116	5029.0	1	5029	312	5341	1	65515	0	10839916	359407	0	1033334	12298172
13	🟢	1	3539	465	4144	1	14716	3679.0	1	3679	465	4144	1	0	1992646	3682809	551499	0	800106	6427060
14	🟢	1	3124	302	3521	1	12876	3219.0	1	3219	302	3521	1	0	0	2799853	609930	0	692699	4102462
15	🟢	1	3506	342	3953	1	14444	3611.0	1	3611	342	3953	1	48178	0	249025	653380	0	968125	1918708
16	🟢	1	3032	314	3442	1	12512	3128.0	1	3128	314	3442	1	0	0	156335	729966	0	775511	1661812
17	🟢	1	2985	317	3396	1	12316	3079.0	1	3079	317	3396	1	0	0	206006	711338	0	869418	1786762
18	🟢	1	1777	122	1943	1	7284	1821.0	1	1821	122	1943	1	0	0	0	70208	0	449513	519721
19	🟢	1	3087	169	3319	1	12600	3150.0	1	3150	169	3319	1	0	0	6362737	134165	0	1102521	7599423
20	🟢	1	2745	162	2966	1	11216	2804.0	1	2804	162	2966	1	0	0	2060210	160190	0	1315912	3536312

Post-nightly pipeline Data Products

- * Recalibration (if necessary)
- * Difference images (eg. in WSA: GPS)
- * Deep stacks (eg. in WSA: DXS)
- * Mosaics (eg. in WSA: UDS; Nottingham)
- * Quality control by checking correct magnitude limits



Provenance, Quality Control

- ★ *Provenance*: Information about file provenance
- ★ *Quality Control*: done by archive scientists, some files marked as deprecated, e.g.:
 - 0 = good data
 - 60-70 = eyeball check deprecation
 - >127 = frame superceded by reprocessing
- ★ Update of *Quality Error Bit Flags* (ppErrBits): deblended sources, bad pixel, boundary sources, saturated sources, cross-talk artefacts.



Source Merging & Reseaming

- ★ *Source merging*: creation of multi-colour, multi-epoch source lists from individual passband detections
- ★ *Reseaming*: flagging of duplicate objects as primaries or secondaries (relies on quality error bit flag).



Neighbour Tables

- * Neighbour tables between sources and detections to quickly generate light curves and look for variable objects
- * Neighbour tables with external catalogues
 - SDSS, MGC (Millennium Galaxy Catalogue), NVSS (NRAO VLA Sky Survey), 2XMM, 2MASS, ROSAT, IRAS, FIRST, GLIMPSE, SSA, WSA, ...



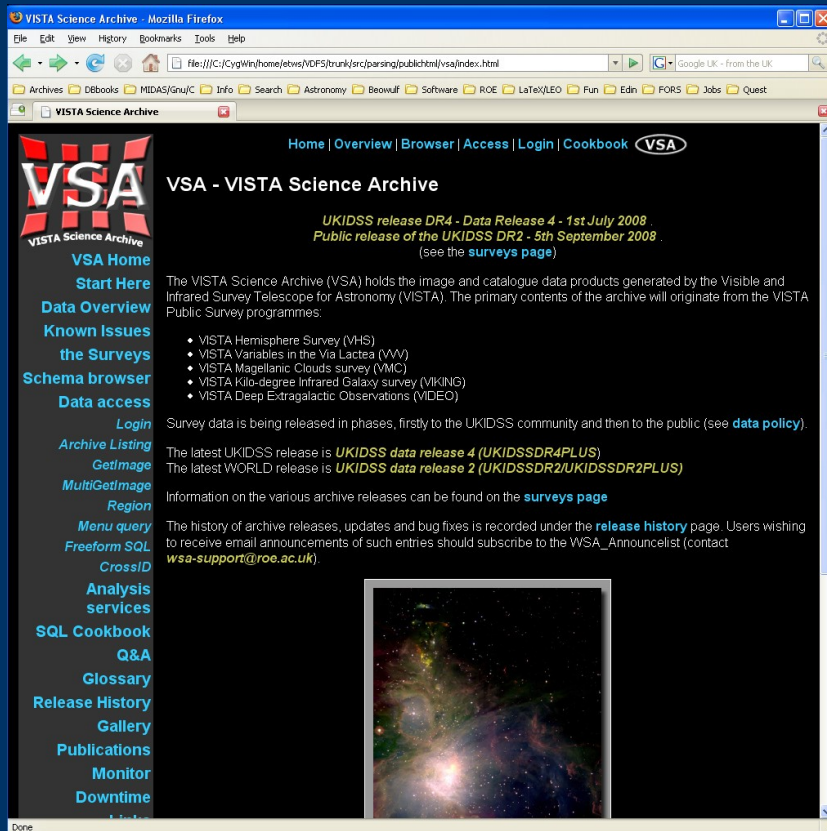
Release

- ★ Creation of the Release Database
 - Created in individual file-groups for faster access
- ★ Copying to Public Catalogue Server
 - Update of VSA Browser pages to reflect changes
- http://horus.roe.ac.uk/vsa/www/vsa_browser.html

VSA Browser

Basic outline can be found at:

http://horus.roe.ac.uk/vsa/www/vsa_browser.html



The screenshot shows a Mozilla Firefox browser window displaying the VSA Science Archive website. The browser's address bar shows the file path: `file:///C:/CygWin/home/etwv/NDFS/trunk/src/parsing/publichtml/vsa/index.html`. The website has a dark background with a navigation menu on the left and a main content area on the right. The main content area features the VSA logo, a navigation bar with links like Home, Overview, Browser, Access, Login, and Cookbook, and a section titled "VSA - VISTA Science Archive". Below this, there is a news item about "UKIDSS release DR4 - Data Release 4 - 1st July 2008" and "Public release of the UKIDSS DR2 - 5th September 2008". A list of survey programs is provided, including VISTA Hemisphere Survey (VHS), VISTA Variables in the Via Lactea (VVV), VISTA Magellanic Clouds survey (VMC), VISTA Kilo-degree Infrared Galaxy survey (VIKING), and VISTA Deep Extragalactic Observations (VIDEO). A sidebar on the left contains various utility links such as Login, Archive Listing, GetImage, MultiGetImage, Region, Menu query, Freeform SQL, CrossID, Analysis services, SQL Cookbook, Q&A, Glossary, Release History, Gallery, Publications, Monitor, and Downtime. At the bottom of the main content area, there is a small image of a galaxy.

VSA Browser: Access

Basic outline can be found at:

http://horus.roe.ac.uk/vsa/www/vsa_browser.html

The screenshot shows the VSA Science Archive website in Mozilla Firefox. The page is titled "VSA - Data Access" and features a navigation menu on the left and a list of data access tools on the right. The navigation menu includes links for VSA Home, Start Here, Data Overview, Known Issues, the Surveys, Schema browser, Data access, Login, Archive Listing, GetImage, MultiGetImage, Region, Menu query, Freeform SQL, CrossID, Analysis services, SQL Cookbook, Q&A, Glossary, Release History, Gallery, Publications, Monitor, and Downtime. The main content area is titled "VSA - Data Access" and contains a list of tools: Archive listing, GetImage, MultiGetImage, Region search, Menu query, Freeform query, and CrossID. Below the list, there is a section titled "The results i.e. rows selected from the database, are displayed in an HTML table and/or written to a file (ASCII, FITS or VOTable format) that can be downloaded." and a section titled "Users should note the following general points:" which lists several limitations and warnings. At the bottom of the page, there is a footer with links for Home, Overview, Browser, Access, Login, Cookbook, Listing, Region, MenuQuery, FreeSQL, Links, and Credits.

VSA - Data Access

Users can access the data held in the VSA through web-based forms. These forms parse the user's input parameters and submit SQL (Structured Query Language) queries to the database.

- **Archive listing** - retrieve listings of the multiframe held in the VSA. Links are returned allowing the user to view the library jpegs and download the FITS files.
- **GetImage** - extract cut-out images around a given position
- **MultiGetImage** - upload a list of coordinates and extract matching cut-out images
- **Region search** - search the VSA object catalogues around a supplied position
- **Menu query** - build simple SQL queries using a series of forms
- **Freeform query** - submit an SQL query directly
- **CrossID** - upload a list of coordinates and extract matching objects

The results i.e. rows selected from the database, are displayed in an HTML table and/or written to a file (ASCII, FITS or VOTable format) that can be downloaded.

Users should note the following general points:

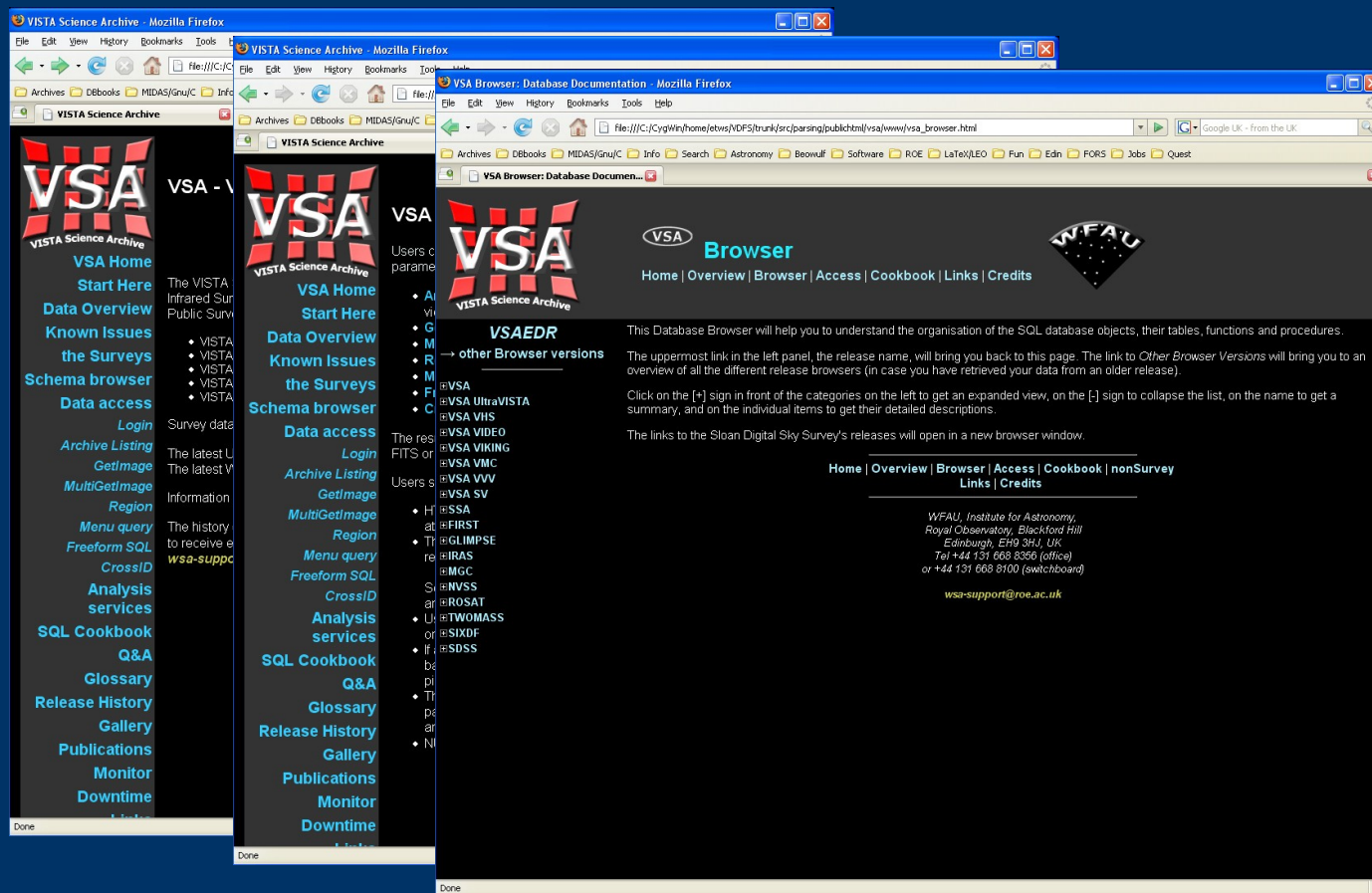
- HTML table output is only intended as a summary and the number of rows displayed is limited to 100. A note at the end of the table informs the user if this was exceeded.
- The number of result rows written to files is also limited and depends on how many parameters have been requested i.e. $maximum\ rows\ written\ to\ file = \min(15000 / no.\ parameters) \times 1000$. So if only three parameters have been requested then the file can contain up to 5 million rows. Again users are warned if the limit was exceeded and an indication of how many extra rows were returned is provided.
- Users who exceed the file row limit should submit their query in parts e.g. by querying sections in RA or Dec or magnitude slices.
- If an email address is supplied queries are allowed to run for 30 seconds before they are placed in the background and the browser window is released. On completion an email is sent informing the user where to pick up the results.
- The tables in the database do not contain any NULL values. Where values are unavailable for a given object parameter default values have been inserted. Users should be aware of this when constructing their queries and when interpreting the results. See the **schema browser** for details of a given parameter's default value.
- NULL values can be returned if users JOIN tables. If this occurs the values are written out as zeroes.

Home | Overview | Browser | Access | Login | Cookbook
Listing | Region | MenuQuery | FreeSQL
Links | Credits

VSA Browser: Schema Browser

Basic outline can be found at:

http://horus.roe.ac.uk/vsa/www/vsa_browser.html



VSA Browser: DB Objects

Basic outline can be found at:

http://horus.roe.ac.uk/vsa/www/vsa_browser.html

The image shows a screenshot of the VSA Browser website, which is a database interface for the VISTA Science Archive. The page is titled "VSA Browser: Database Documentation" and displays the schema for "VSA_MultiframeSchema.sql". The website has a dark theme with a navigation menu on the left and a main content area on the right. The navigation menu includes links for "VSA Home", "Data Overview", "Known Issues", "the Surveys", "Schema browser", "Data access", "Login", "Archive Listing", "GetImage", "MultiGetImage", "Region", "Menu query", "Freeform SQL", "CrossID", "Analysis services", "SQL Cookbook", "Q&A", "Glossary", "Release History", "Gallery", "Publications", "Monitor", and "Downtime". The main content area shows the "VSA Browser" logo and a list of database objects, including "Database Objects", "Functions", "Views", "Tables", and "Views". The "Database Objects" section is expanded, showing a list of tables and their descriptions. The "CurrentAstrometry" table is highlighted, and its description is shown: "Contains the current astrometric calibration coefficients." The "Multiframe" table is also highlighted, with the description: "Contains details of all multiframe frames stored in the archive." The "MultiframeDetector" table is highlighted, with the description: "Contains details of individual detector frames that are part of a multiframe." The "PreviousAstrometry" table is highlighted, with the description: "Contains previous astrometric calibration coefficients." The "PreviousMFDZP" table is highlighted, with the description: "Contains previous zeropoints for each multiframe and detector combination." The "ProgrammeFrame" table is highlighted, with the description: "Contains details of which image(s) belong to which Programme(s)." The "Provenance" table is highlighted, with the description: "Describes the provenance of all combined frames in the archive." The "Latest Revision" is shown as "5016, 2008.08.19 09:55:05 (Tue, 19 Aug 2008)".

VSA Browser: Tables

Basic outline can be found at:

http://horus.roe.ac.uk/vsa/www/vsa_browser.html

The screenshot displays the VSA Browser interface in Mozilla Firefox. The main content area shows a table titled "TABLE ultravistaDetection". The table contains the following data:

Name	Type	Length	Unit	Description	Default Value	Unified Content Descriptor
multiframeID	bigint	8		the UID of the relevant multiframe		ID_FRAME
extNum	tinyint	1		the extension number of this frame		NUMBER
cuEventID	int	4		UID of curation event giving rise to this record		REFER_CODE
seqNum	int	4		the running number of this detection (catalogue TType keyword: Sequence_number)		ID_NUMBER
filterID	tinyint	1		UID of combined filter (assigned in VSA: 1=Z,2=Y,3=L,4=H,5=K,6=blank)		INST_FILTER_CODE
isoFlux	real	4	ADU	instrumental isophotal flux counts (SE_FLUX_ISO) (catalogue TType keyword: isophotal_flux)		PHOT_INTENSITY_ADU
isoMag	real	4	mag	Calibrated isophotal magnitude		PHOT_INT-MAG
x	real	4	pixels	X coordinate of detection (SE_X_IMAGE) (catalogue TType keyword: X_coordinate)		POS_PLATE_X
xErr	real	4	pixels	Error in X coordinate (SE_ERRX2_IMAGE%) (catalogue TType keyword: X_coordinate_err)		ERROR
y	real	4	pixels	Y coordinate of detection (SE_Y_IMAGE) (catalogue TType keyword: Y_coordinate)		POS_PLATE_Y
yErr	real	4	pixels	Error in Y coordinate (SE_ERRY2_IMAGE%) (catalogue TType keyword: Y_coordinate_err)		ERROR

Summary

- * ESO's release policy
 - Major data releases with uniform photometric and astrometric calibrations at least once per year
 - First delivery expected no later than 18 months after beginning of the observations.
- * DBs are in place and awaiting data
- * VSA Browser is available
 - Test functionality with WSA DR2

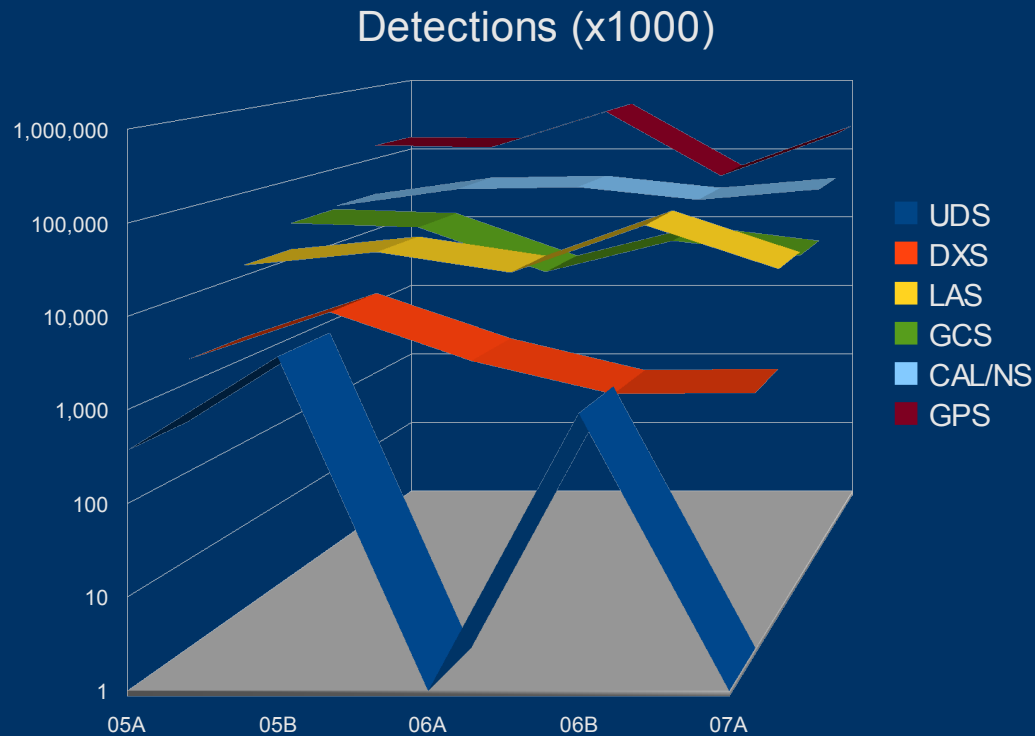


For questions & feedback please contact
nch@roe.ac.uk
etws@roe.ac.uk



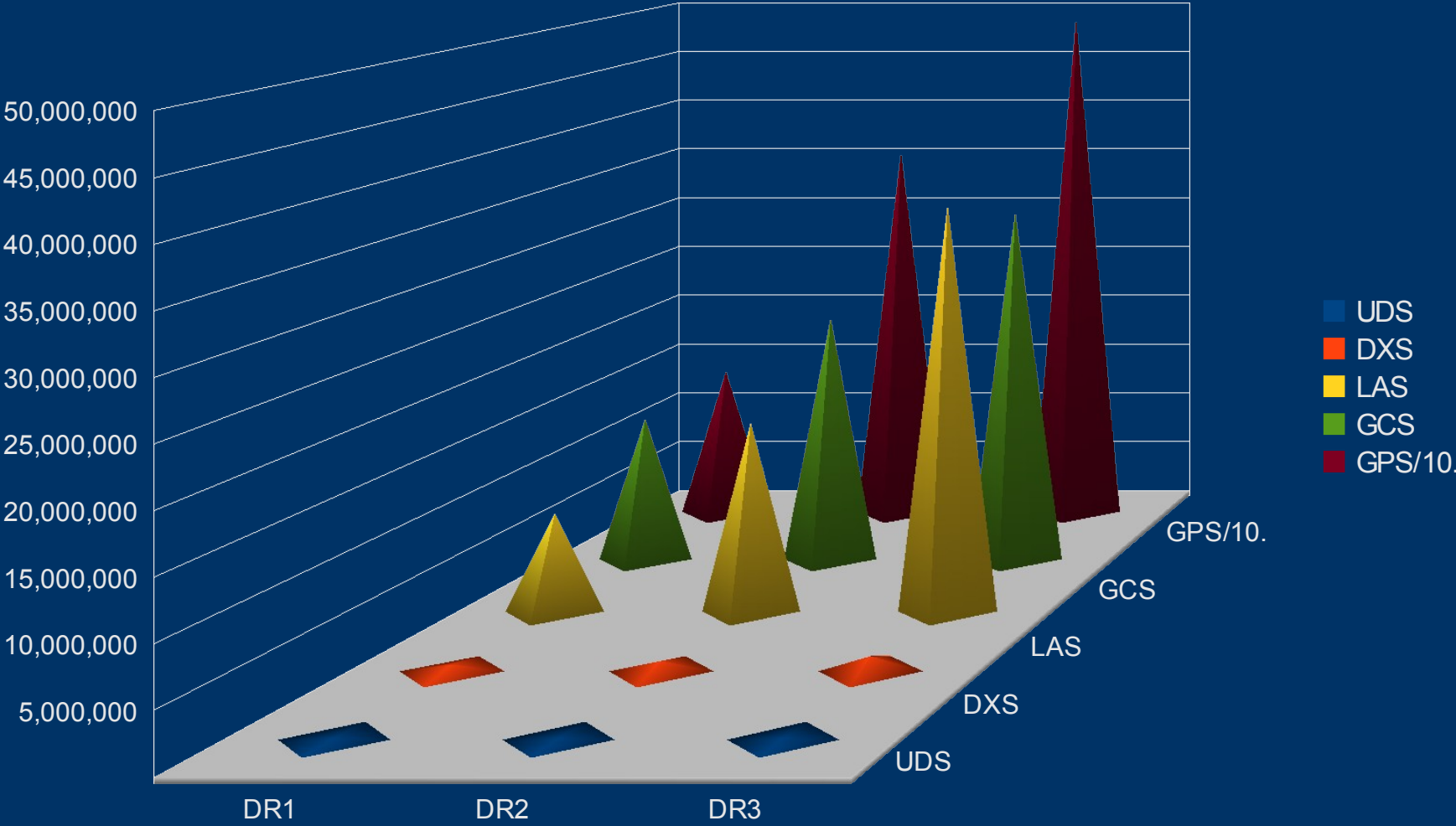
Some Ingest Statistics

		Obs. Nights	Transfer/Ingest (cat.)	JPEGs	Detections (x1000)
DR1	05A	72	220,581 (4%)	844,748	225,502
	05B	140	280,296 (8%)	1,575,638	249,257
DR2	06A	88	223,824 (6%)	845,356	525,721
	06B	121	356,625 (7%)	1,325,244	154,757
DR3	07A	51	168,483 (6%)	633,640	292,503



More Statistics

Number of Sources per Survey



Timeline

DR3

Transfer	7h per observation night	50d (~7d after last pipeline proc.)
JPEGs	2h calc. + 0.3h ingest per observation night	11d (~7d in parallel mode) + 2d ingest
Metadata	0.2h calc. + 0.2h ingest per observation night	1.2d calc. (0.6d parallel) + 1.2d ingest

Catalogue data	1h calc. + 1h ingest per observation night	7d calc. (4d parallel) + 7d ingest
----------------	--	------------------------------------

Recalibration		4d
GPS diff. images		1.5d
UDS stacks		1d
DXS mosaics		3d
Provenance		1d
Quality Control	depending on PI input	28d
Error bit flags		2d
Source merging	UDS+DXS: 0.5h; GCS: 1d; LAS: 2d; [GPS: 3d]	3d
Resampling	UDS+DXS: 2h; GCS: 2d; LAS: 0.5d; [GPS: ~14d]	2.5d
Neighbour tables	UDS+DXS: 1h; GCS: 1h; LAS: 4d	4d
Release database	15h creation; 7h copying	1d

Total Computational Time		80d
--------------------------	--	-----