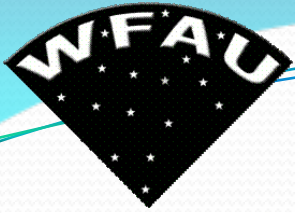
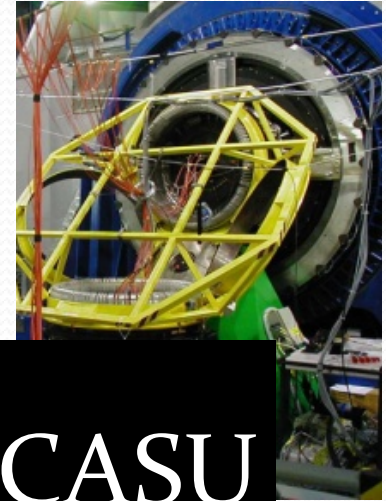


WFAU Science Archives: Facilitating Survey Astronomy

Nicholas Cross, Rob Blake, Ross Collins, Clive Davenhall, Mark Holliman, Mike Read, Eckhard Sutorius, Nigel Hambly, Dave Morris, Stellios Voustanis, Bob Mann, Keith Noddle
Wide Field Astronomy Unit, Institute for Astronomy, University of Edinburgh, UK
njc@roe.ac.uk



WFAU Science Archives



EARLIER PROCESSING AT CASU
See Gonzalez-Solares talk



WFAU

Roles of the Science Archive

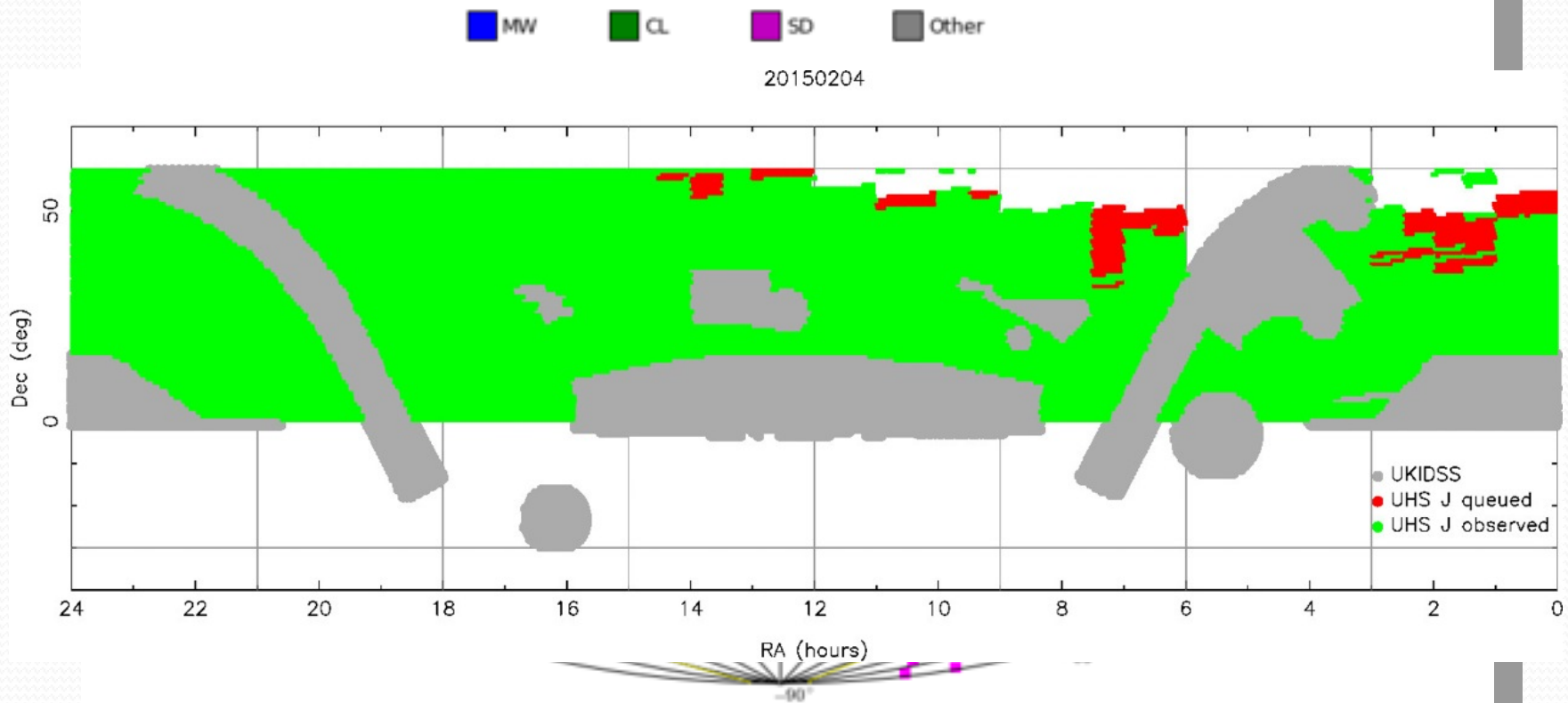
- **Repository** of (e.g. VISTA) data from reduced images to complex, catalogue products.
- Interface to do **science – select data**.
- Interface for teams to check data for **quality control**.
- Requires both:
 - a continuously updated **dynamic main-DB**: new data, better calibration, reprocessing, quality control, higher order products.
 - Static, well documented **release-DBs** that can be referred to in publications.

Common Design

- WSA, VSA, OSA developed through **VDFS** – VISTA Data Flow System
- MS-SQL database system based on **SDSS SkyServer**
- Similar inputs: **FITS files generated by CASU**
- **Common software architecture**, data pipeline, user interfaces, hardware architecture.
- **GES is significantly different**: spectroscopic, wider range of data from multiple groups, different interfaces

Surveys

WF VST: ATLAS, [VPHAS+] DRING, VIDEO



Archives: <http://surveys.roe.ac.uk/>

4.2 Selecting nearby cross-matches

Suppose a user wishes to select the identifiers and co-ordinates of all Faint Source Catalogue sources from the ROSAT All-Sky Survey that are within 6 arcsec of a VHS source. The SQL to do this is as follows:

```
SELECT vhs.sourceID, vhs.ra, vhs.dec, fsc.seqNo, fsc.ra, fsc.dec
FROM vhsSource AS vhs, ROSAT..rosat_fsc AS fsc, vhsSourceXrosat_fsc AS x
WHERE x.masterObjID=vhs.sourceID AND x.slaveObjID=fsc.seqNo AND x.distanceMins<0.1
```

Note: i) the use of table aliases *vhs* and *fsc* which are a convenient short-hand; ii) the use of attribute prefixes (e.g. *vhs.ra*) to distinguish between attributes in different tables that happen to have the same name; iii) the table joining condition in the *WHERE* clause, which selects associated rows (if you omit this condition, you will get all rows of each table joined to all other rows, i.e. an extremely large and useless dataset!); and finally iv) the specification of a maximum radius of 0.1 arcmin (=6 arcsec) for this query, where the maximum available for ROSAT joins is 30 arcsec. You can see the maximum joining radii (in arcsec) available to you in each neighbour/cross-neighbour table by issuing the following query:

```
SELECT neighbourTable, joinCriterion*3600.0 as maxRadius
FROM RequiredNeighbours
```

When selecting *nearby* cross-matches, users should note that one or more than one rows may result for each master object, since there may be more than one slave neighbour within the neighbourhood defined by the maximum join criterion specified. If you want the nearest match only, then the next section explains how to do this.

4.3 Selecting the nearest cross-match

Suppose a user wishes to select the infrared and optical point source photometry for a sample (say the first 50 rows for speed) in the VHS and SDSS DR2 overlap, taking the *nearest* SDSS source to each VISTA source as being the most likely match, where the maximum positional tolerance is set to 2 arcsec. The following query will do the job:

```
SELECT TOP 50 vhs.ra, vhs.dec, yAperMag3, jAperMag3, hAperMag3, ksAperMag3, psfMag_u, psfMag_g, psfMag_r, psfMag_i, psfMag_z
FROM vhsSource AS vhs, BestDR2..PhotoObj AS dr2, vhsSourceXDR2PhotoObj AS x
WHERE masterObjID=vhs.sourceID AND slaveObjID=dr2.ObjID AND distanceMins<0.033333 AND sdssPrimary=1 AND distanceMins IN (
SELECT MIN(distanceMins)
FROM vhsSourceXDR2PhotoObj
WHERE masterObjID=x.masterObjID AND sdssPrimary=1)
```

Imaging Surveys Pipeline

- **Ingest** from CASU (and teams)
- **Provenance** - link related images
- **Quality Control**: Automated + input from teams.
- Process data for **semester** - done **per programme**:
 - Produce and ingest **deep** stacks/tiles/mosaics + catalogues
 - **Merge** pass-band catalogues to create source tables
 - Create **neighbour** tables to link external catalogues
 - Link multi-epoch data and calculate **variability** statistics
- **Release** a documented, static data product to users
- Convert data to ESO format and **copy to ESO** (VISTA, VST)

Gaia-ESO Spectroscopic Archive

- Spectroscopic survey of **~100,000 stars**, covering all major components of MW using **VLT-FLAMES**
- Archive includes:
 - list of **targets** with photometry used in the target selection
 - **spectra** processed by CASU
 - atomic and molecular **line lists**
 - **Recommended astrophysical parameters** derived by various working groups:
 - T_{eff} , $\log(g)$, abundances ...
 - **Neighbour tables** to other surveys: SuperCOSMOS, VST-ATLAS, 2MASS, VISTA-VHS
- No public release yet. 3rd internal data release.

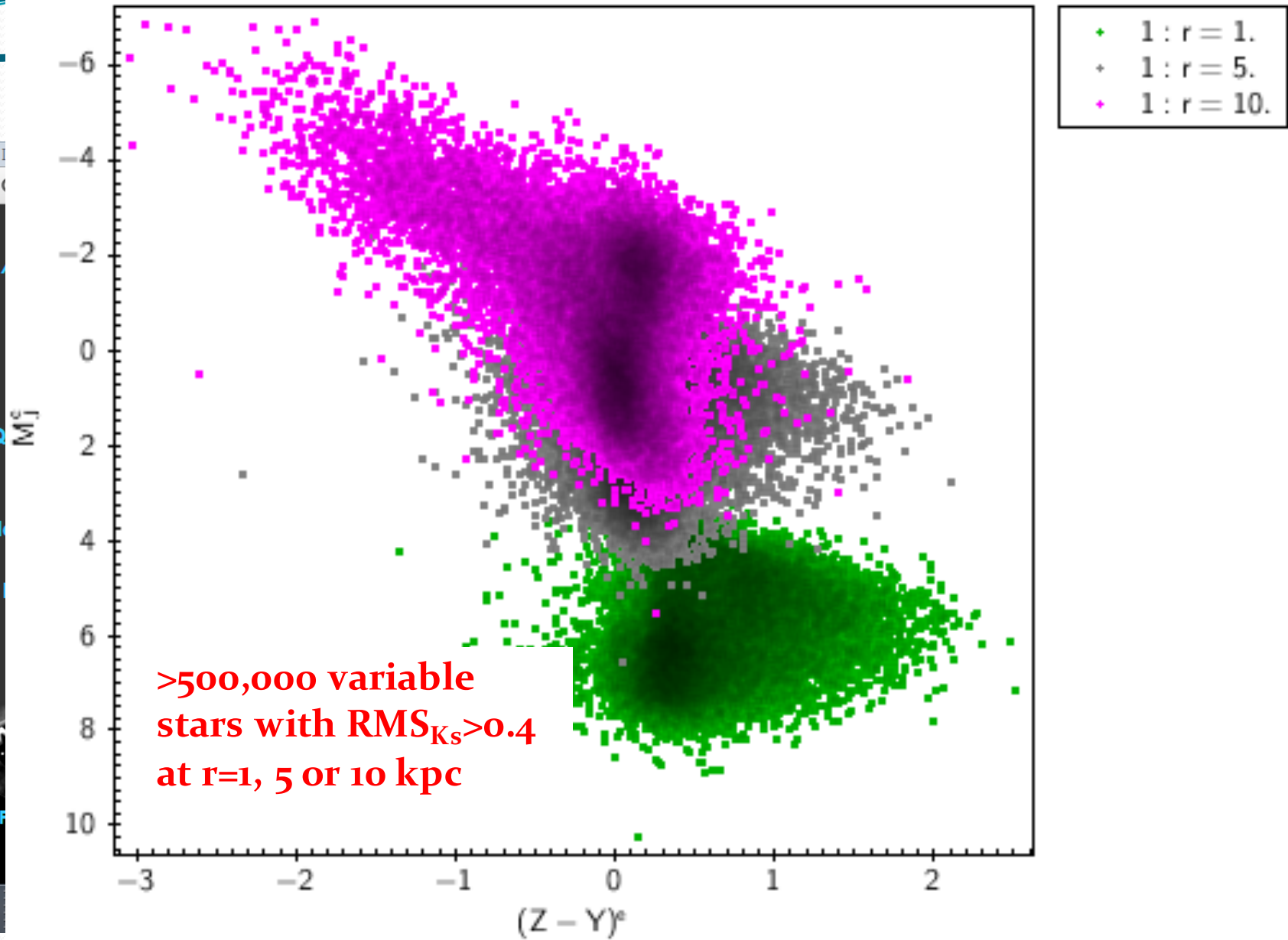
Value Added Data

- Data created by science teams:
- VMC :
 - PSF catalogues
 - Variable types (Cepheid, Eclipsing Binary)
- VVV:
 - 3D extinction maps
 - PSF catalogues (multi-band 1 epoch) + DIA light-curves [Discussions with Javier Alonso-Garcia and Eamonn Kerins]
 - Proper motion + Parallax [Not yet delivered]
- GES:
 - Wide range of parameters from multiple working groups.
- Work to improve connections of team generated data with CASU/WFAU generated to maximise usefulness.

3D Extinction Maps

- Included VVV extinction maps of bulge (Chen et al. 2013)
- **Match** to vvvSource.
- **Filter conversions** for VISTA, UKIRT, VST, Spitzer and WISE filters
- Will add in more (IPHAS, PanSTARRS ...)

3L



Survey Interfaces

- **Coverage Maps** – Multi-Order Coverage
- **Archive Listing** - Image/Flat file download
- **GetImage** / Multi-GetImage – Image cutouts
- **Region** – Cone search
- **CrossID** – Catalogue matching
- **Freeform SQL** – Most powerful catalogue query tool
- **ADQL** (OSA only so far) – Astronomical Data Query Language. To do distributed queries between archives (e.g. Gaia, PanSTARRS, LSST)
- **Firethorn** – MyDB works with ADQL for distributed multi-wavelength work (Under development)

Coverage Maps

Firefox 5th VV Science M... Aladin v7.5

Firefox 5th VV Science M... Aladin v7.5

horus.roe.ac.uk/vsa/coverage-maps.html

Gallery
Publications
Monitor
Downtime
Links

WFAU
IFA ROE

- DR1
- VISTA Kilo-Degree
- VISTA Deep Extra
- Other WFAU Hosted
 - UKIDSS Surveys
 - OmegaCAM Surveys
 - Other Surveys M...
 - Deep Near-Inf...
 - Expérience po...
 - Faint Images
 - Glimpse Surveys
 - Infrared Astro...
 - MACHO Surveys
 - Magellanic Cl...
 - Millennium Ga...
 - ROSAT
 - Sloan Digital S...
 - Sloan Digital S...
 - Spitzer Space
 - 2XMM Survey
 - Wide-field Inf...

File Edit Image Catalog Overlay Tool View Interop Help

Location 17:55:15.61 -29:07:42.1

Frame ICRS

Optical DSS Simbad NED

WFAU DR2

select cont
pan pixel
zoom prop
dist del
phot
draw tag
filter
x-y
rgb
assoc
crop

Mouse controls:
Left: source selection.
Middle: quick panning.

2XMM
GPS Any Band DR2
WFAU DR2

size - +
opac. - +
zoom - +

Frame: ICRS

+90
+180
-90
-180

17:44:47.32 -31:35:11.3
14.92° x 15.7°

Search

(c) 2012 UDS/CNRS - by CDS - Distributed under GNU GPL v3

0 sel / 0 src 33Mb

EN 21:12
02/04/2014

An Example: Selecting point source variables in VIDEO

The screenshot displays a web browser window with the URL `horus.roe.ac.uk:8080/vdfs/VSQL_form.jsp?`. The search bar contains the ID `644245293554`. The database interface shows a SQL query editor with the following text:

```
SELECT  
v.tbl  
(v.i  
v.k:  
(v.i  
FROM  
s /  
WHEI  
/* )  
s.i  
v.v:  
/*  
at :  
((:  
AND  
AND  
((j)  
jnG  
AND  
((k:  
ANI
```

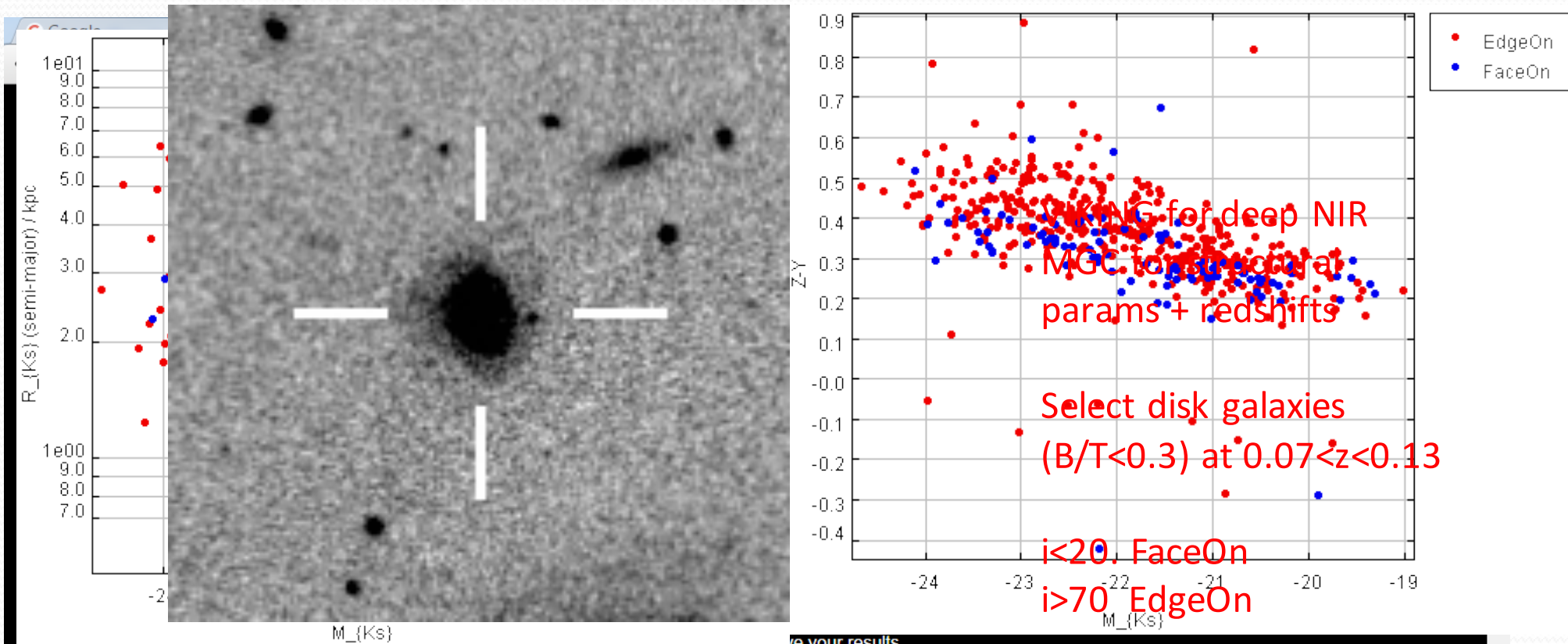
A red dot is visible in the top right corner of the video frame, and a white circle highlights a specific point source. The browser's taskbar at the bottom shows the date `28/03/2012` and time `22:52`.

Select high amplitude point source variables from VIDEO

Surveys are not independent entities

- Importance of multi-wavelength astronomy
- Common survey fields
 - VIKING and KiDS
 - VST-ATLAS and VHS
 - UKIDSS LAS/UHS and SDSS
 - VVV, GPS, IPHAS, VPHAS+
- Need for data integration
 - Cross-neighbours tables, publishing to VO
 - ADQL + Firethorn to be able to link to any published.
 - Matched aperture photometry

External Datasets



Monitor Downtime Links

Email Address: [redacted] the results of long running queries will be sent by email.

- HTML table summary (results are NOT saved to file)
- ASCII FILE (downloadable with HTML table summary on screen)

KsFaceOn.jpg | KsEdgeOn.jpg

Show all downloads...

EN 10:30 07/10/2015

Data Volumes

- Images:
 - WFCAM 1720 raw image frames a day (4 2kx2k)
 - VISTA 580 raw image frames a day (16 2kx2k)
 - VST 100 raw image frames a day (32 4kx2k)
 - **~1PB of image data**
- Catalogues:
 - VVV completely dominates: combination of high source density, short exposure time.
 - Next VVV release will be 45TB (**Few weeks away**)
 - 55 billion detections.
 - Largest astronomical SQL database?
 - **Final VVV+VVVX (?) >100TB**

WFAU VDFS Publications

Hambly et al. 2008, MNRAS, 384, 637 (WSA)

Cross et al. 2009, MNRAS, 399, 1730 (Multi-Epoch processing)

Cross et al. 2012, A&A, 548A, 119 (VSA)

Various other technical documents:

<http://surveys.roe.ac.uk/vsa/publications>

Future work – some highlights

- **Matched Aperture Pipeline:**
 - imcore_list & SExtractor
 - Between surveys, inter survey, user defined
- **Variable Star Selection:**
 - Add in improved indices for variable selection
- Much more **value added data**
- **MOC driven cone-searches**
- **MOONS?**

Archive Usage

