Archive and Data products

Presented by Magda Arnaboldi
Head - Archive Science Group
ESO Survey Team leader

marnabol@eso.org

ASG (M. Arnaboldi, N. Delmotte. L. Mascetti, A. Micol, J. Retzlaff)

Back-End Operation Department

EST (M. Arnaboldi, M. Hilker, G. Hussain, M. Petr-Gotzens, M. Rejkuba)

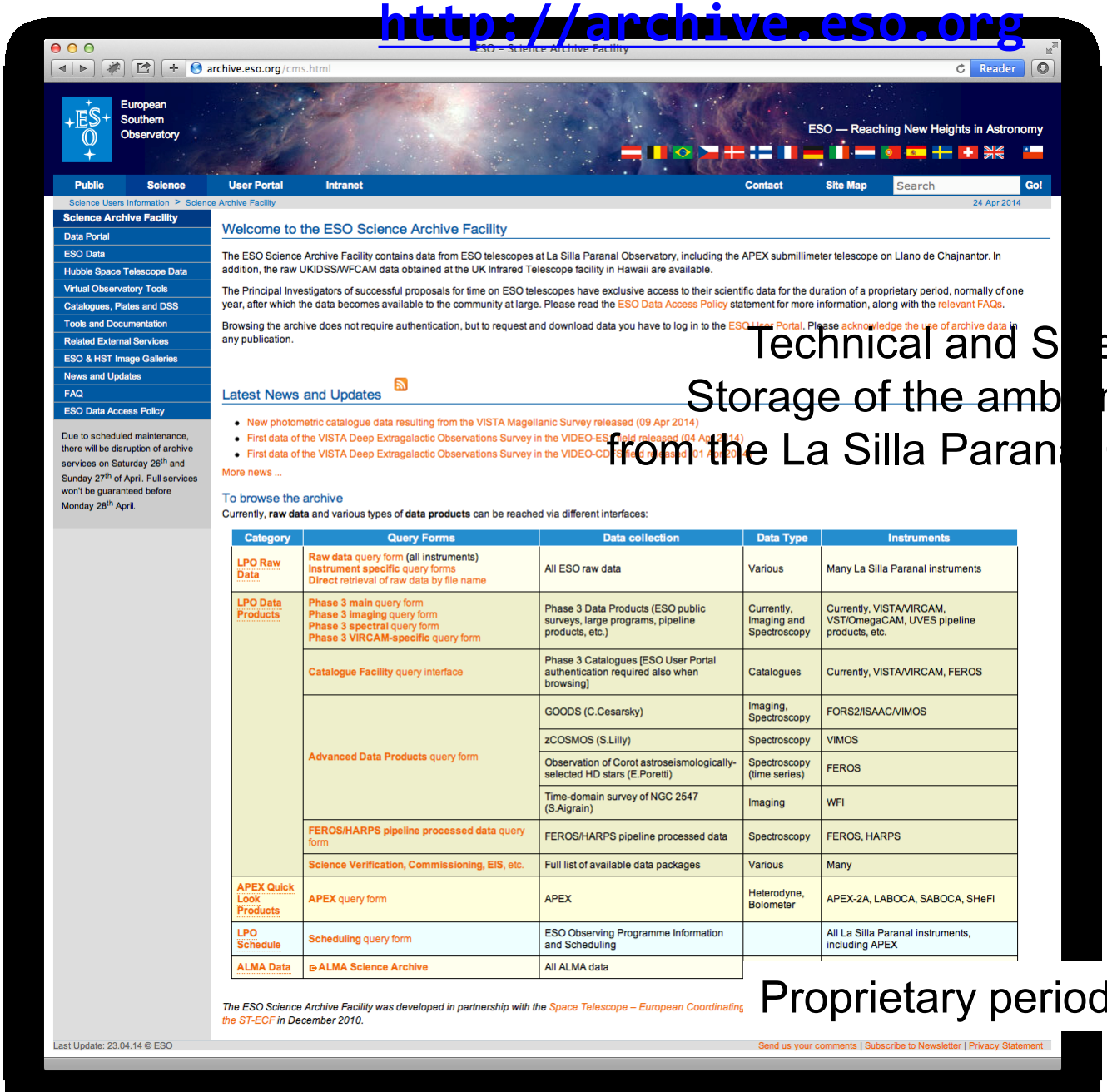
Outline

- ESO Science Archive Facility - SAF
- ESO Telescopes & Public Surveys
- Science Data Products
  - ✓ In house reprocessing
  - ✓ Contributed products
- Phase 3 data releases & query forms
- Building a community, impact & statistics from ESO SAF, highlights

ESO Science Archive Facility

ESO Science Archive Facility (SAF)

http://archive.eso.org



Since 1998

Technical and S ence Archive

Storage of the amb nt conditions

from the La Silla Paran Observatory

Proprietary period is one year

ESO Science Archive Facility (SAF)

http://archive.eso.org

- Data generated at the telescopes need to be processed before science information can be extracted
- What does ESO do?
  - Archive the data!
  - Provide science-ready processed data
  - Support customized data processing by end users:
    - Pipelines
    - Workflows
- Why does ESO do it?
  - ESO's mission: provide the best **science data** to our community, to investigate and understand our Universe
  - Increasing complexity of instruments and data → need to support the community in exploiting the information

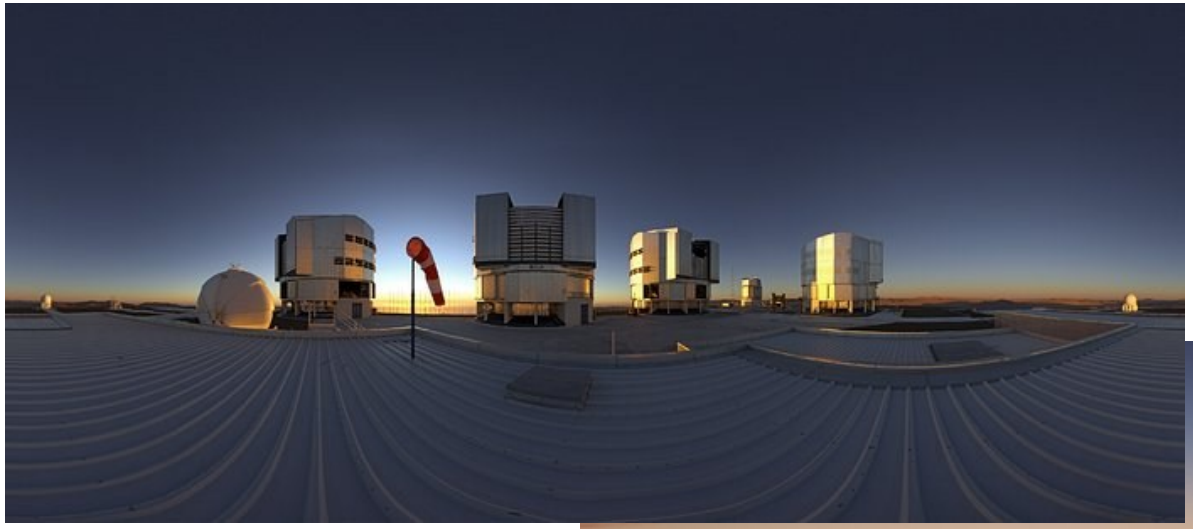
Operations @ the ESO SAF

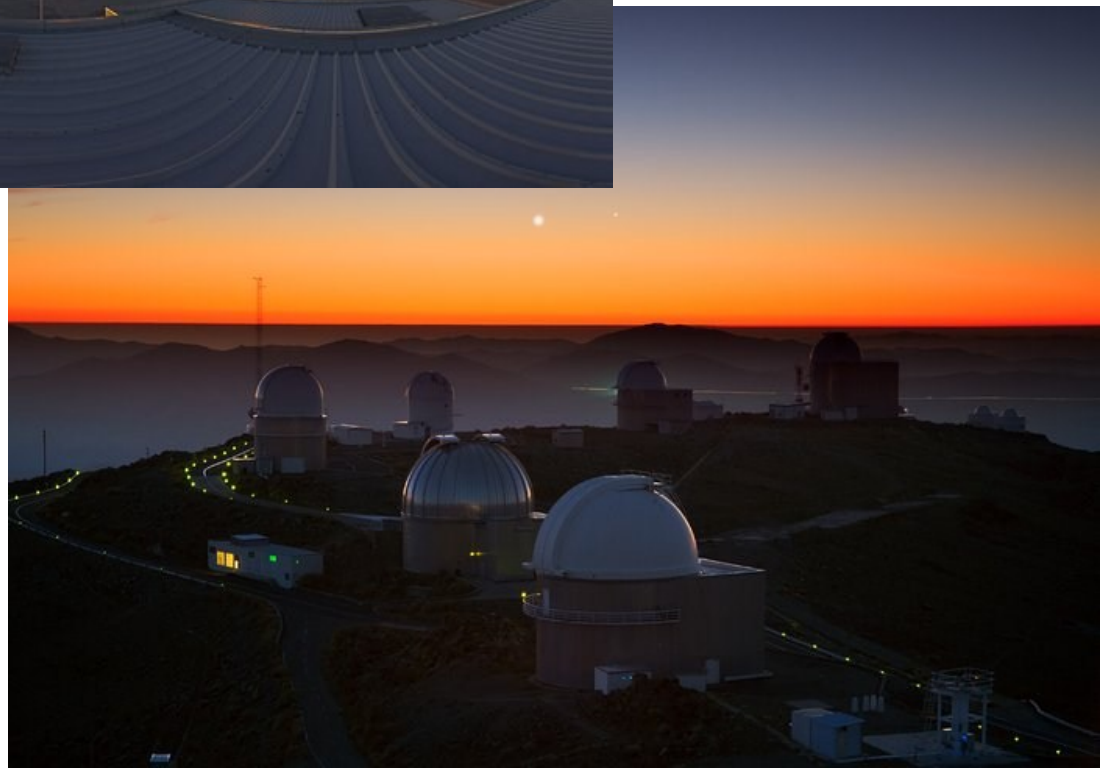
- Operations of the Archive
  - Inflow: ~10 TB/month; outflow ~15 TB/month
  - Total archive holdings: ~630 TB of data in 29 million files and ~26.5 billion database rows to store header keywords
  - Infrastructure: on-going studies to support the evolution of software and hardware – storage and browsing infrastr.
- Phase 3 process
  - Ingestion, validation and publication of science data products in routine operation, from external users and in-house repr.
  - Upgrade near completion, to enhance capability and user support
  - Standard evolved to include new products – flux maps, 3D spectroscopy, VLTI

ESO Telescopes and Public Surveys

Neon School Garching, May 2016

ESO Telescopes

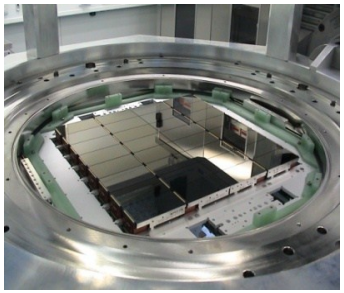




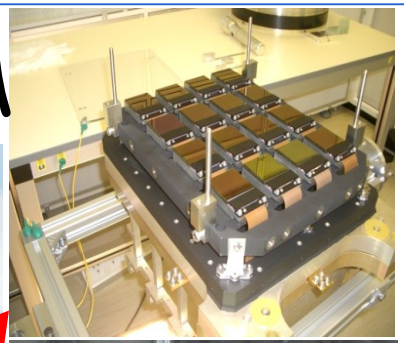
205 Gb/night of compressed data; Transferred within 24 hrs (EVALSO cable link)

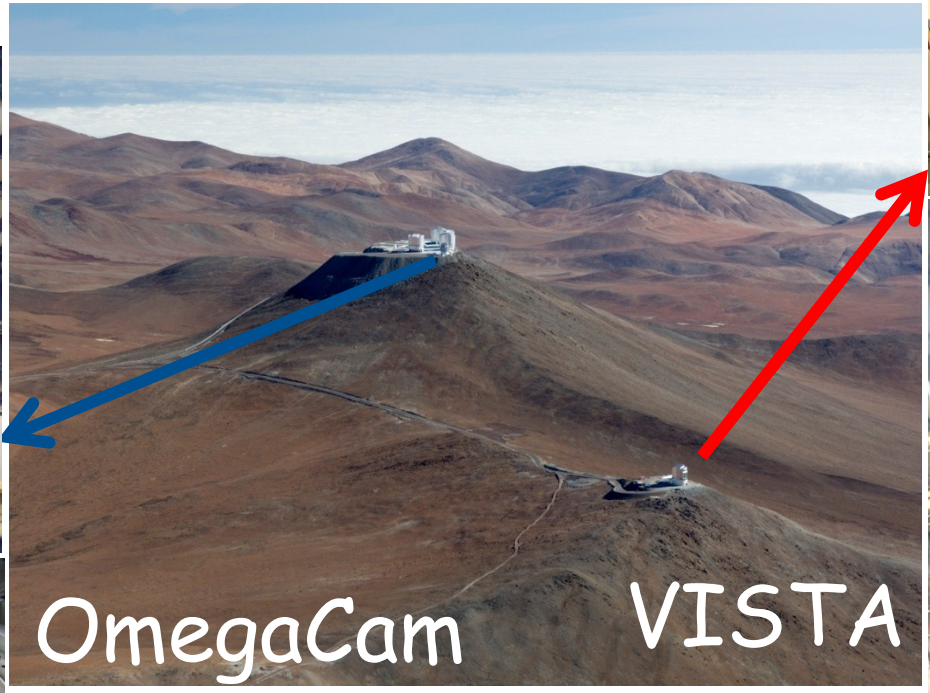
ESO Telescopes for surveys: VISTA & VST

VST



VIRCAM





- Observations carried out in service mode
- 1800 – 3400h per programme, several thousands Obs each semester
- Expected to run ~6-8 yrs

Optical & NIR surveys and telescopes

	Telescope Diameter	FoV	Etendu [m^2 deg^2]	Pixel size [arcsec/pixel]	Wavelength range [μm]
Ground					
SkyMapper	1.4	5.7 deg^2	6.6	0.500	0.32-0.95
Pan-STARRS 1	1.8	2.8ºx2.9º	16.3	0.300	0.4-1.15
SDSS	2.5		6.0	0.394	0.33-1.0
MegaCam	3.6	0.96ºx 0.94º	7.6	0.187	0.34-0.95
CTIO (DES)	4.0	4.0 deg^2	16.0	~0.20	0.6-1.0
VISTA	4.0	1.65ºdeg circ	5.2	0.339	0.88-2.15
SuprimeCam	8.0	34'x27'	13.5	0.200	0.36-1
VLT HAWK-I	8.2	7.5'x7.5'	0.81	0.106	1-2.2
VLT-VIMOS	8.2	4 x 7'x8'	3.2	0.205	0.36-1
LBC@LBT	8.4	23'x23'	8.1	0.230	0.32-1.0
VST	2.6	1 deg^2	5.5	0.210	0.33-1.0
Subaru HyperSuprime	8.0	1.5 deg circ	94	0.17	0.36-1
LSST	8.4	3.5 deg circ	319	0.200	0.33-1.07

ESO Telescopes for surveys: Flames@UT2, VIMOS@UT3 and EFOSC/SOFI@ NTT

FLAMES@UT2
VIMOS@UT3

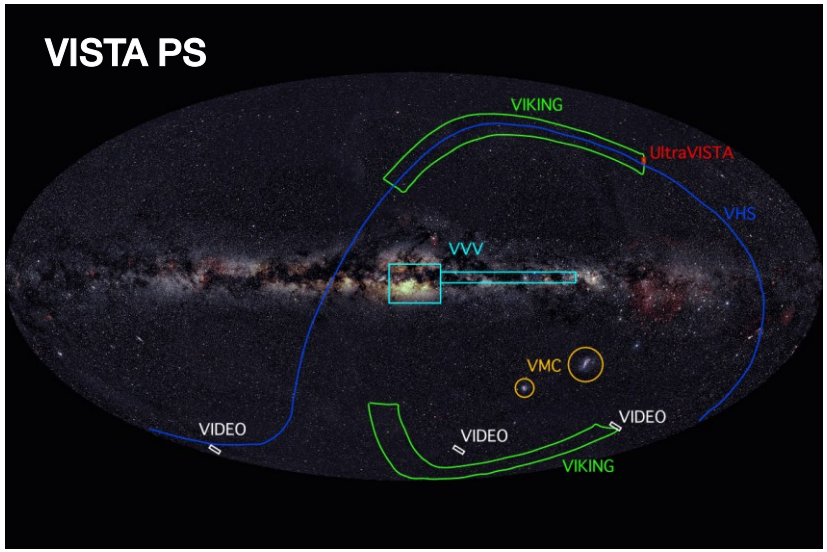
EFOSC&SOFI @ NTT

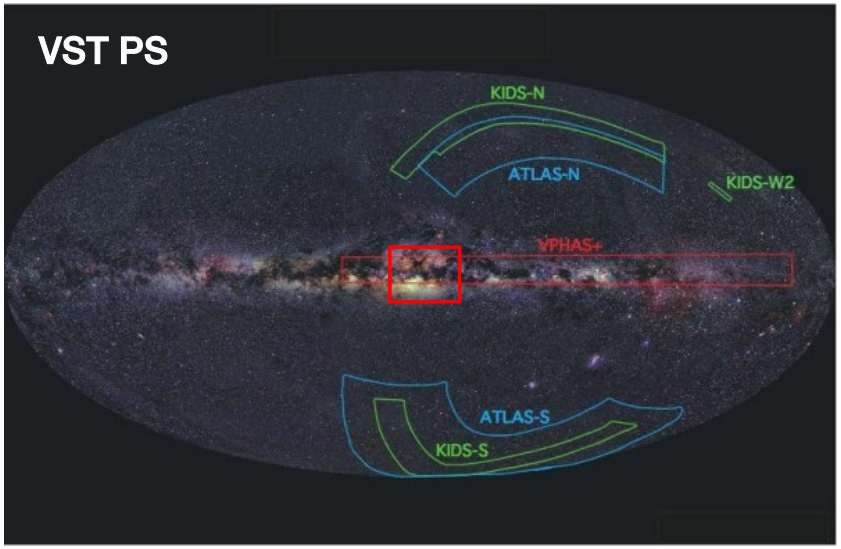
- Observations carried out in visitor mode
- Expected to run for 4-5 yrs

Future developments: MOONS@VLT, 4MOST@VISTA

ESO Public Surveys

- Public Surveys
  - Observing programs lasting more than 2 years – time allocation O(1000) hrs
  - Legacy value for astronomical community at large
  - **All raw observations are immediately public**
  - Survey teams commit to deliver reduced images/spectra and catalogues within ~yearly releases





On-going selection for the second VISTA call – Observations up to 2020

VISTA Public Surveys

Surveys	Area (deg^2)	Filters	Magnitude limit 5σ (AB), 10σ (AB) x VMC	Observation hours taken (Apr 16)
Ultra-VISTA	1.7 deep 0.73 ultra-deep	Y J H K_s Y J H K_s NB118	25.7, 25.5, 25.1, 24.5 26.7, 26.6, 26.1, 25.6 26.0	1801
VHS	17800	Y J H K_s	21.2, 21.1, 20.6, 20.0	4710
VIDEO	12.0	Z Y J H K_s	25.7 24.6 24.5 24.0 23.5	1942
VVV	560	Z Y J H K_s	21.9 21.1 20.2 18.2 18.1	Comp.
VIKING	1500	Z Y J H K_s	23.1 22.3 22.1 21.5 21.2	2200
VMC	180	Y J K_s	21.9, 21.4, 20.3	1887

Deep high z Whole Sky Galactic Extragalactic Resolved SFH

VST Public Surveys

Survey	Area (deg^2)	Filters	Magnitude limits	Depth measure	Observation hours taken (Oct 15)
KIDS(*)	1500	u' g' r' i'	24.1, 24.6, 24.4, 23.4	10σ (AB)	3225
ATLAS	4000	u' g' r' i' z'	22.0, 22.2, 22.2, 21.3, 20.5	10σ (AB)	1390
VPHAS+	~2000	u' g' Hα r' i'	21.8, 22.5, 21.6, 22.5, 21.8	10σ (AB)	899

Deep high z Whole Sky Galactic Extragalactic Resolved SFH

(*) Same area as VIKING

ESO Public Spectroscopic Surveys

- Gaia - ESO (PIs: Gilmore, Randich) – UT2 FLAMES
  - 10^5 stars in all major components of the MW, 100 open clusters
  - Synergy with Gaia – phase space structure and abundances for Milky Way stellar populations; Target selection: VHS, VVV, WFI and other optical photometry
- PESSTO (PI: S. Smartt) - EFOSC2 and SOFI on NTT
  - Follow up of ~150 transient objects in an unbiased sample of nearby galaxies drawn from ongoing surveys - SN explosion physics, SN progenitors
- VANDELS (PIs: R. McLure, L. Pentericci) – UT3 VIMOS
  - Deep survey of high z gals in CANDELS UDS and CDFS; about 2500 spectra (SFRgals $2.5 < z < 7.0$ and passive gals $1.5<z<2.5$).
  - Metallicities and velocity offsets from absorption and emission lines, for a detailed investigation of the physics of galaxies in the early Universe. - Synergy with multi-wavelength imaging data available in these fields

ESO Public Spectroscopic Surveys (cont.)

- LEGA-C (PI: A. van der Wel) – UT3 VIMOS
  - Survey of 3000 early type galaxies in the COSMOS field in the z-range $0.6 < z < 1.0$ and with a K-band total magnitude limit running from $K_{AB} = 21.1$ at $z = 0.6$ to $K_{AB} = 20.4$ at $z = 1$. From UltraVISTA.
  - Understanding the gals growth through the measurement s of σ^* (dynamical masses) , stellar ages and abundances.
- Large Programmes!

Science Data Products

- ✓ In-house reprocessing
- ✓ Contributed products

Science Data Products

- Science ready data products through the Science Archive Facility to foster their quicker, wider use
- Two channels to feed the archive with data products
  - **Internal** automated processing with scientifically validated pipelines
    - **UVES echelle**, **X-Shooter echelle**, **HARPS echelle**, **FLAMES-MEDUSA**, then HAWK-I and VIMOS IMG (UK in-kind), MUSE, PIONIER, FEROS, …
    - Migration into Phase 3 of legacy historical Advanced Data Products completed
  - **External** Principal Investigators of Public Surveys, Large Programmes, … provide high-level products (mosaics, source catalogues, …) that we validate and integrate
    - All 11 Public Surveys have returned data products
    - Summary in December 2013 Messenger: papers from ESO, the Survey Teams and archive users

- Building high-quality, extensive content

http://www.eso.org/sci/observing/phase3/data_releases.html

****Legacy value: archival data may be useful for science exploitation independent for original scientific goal!**

In-house generation of science processed data

Extraction of science content

Data processing

Science users

Processed by the user

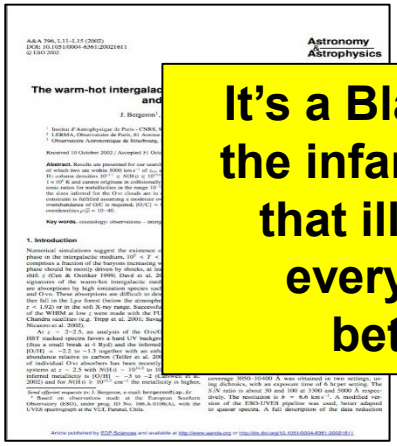
Processed by ESO & fed into the archive

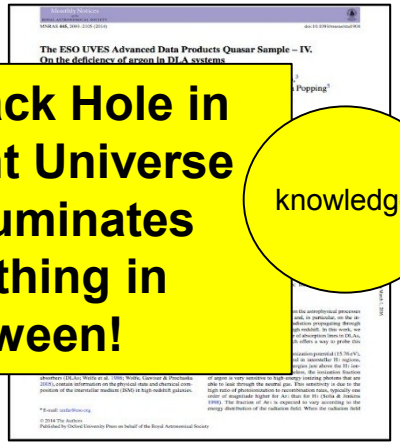
Processed by the community & fed into the archive

Raw data

PI science

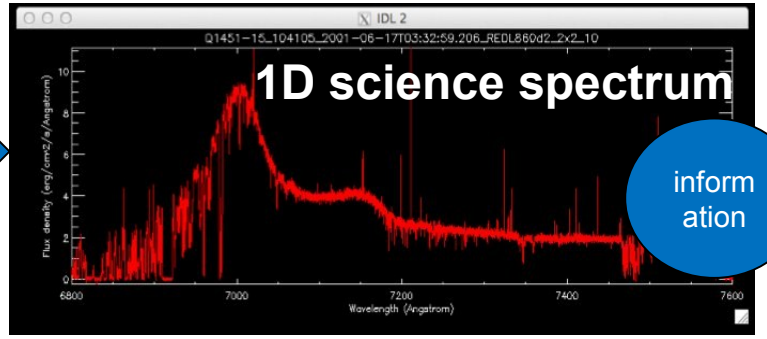
Archive science



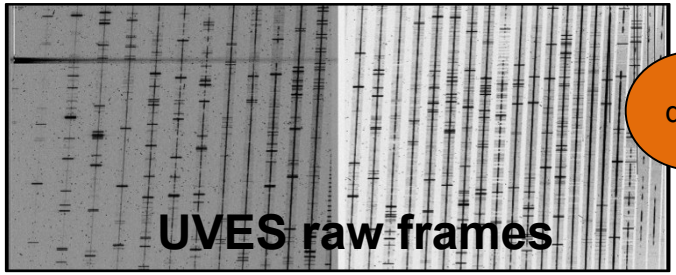


It's a Black Hole in the infant Universe that illuminates everything in between!

knowledge



information



data

In-house generation of science processed data

- QC group: Processing data at ESO & feed the science archive (since 2013)
  - UVES Echelle (~110,000 individual spectra, covering hundreds of science questions)
  - XSHOOTER Echelle (~40,000 individual spectra, from UV to NIR)
  - HARPS (~240,000 individual spectra)
  - FLAMES/GIRAFFE MOS (Medusa) (~1,300,000 individual spectra)
  - Coming soon: MUSE, HAWK-I, VIMOS, PIONIER, FEROS
- Requirements for quality-oriented processing of science data
  - Stable science operations
    - Calibration plan
    - On-site QC0
    - Instrument status well known (scores, Health Check monitor)
    - QC loop
  - Understood and certified data processing tools (pipelines)

Contributed processed science data

Extraction of science content

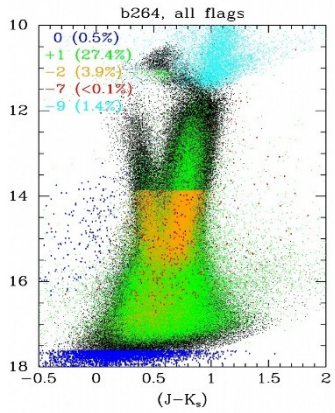
Science users

Data processing

Processed by the end user

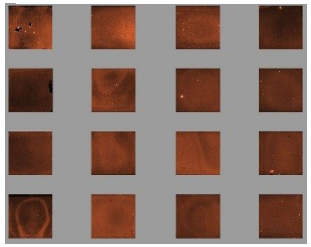
Processed at ESO to feed the science archive

Contributed by the community to feed the science archive



Source lists at several wavelengths

Raw data

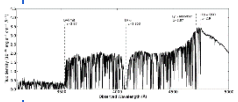


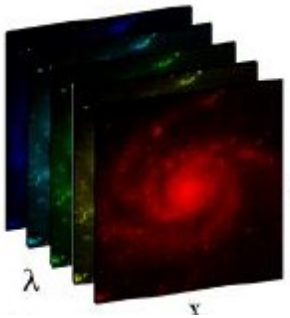
VIRCAM raw frame

Contributed processed science data

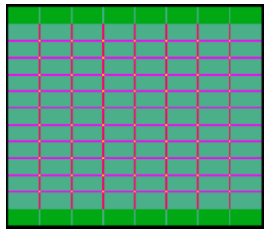
- Processed data contributed by the community to the science archive – routinely since 2011
  - Mandatory for datasets with high legacy value (Public Surveys and Large Programmes) - Arnaboldi+2014, Messenger,156, 24
    - Open to contributions from the community at large
- Goals of the quality control on the contributed data: build high quality, trusted content
  - Homogeneity and completeness of the published data
  - Coherent and comprehensive high quality description of the data
  - Ensure homogeneous, reliable user documentation of the data

ASG - Phase 3 science data standard





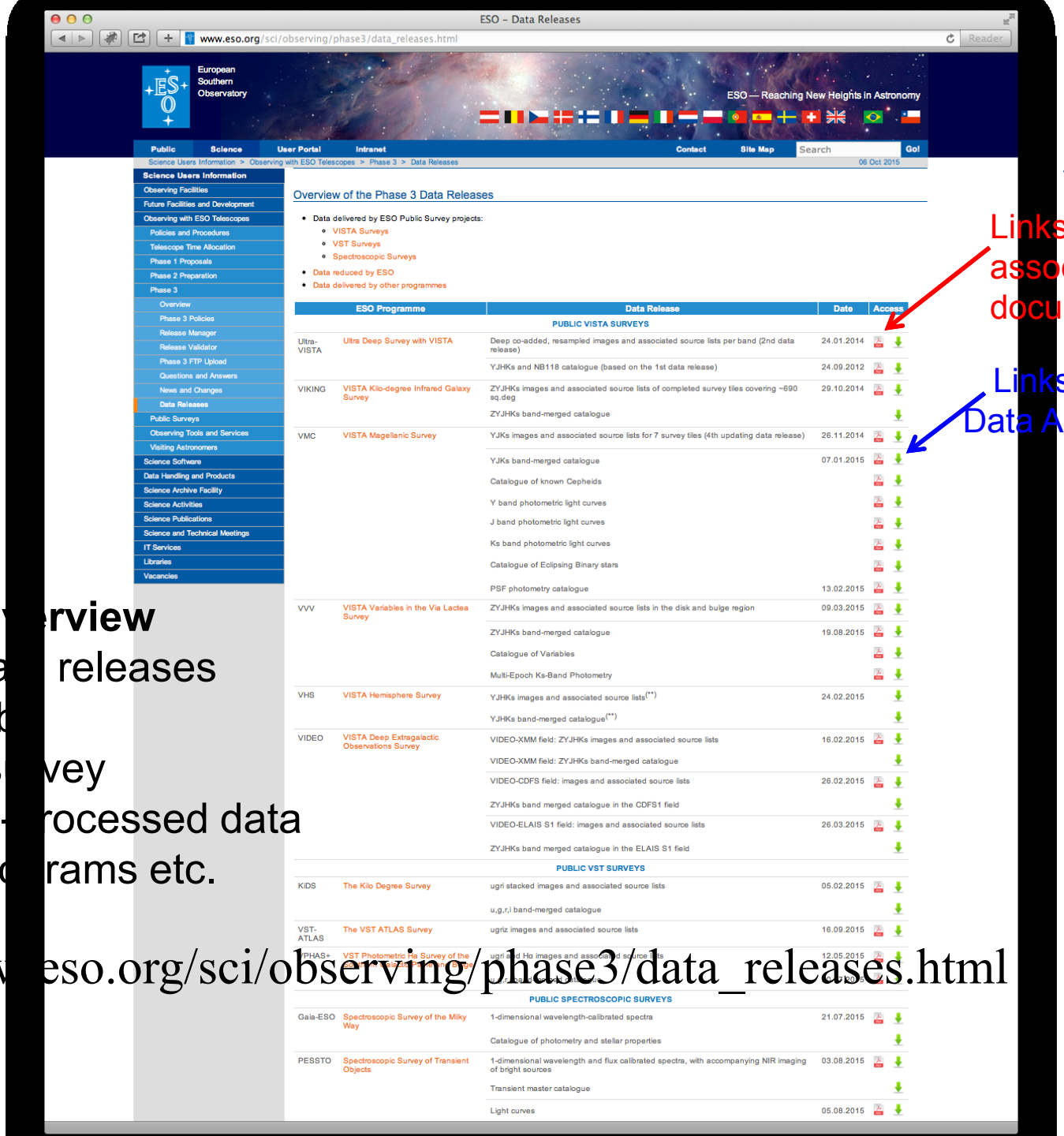
3D cube



VISTA survey tile.

Content validation by ASG

Retzlaff+2014, SPIE, 9149, 3



Links to the associated data documentation

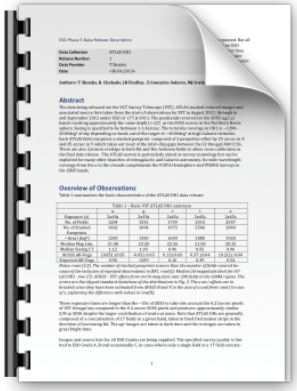
Links for Data Access

Release Overview

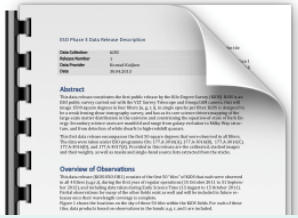
Phase 3 data releases organized by

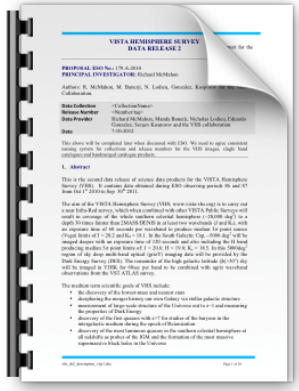
- Public survey
- pipeline-processed data
- large programs etc.

http://www.eso.org/sci/observing/phase3/data_releases.html



atlas_dr1.pdf





vhs_dr2.pdf

Data Release Description

- Provide short broad overview of the program, with an overview/layout of the observations

Essential input for data content validation.

Release content

- Extended listing for each sky position, filters, exposure times, seeing

Release notes

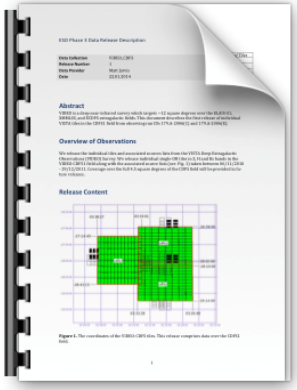
- Reduction method used, calibration procedures, data quality

Data format

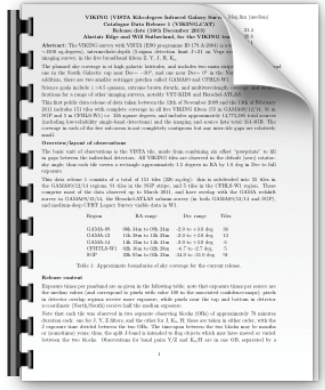
- Description of files in this data release , associated files, and naming conventions

Acknowledgements

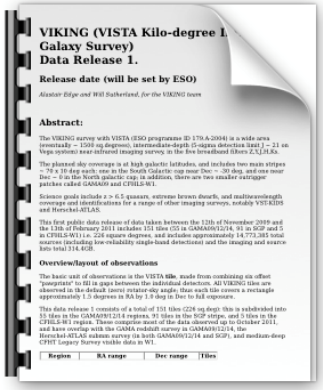
- Bibliographic reference to be included when using these data.



video_cdfs_dr1.pdf

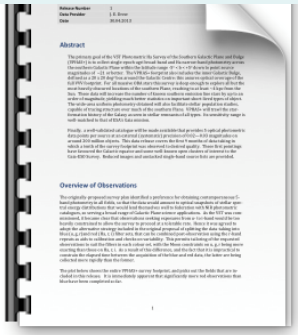


viking_cat_dr1.pdf

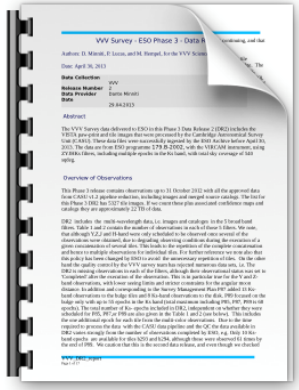


viking_dr1.pdf

vmc_dr2.pdf



vphasplus_dr1.pdf



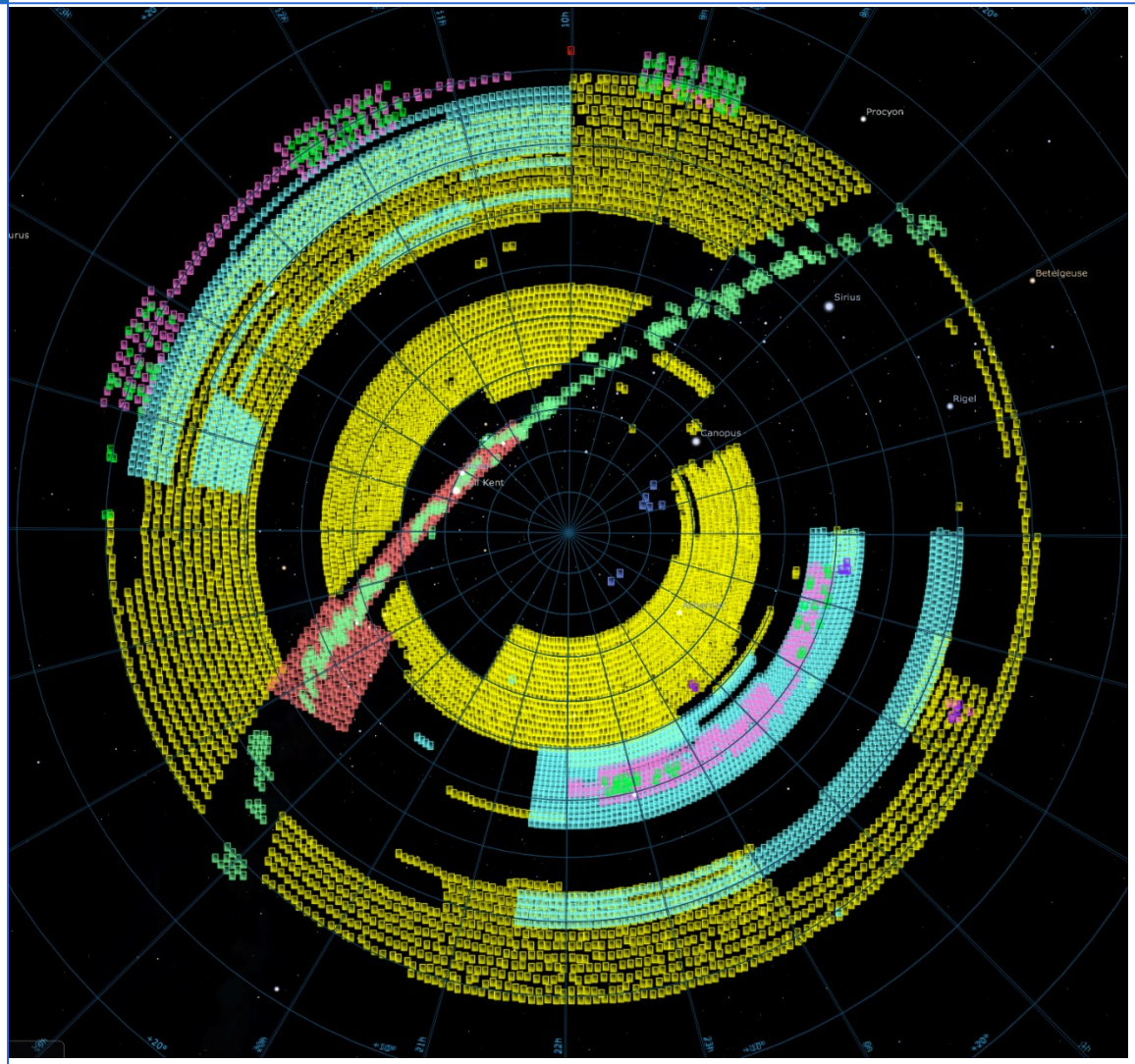
vvv_dr2.pdf

Phase 3 releases and query forms

Phase 3 releases from ESO Public Surveys and in-house reprocessing

- Phase 3 process completed for 11 Public
  - VISTA DR3, VST DR2
  - SPSs: GaiaESO, PESSTO DR2
- Sky coverage >11500 sq.deg
  - Opt./NIR: 4336 / 9445 sq.deg
- Tot. data volume: >35 TB
  - 270k+ files
  - >29k spectra
- High-level science catalogs for all surveys
  - Including light curves for variable sources!
- UVES, Xshooter, Harps, Giraffe, APEX (LABOCA)

Phase 3 releases from ESO Public Surveys

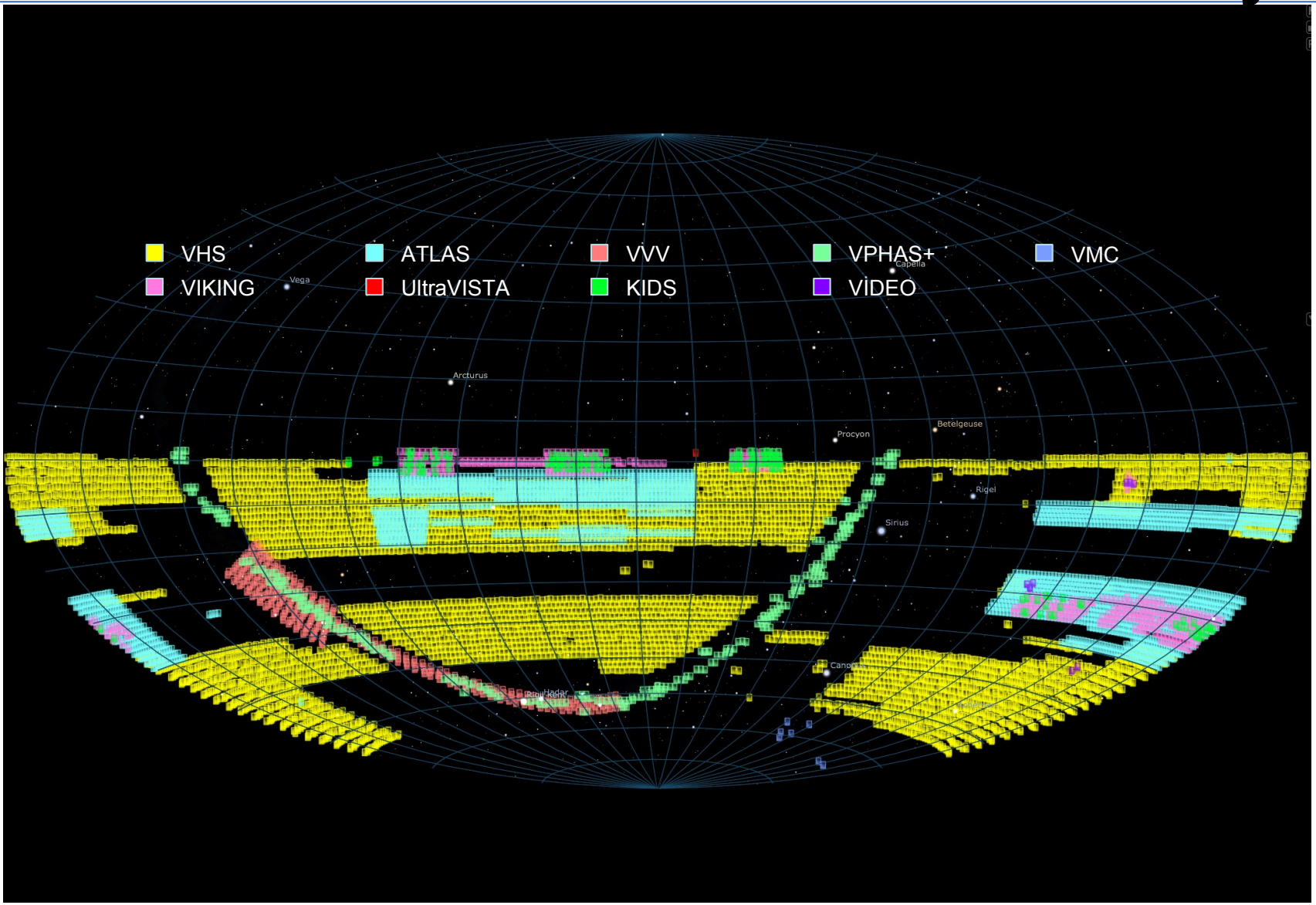


The VISTA & VST public survey DPs released through the Phase 3 process in 2014/2015 cover almost **11500 square degrees** of the Southern Hemisphere.

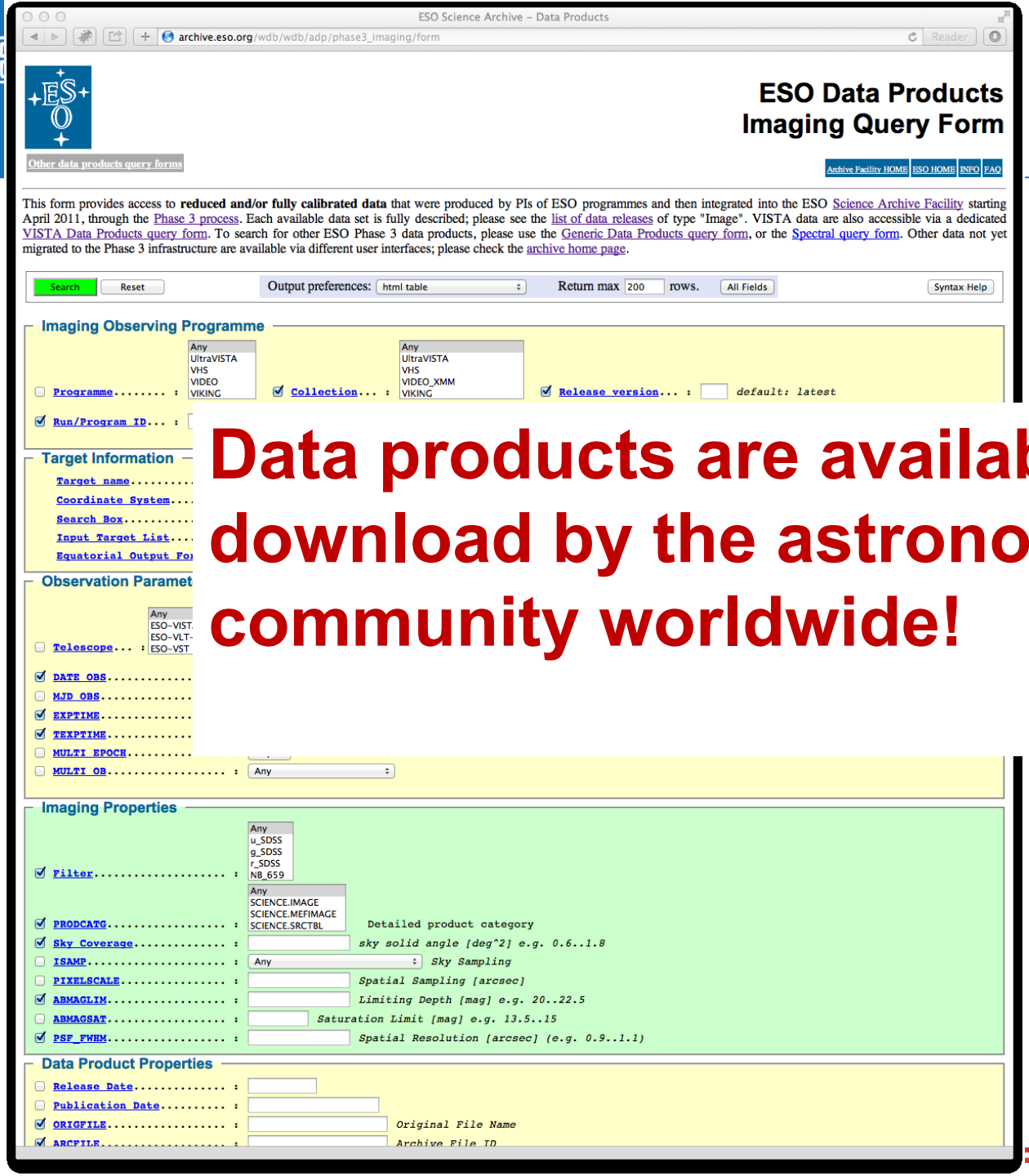
- VPHAS+
- VIDEO
- VVV
- VHS
- KiDS
- VIKING,
- VMC - blue
- UltraVISTA
- ATLAS

STATUS: more than 35 TB of SDPs have been ingested and published through the ESO SAF in 2014/2015

Phase 3 releases from ESO Public Surveys



ESO Archive Query Forms



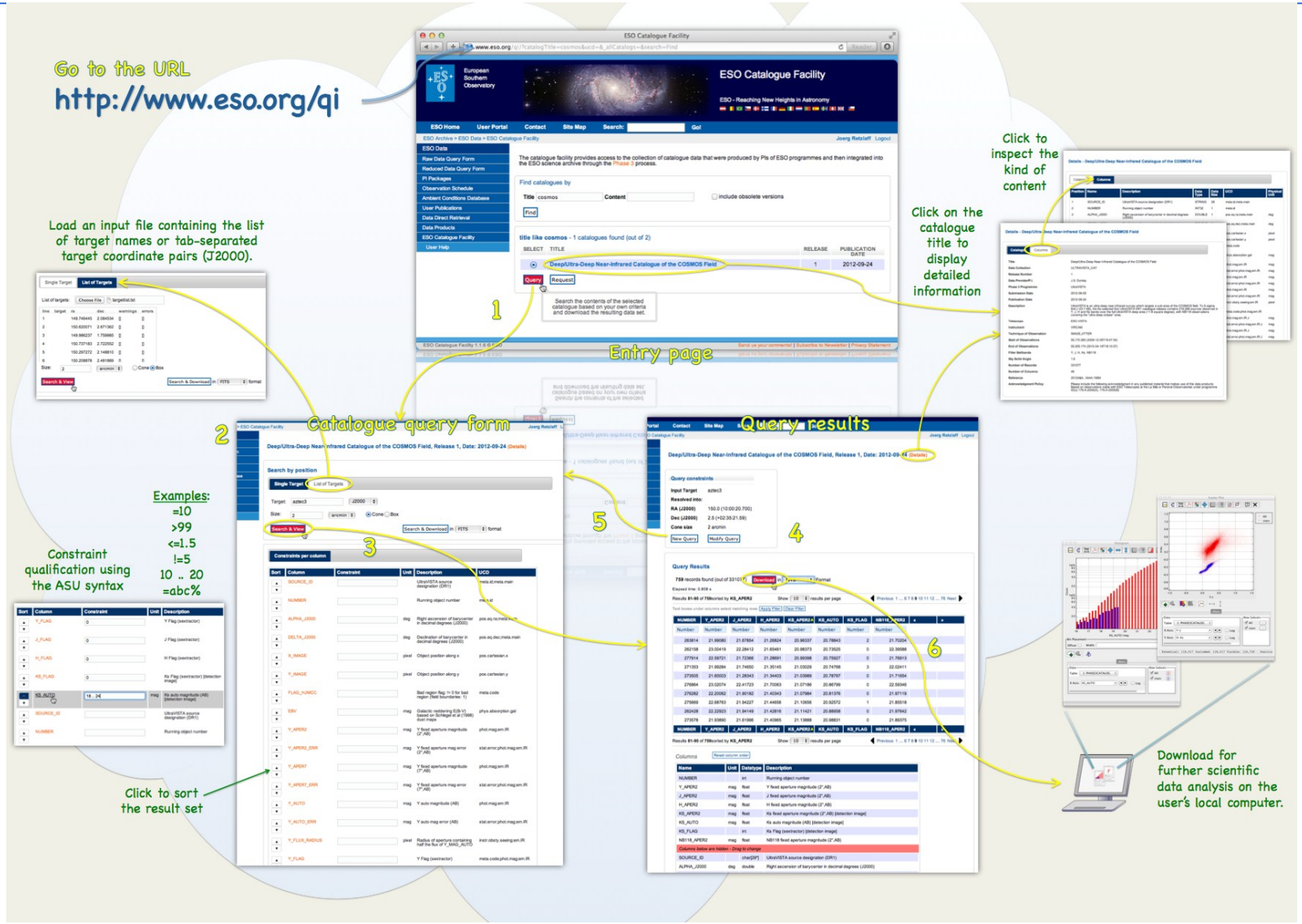
Data products are available for download by the astronomical community worldwide!

- Seamless access to all ESO data products via top-level query form;
- Including ESO-
  stream of
  ntly
  elle;
  chelle,
  ffe/Medusa.

by high-
ality
parameters:

- MAGLIM
- PSF_FWHM

Catalogue Facility Query Interface



Welcome to the ESO Science Archive Facility

The ESO Science Archive Facility contains data from ESO telescopes at La Silla Paranal Observatory, including the APEX submillimeter telescope on Llano de Chajnantor. In addition, the raw UKIDSS/WFCAM data obtained at the UK Infrared Telescope facility in Hawaii are available.

The Principal Investigators of successful proposals for time on ESO telescopes have exclusive access to their scientific data for the duration of a proprietary period, normally of one year, after which the data becomes available to the community at large. Please read the ESO Data Access Policy statement for more information, along with the relevant FAQs.

Browsing the archive does not require authentication, but to request and download data you have to log in to the ESO User Portal. Please acknowledge the use of archive data in any publication.

Latest News and Updates

- First release of PESSTO spectral data products (20 Jan 2014)
- DSS and Skycat catalogues services migrated (17 Jan 2014)
- New data of the VISTA Deep Extragalactic Observations Survey in the VIDEO-XMM field released (17 Jan 2014)

More news ...

To browse the archive

Currently, **raw data** and various types of **data products** can be reached via different interfaces:

Category	Query Forms	Data collection	Data Type	Instruments
LPO Raw Data	**Raw data** query form (all instruments) **Instrument specific** query forms **Direct** retrieval of raw data by file name	All ESO raw data	Various	Many La Silla Paranal instruments
LPO Data Products	**Phase 3 main** query form **Phase 3 imaging** query form **Phase 3 spectral** query form **Phase 3 VIRCAM-specific** query form	Phase 3 Data Products (ESO public surveys, large programs, pipeline products, etc.)	Currently, Imaging and Spectroscopy	Currently, VISTA/VIRCAM, VST/OmegaCAM, UVES pipeline products, etc.
	Catalogue Facility query interface	Phase 3 Catalogues [ESO User Portal authentication required also when browsing]	Catalogues	Currently, VISTA/VIRCAM, FEROS
	Advanced Data Products query form	GOODS (C.Cesarsky)	Imaging, Spectroscopy	FORS2/ISAAC/VIMOS
		zCOSMOS (S.Lilly)	Spectroscopy	VIMOS
		Observation of Corot astroseismologically-selected HD stars (E.Poretti)	Spectroscopy (time series)	FEROS
		Time-domain survey of NGC 2547 (S.Aigrain)	Imaging	WFI
	FEROS/HARPS pipeline processed data query form	FEROS/HARPS pipeline processed data	Spectroscopy	FEROS, HARPS
	Science Verification, Commissioning, EIS, etc.	Full list of available data packages	Various	Many
APEX Quick Look Products	**APEX** query form	APEX	Heterodyne, Bolometer	APEX-2A, LABOCA, SABOCA, SHeFI
LPO Schedule	**Scheduling** query form	ESO Observing Programme Information and Scheduling		All La Silla Paranal instruments, including APEX
ALMA Data	**ALMA Science Archive**	All ALMA data	Cube	ALMA

The ESO Science Archive Facility was developed in partnership with the Space Telescope – European Coordinating Facility (ST-ECF). It was operated jointly until the closure of the ST-ECF in December 2010.

archive.eso.org

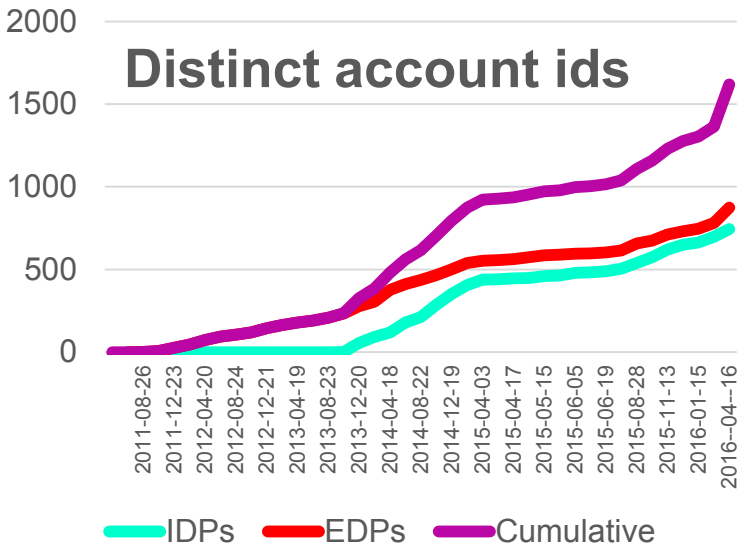
Building a community
Impact & statistics from ESO SAF
Highlights

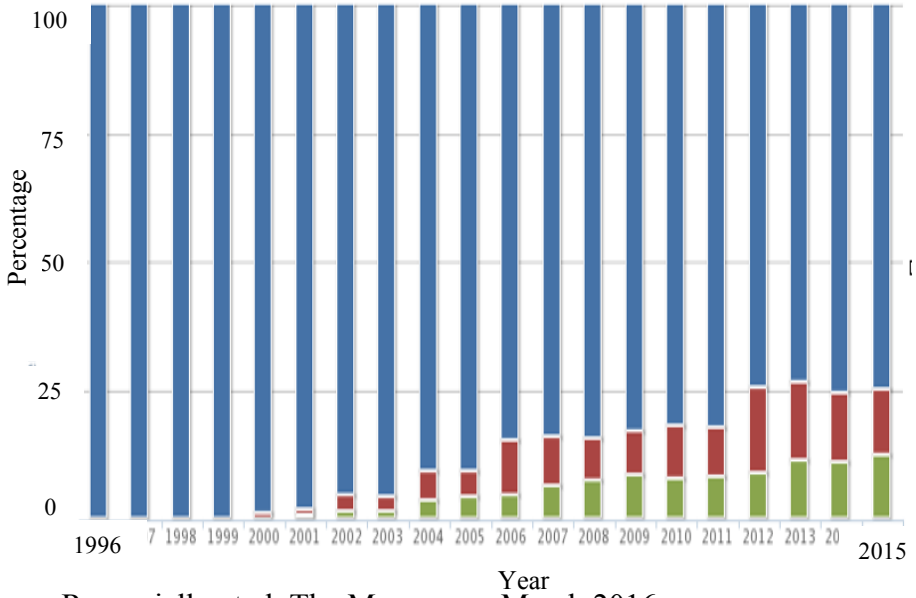
Neon School Garching, May 2016

Highlights: the impact of quality archived science data

- **Building a community**
  - ~1600 unique users of science ready data…and counting
    - ~7 requests per user
  - ~30% of them are new to ESO, having not applied for time

- **Refereed papers**
  - ~25% of ESO's output
  - ~5% of La Silla Paranal data only published as archive papers





Romaniello et al, The Messenger, March 2016

Highlights

Refereed publications from ESO Public Surveys

(ESO/Telbib census; Dec 2015)

- VHS 21, UltraVISTA 44, VVV 79, VIDEO 22, VIKING 17, VMC 22
- V-ATLAS 6, KIDs 9, Vphas+7
- PSS : Gaia ESO 32, PESSTO 24
- Archival : 45

Cumulative queries reduced products

- Number of distinct requests: 11302
- Number distinct account IDs:1620

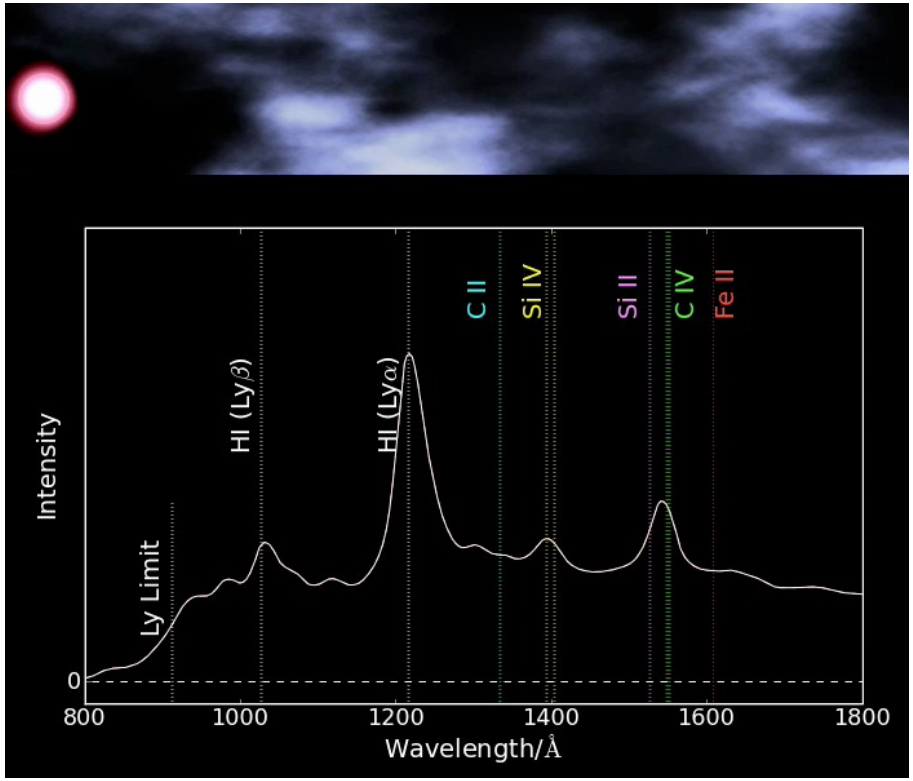
Implications of the above figures of merit

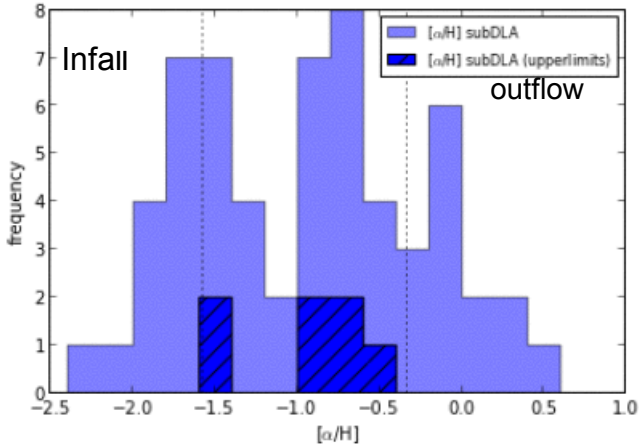
- Archive users query for reduced products more than once(~7 times, on average)
- The number of archive users using reduced products for their science is more than **twice** the number of listed PIs/CoIs of ESO public surveys' proposals

``Archive'' Surveys become possible….

ESO UVES Quasar sample – PI C. Peroux (LAM, Fr)

- 250 Quasar spectra, equivalent to 1500 hrs of VLT time
- 150 DLAs/sub-DLAs
- (Zafar et al. 2013a,b, 2014, 2015, ...)





- sub-DLAs at mean z~2
- Hint of bimodality infall/outflow
- Quiret et al. 2015 in prep

Archive science makes press releases!

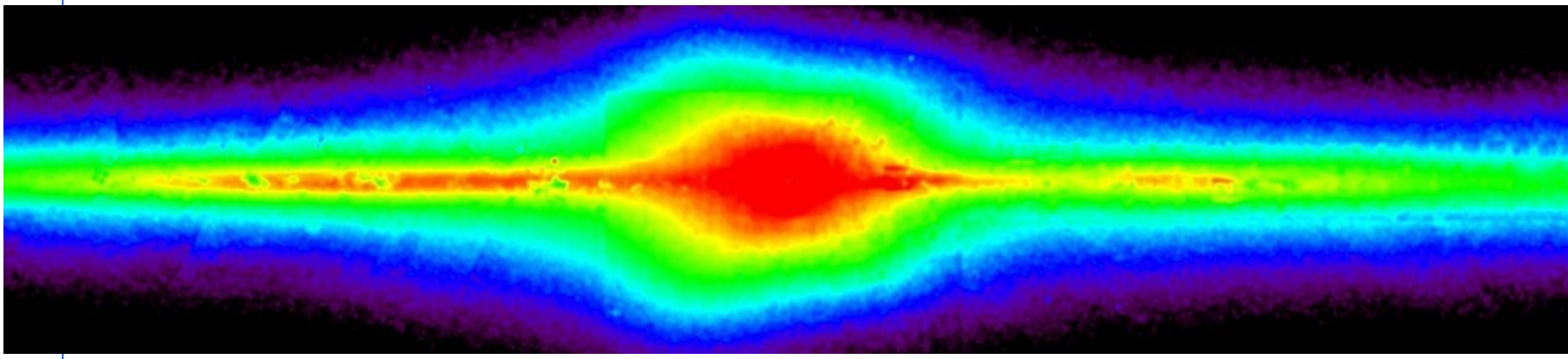
- ESO PR 1339 - 12 September 2013: **The Peanut at the Heart of our Galaxy**

C. Wegg, O. Gerhard from MPE. Used the VVV DR1 to build the 3D map of the galactic bulge. The inner region of our Galaxy has the shape of a peanut in its shell from the side, and of a highly elongated bar from above.

Wegg, C, Gerhard, O. 2013MNRAS.435.1874W

Archive science makes press releases!

The structure of the Milky Way outside the Bulge(*)



This image shows the surface density of stars in the Milky Way as seen from the Sun, taken from three different surveys (UKIDSS, VVV, 2MASS, and GLIMPSE) and corrected for extinction. The bulge is the thicker region near the center; it is asymmetric because it is barred. The asymmetry in the disk towards the left of the image is due to the thinner long bar outside the bulge. Wegg, Gerhard & Portail 2015, MNRAS, 450, 4050

* MPE PR http://www.mpe.mpg.de/6333402/News_20150521

End!

Many thanks to the PIs of ESO PS and their teams, the ESO colleagues and the BOD members for their hard work & support!