



ESO Phase 3 Operations & the 3D Data Standard

Jörg Retzlaff

Archive Science Group,
ESO Back-end Operations (DMO)



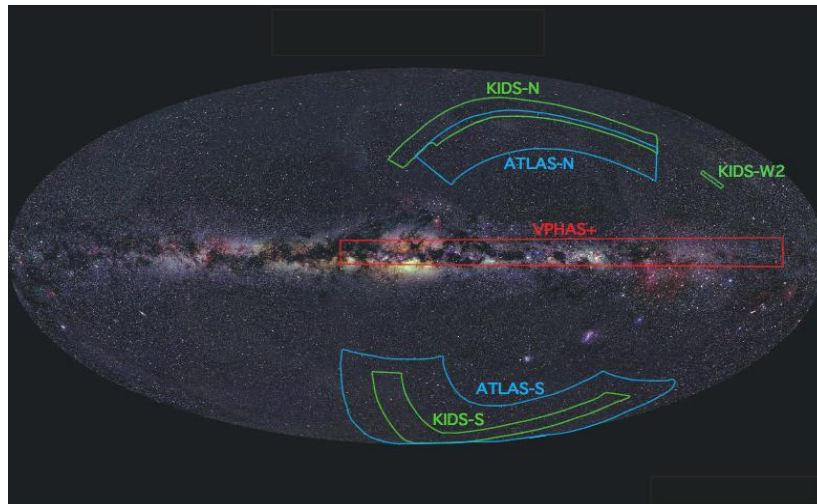
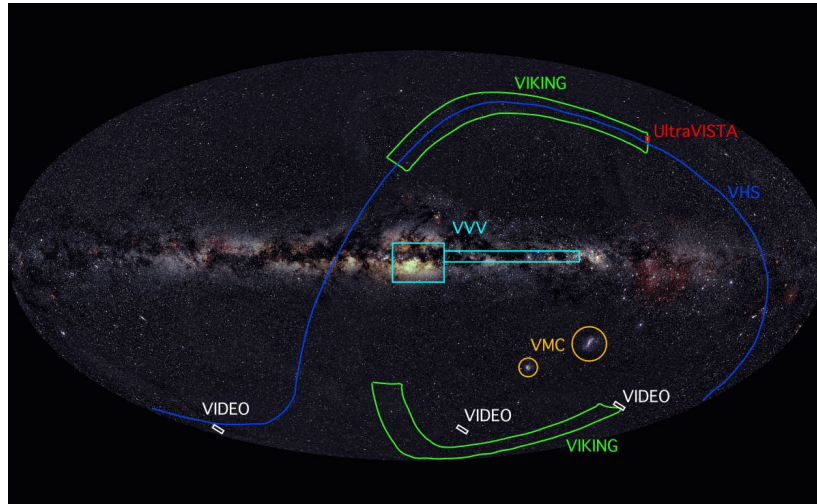
ESO PHASE 3 OPERATIONS



Publication of Science Data Products through the ESO Archive

- **Aim:** to facilitate the exploitation of ESO science data, in particular beyond the objectives of the observing programme for which the data were originally acquired.
- Reduced data is being generated by:
 - PIs of ESO Public Surveys, large programmes, etc. who return reduced science data to ESO (often in collaboration with specialized data centres)
 - ESO processing the stream (including the archive history) of observational data for selected instruments and modes on a regular basis.
- ESO/ASG makes the data **accessible** to the community at large while guaranteeing long-term availability.

ESO Public Surveys



VISTA: 6 surveys started Apr 2010 (P85)

- VHS - 20000 deg² YJHKs (Ks<20 AB)
- VIKING - 1500 deg² ZYJHKs (Ks<21.2 AB)
- VIDEO - 3 Deep Extragalactic Fields
- Ultra-VISTA - Ultra-deep ZYJHKs + NB118 in the COSMOS field
- VVV - Variability study of 520 deg² in bulge+plane plus multi-color map
- VMC - Magellanic Survey

VLT Survey Telescope: 3 surveys, started 15 Oct 2011 (P88)

- VST-Atlas – 4500 deg² UVRIZ, like SDSS
- KIDS – 1500 deg² UVRI, 2.5 mag deeper than SDSS
- VPHAS+ 1800 deg² UVHaRI in the Southern Galactic Plane

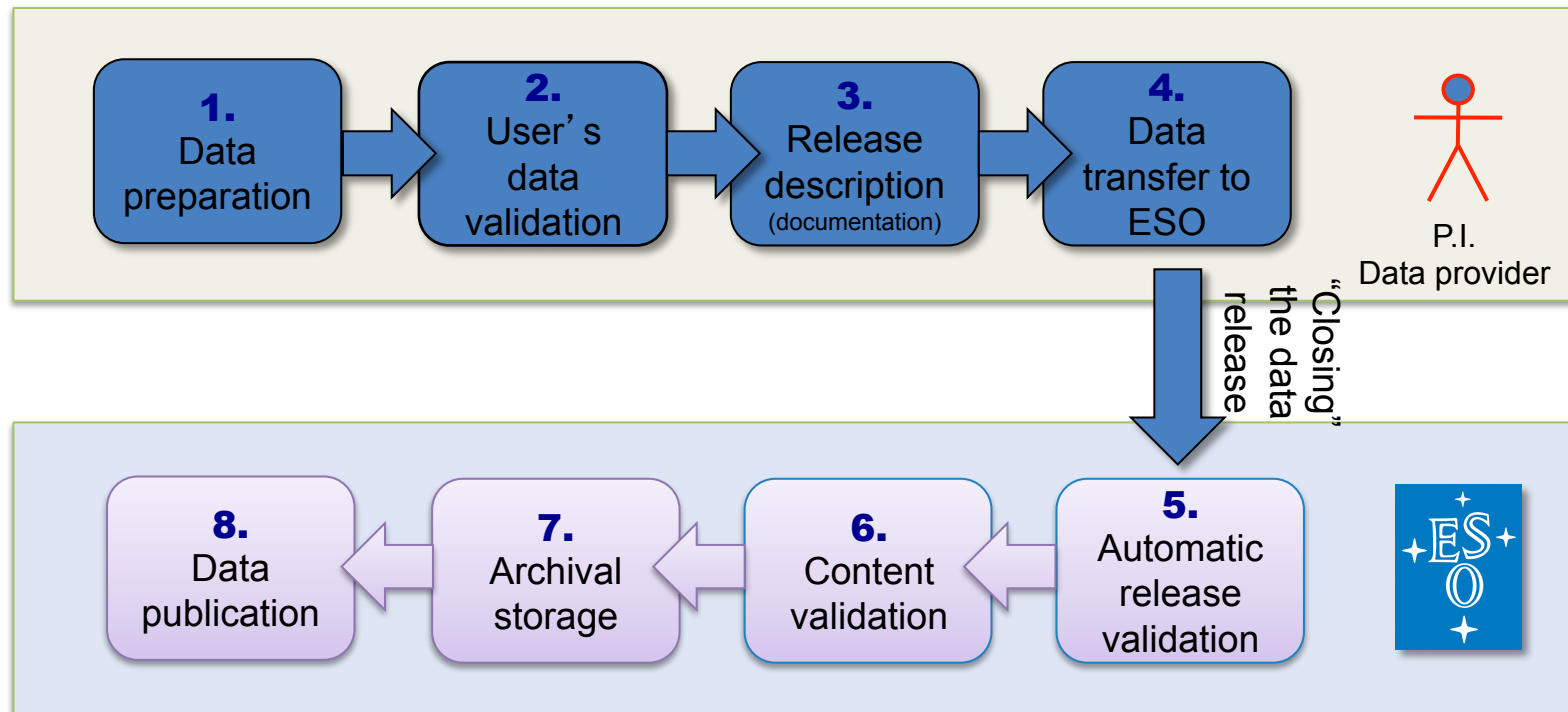
Spectroscopic Surveys: started 1 Jan 2012

- PESSTO: 30+60n on NTT (SOFI+EFOSC)
- Gaia-ESO: 30+30n on VLT-UT2 (FLAMES)
- LEGA-C + VANDELS @VIMOS (as of P94)



Phase 3 in a Nutshell

Phase 3 denotes the process of preparation, submission, validation and ingestion of science data products in the ESO Science Archive Facility.

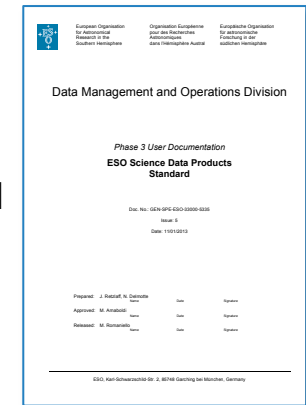




Data Preparation

ESO Science Data Product Standard

- Compliance with data format standard required:
 - Common keywords enable queries across instruments and data types independent of the details of data processing.
 - Uniform structure of data for each type (image, spectrum, IFU cube, catalog) and standard keywords facilitate scientific exploitation.
 - Data interface for reduced science data based on established standards whenever possible (FITS, IVOA, OIFITS).
- Variety of data product types covered:
 - Imaging (astrometrically/photometrically calibrated)
 - catalogs of sources per image (“source tables”)
 - 1-d extracted calibrated spectra (inspired by IVOA Spectral DM, version 1)
 - sub-mm flux maps (e.g. APEX/LABOCA)
 - IFU data cubes incl. error and quality information (=> MUSE)
 - Catalogs (high-level): photometry, light curves, spectroscopic surveys
 - Soon: Optical interferometric data
- Association scheme for ancillary data with science data products
 - Quality and error information (e.g. weightmaps), previews, 2d spectral frames



-
- The image is a composite of three screenshots illustrating the Phase 3 Release Manager workflow:
- Top Left:** A screenshot of the "Phase 3 Release Manager" web interface. It shows a "Progress" bar and a table of release data. The table has columns for "File", "Size", and "Status". The status column shows various release types like "RELEASE", "RELEASED", and "RELEASED".
 - Top Right:** A terminal window showing the output of the command `phase3ReleaseManager -info -list -list`. The output lists the release status for various files, including "RELEASE", "RELEASED", and "RELEASED".
 - Bottom Left:** A screenshot of the "Phase 3 Release Manager" web interface, showing a table of release data. The table has columns for "File", "Size", and "Status". The status column shows various release types like "RELEASE", "RELEASED", and "RELEASED".
 - Bottom Right:** A screenshot of a file explorer window showing the contents of a directory. The files are listed in a table with columns for "Name", "Size", and "Type". The files are named with the release ID and the file type, such as "RELEASE", "RELEASED", and "RELEASED".
- A large yellow arrow points from the "Phase 3 Release Manager" web interface to the file explorer window, indicating the flow of data from the web interface to the local file system.
- Phase 3 Release Manager**
<http://www.eso.org/rm>
- Phase 3 FTP server**



ESO - Phase 3

www.eso.org/sci/observing/phase3.html

European Southern Observatory

ESO — Reaching New Heights in Astronomy

Public Science User Portal Intranet Contact Site Map Search Go!

Science Users Information > Observing with ESO Telescopes > Phase 3

06 Oct 2015

Phase 3

16 September 2015

- New release of stacked reduced images and source lists for the VST Public Survey ATLAS [Read More]

21 August 2015

- New Data Release of VVV Photometric Catalogues [Read More]

6 August 2015

- New Release of PESSTO Spectroscopic Public Survey Data [Read More]

[More News...]

In a nutshell, Phase 3 denotes the process of preparation, validation and ingestion of science data products (SDPs) for storage in the ESO science archive facility, and subsequent data publication to the scientific community. SDPs are produced by 1) principal investigators of ESO observing programmes, and 2) ESO pipelines as part of the quality control (QC) process or from specific, dedicated, re-processing projects for homogeneous raw data sets.

ESO's policies governing Phase 3 are specific to the type of observing programme. Phase 3 is mandatory for ESO Public Surveys and for ESO Large Programmes since period 75. For other ESO programmes there is no obligation but PIs are invited to take advantage of the Phase 3.

To ensure the successful integration of SDPs into the archive, ESO supports the users in carrying out the Phase 3 process by defining ESO/SDP data standards, by devising procedures and providing the infrastructure for the delivery of SDPs, and by supplying tools for the data preparation.

The description of the policies, the data standard and the procedure for the submission of reduced data products applicable to the ESO Phase 3 process given here is intended to provide the information for the preparation and successful completion of the ESO Phase 3 process.

The target audience consists of 1) principal investigators and their collaborators who return reduced data products resulting from ESO observations for public release to the astronomical community through the ESO archive, 2) ESO scientists involved in the QC process or in specific re-processing projects, 3) instruments scientists and pipeline developers for the new and existing ESO instruments.

Instructions and Documentation

- Overview of the Phase 3 process
- Phase 3 Policies for ESO Public Surveys
- ESO Science Data Products standard [PDF]
- Phase 3 User Guide to the Data Submission [PDF]

Phase 3 Infrastructure and Software Tools

- Phase 3 Release Manager
- Release Validator
- FTP upload (phase3ftp.eso.org)

Further Information

- Questions & Answers
- News and Changes

In March 2011, Phase 3 operations have started to support the validation, transfer and publication of the first data products from ESO public surveys. Other ESO observing programmes will be handled through the Phase 3 system as soon as the underlying data products standard and infrastructure have been extended to cover the respective instruments and data product types.

Contact the Phase 3 Helpdesk

For any questions regarding Phase 3, its policies, the data content and format, or the submission process of data products to the ESO Archive, please feel free to contact the ESO Archive Science Group via email at

usd-help@eso.org, subject: Phase 3

Phase 3 Quick Links

- Phase 3 Main Level
- ESO SDP Standard (v5) [PDF]
- APEX Sub-mm Sky Maps [PDF]
- Integral Field Spectroscopy: 3D Data Cubes [PDF]
- Phase 3 User Guide [PDF]
- Phase 3 Release Manager
- Release Validator
- Phase 3 FTP upload
- Get template for the data release description
- Contact the Phase 3 Helpdesk

Overview of data published via Phase 3

How to prepare and submit the data..

Get the Science Data Product Standard (data format & required keywords):

1. Images+Spectra (main doc.)
2. APEX sky maps (appendix)
3. IFU Data cubes

Phase 3 user support



Survey data releases 2015

Date	Survey	Data
Nov 2015	VST-Atlas	Band-merged ugriz catalog (87 Mio. objects)
Sep 2015	VST-Atlas	Stacked ugriz images and source lists. 2000-3000 sq.deg. observations until Sep 2013. 5 TB.
Aug 2015	VVV	ZYJHKs photometric catalogues. 418 Mio. sources. 2.5 Billion Ks photom. points. 13 Mio. variables. Including observations until Sep 2011.
Aug 2015	PESSTO	1728 EFOSC2 and SOFI spectra and 392 reduced SOFI images. Associated catalogs.
Jul 2015	VPHAS+	Band-merged catalog based on DR2 imaging includes observations until Sep 2013.
Jul 2015	Gaia-ESO	27359 spectra of point-like sources obtained with FLAMES until Dec 2013.
May 2015	VPHAS+	Images and single-band source lists for data acquired until Sep 2013.
Mar 2015	VIDEO	ELAIS-S1 field
Mar 2015	VVV	Images and source lists including observations until Sep 2013. Ks variability survey.
Feb 2015	VIDEO	CDFS & XMM-LSS field
Feb 2015	VMC	PSF photometry
Feb 2015	VHS	14 TB imaging data including observations until Sep 2013 + band-merged catalog of 590 Mio sources
Feb 2015	KIDS	Images and catalogues including data acquired until Sep 2013. 148 sq.deg. source catalog 17 Mio. sources.
Jan 2015	VMC	YJKs photometric catalogue from 7 completed tiles (10.5 sq.deg) + multi-epoch photometric data + catalogs of Cepheids and Eclipsing binaries.
Oct 2014	VIKING	Images and catalogues including data until Sep 2013. 690 sq.deg + 46 Mio. sources catalog with band-merged photometry.



Published Data Products

Public Survey Data Releases

Reduced data from 11 ESO public surveys

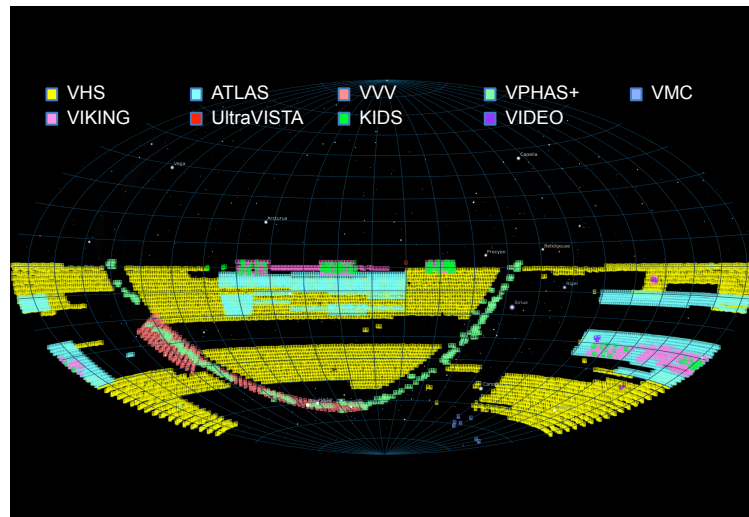
VISTA DR3, VST DR2, GaiaESO, PESSTO DR2

Sky coverage total >11500 sq.deg

Opt./NIR: 4336 / 9445 sq.deg

Tot. data volume: >35 TB

270k+ files, >29k spectra



ASTERICS DADI ESFRI Forum, Trieste, 03/12/15

Science data processed by ESO

	No. of spectra	GB
UVES-Echelle	107,469	280
X-Shooter	47,919	878
HARPS	237,062	2,362
Giraffe-Medusa	1,142,908	5,308

26 Nov 2015

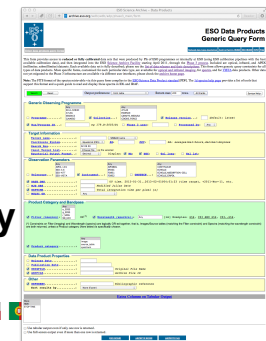
Data products become publicly accessible once the proprietary period (usually one year) has expired;

before that date, access to the data products is restricted to the P.I. of the observing programme and her/his collaborators.

Serves both PI's and archival research interests

<http://archive.eso.org>

The Science Archive Facility provides seamless access to all data products whether provided by PIs or processed by ESO.





Catalogue Facility Query Interface

Go to the URL
<http://www.eso.org/qi>

Load an input file containing the list of target names or tab-separated target coordinate pairs (J2000).

Click to inspect the kind of content

Click on the catalogue title to display detailed information

Entry page

Catalogue query form

Query results

Examples:
=10
>99
<1.5
!=5
10..20
=abc%

Constraint qualification using the ASU syntax

Click to sort the result set

Download for further scientific data analysis on the user's local computer.

- Dedicated user interface providing access to high-level catalog data for ESO public surveys and other programmes.
- Full catalog data search supported.
- Currently 20+ catalogs, >3 billion records, with variety of data (photometry, variables, redshifts etc.)
- Unique links between catalog records and spectra.
- Collaboration with CDS/Vizier to provide additional services like aperture-matched multi-wavelength catalogs (eg. VHS-DES)



Lessons learnt

- Validation is required, compliance of submitted Phase 3 (meta)data with the data interface ('standard') must be checked.
- Processing provenance: keeping track of the link between products (reduced data) and raw data is essential.
 - To implement access control for data under proprietary protection (usually for 1 year)
 - To allow consistency checks of the products MD (time, position, energy etc.)
 - Also to implement high-level archive services like overlaying image and detected sources.
- Version management is inevitable
 - Reprocessing takes place but not always/completely; challenging for very large releases
 - Requirement: focus on the most recent ("best") version of processing.



Phase 3 Upgrade

■ Main Goals

- Minimizing the overall time to data publication
 - Optimize the feedback loop with the data provider regarding the data validation status
 - and uploaded data content
 - Reducing workload for data providers as much as possible
- Implementation of missing operational requirements
 - Better support for very large data releases using submission via multiple 'batches'
 - New Survey manager role
 - Improved communication and tracking of data-related issues between data providers and ESO ("Phase 3 RIX'es mgt.")

■ Released planned for 2016Q1.



ARCHIVE SCIENCE AND DATA ACCESS

Example: VVV ESO public survey

eso1339 — Science Release

SPACE SCOOP

The Peanut at the Heart of our Galaxy

ESO telescopes create the best 3D map yet of central bulge of the Milky Way

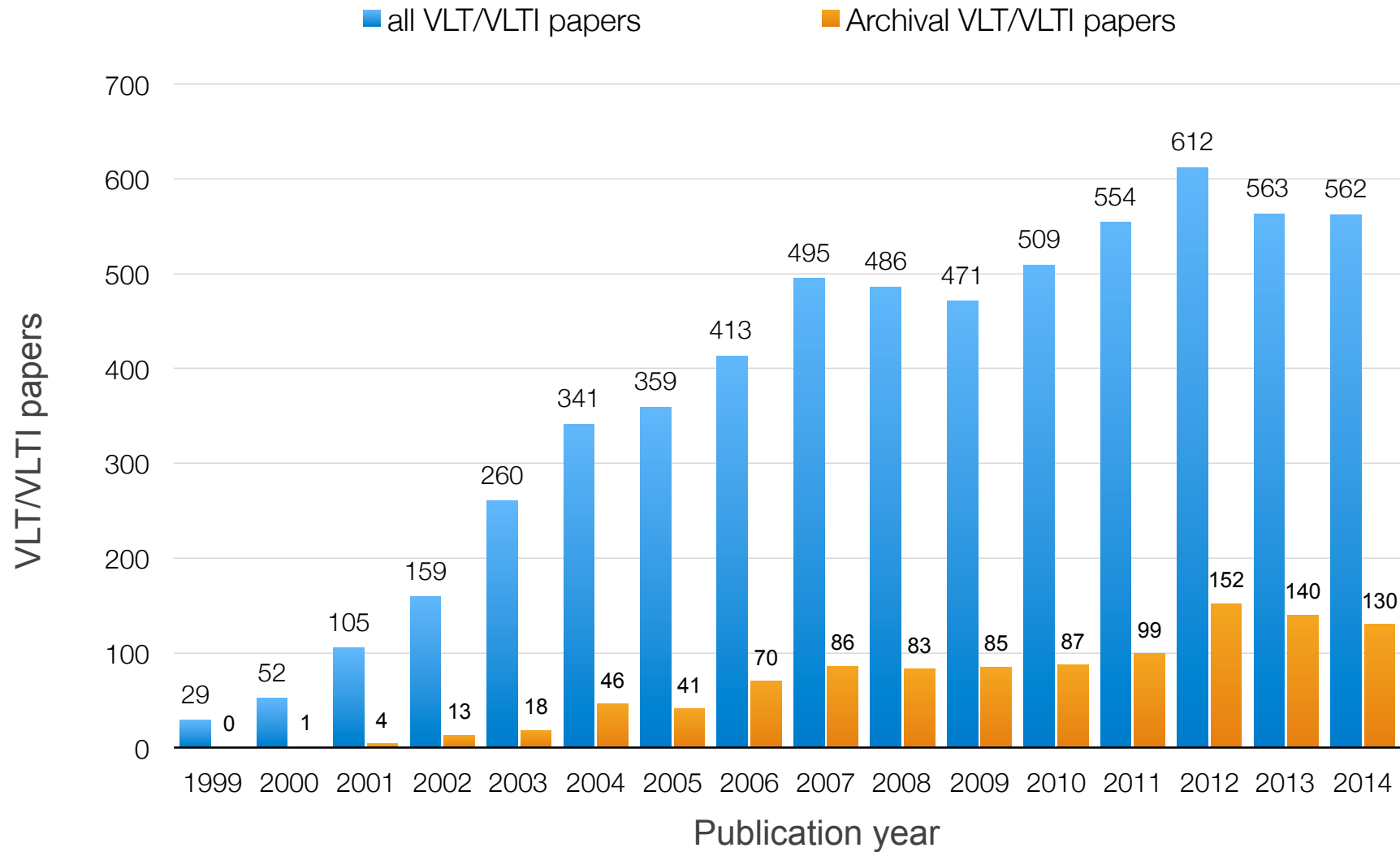
12 September 2013



Wegg & Gerhard 2013, archival data only
also: Wegg, Gerhard, Portail 2015 (arXiv:1504.01401)



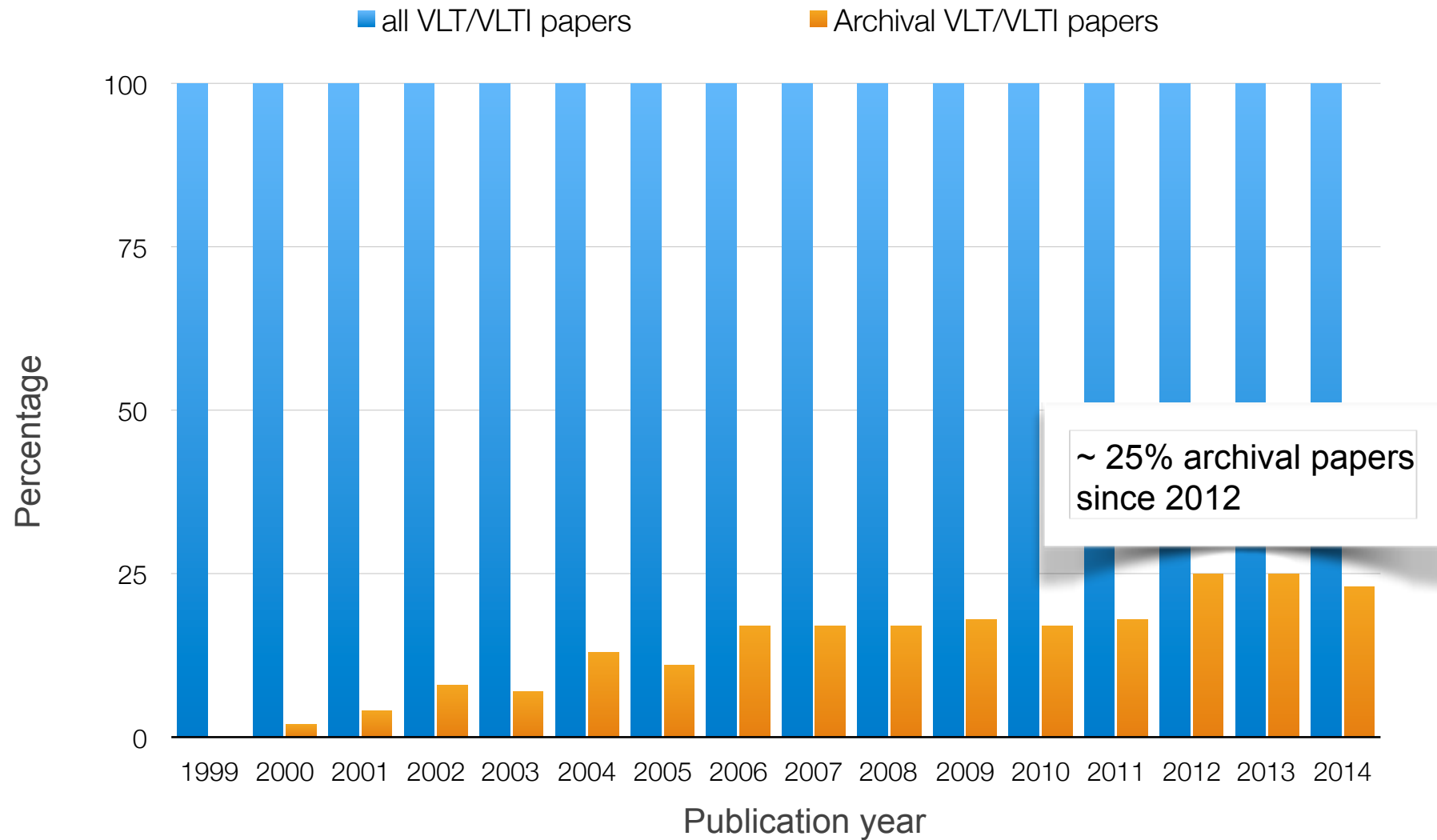
Publications based on ESO Data



Courtesy: U. Grothkopf/ESO Library



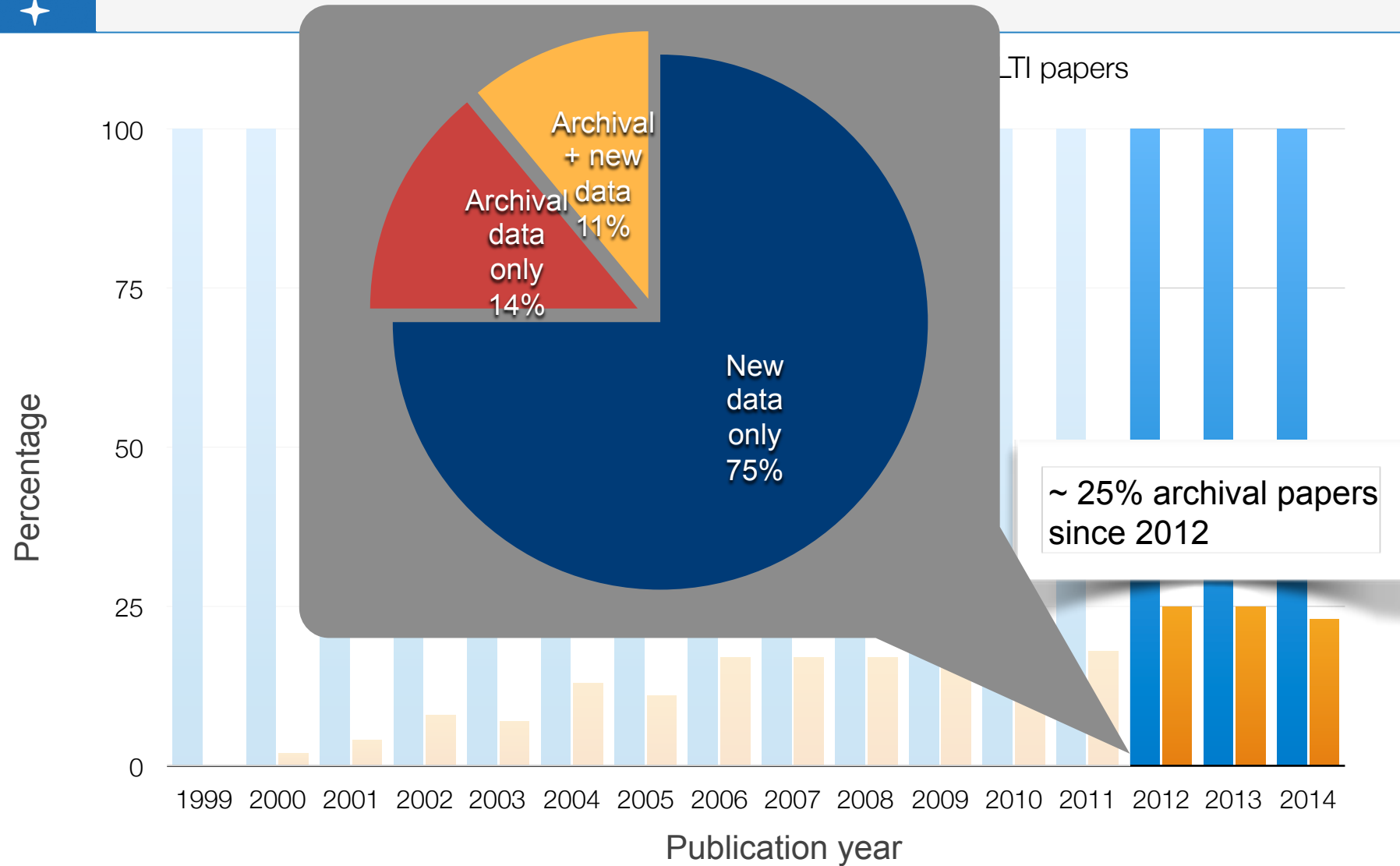
Archival papers - normalized



Courtesy: U. Grothkopf/ESO Library



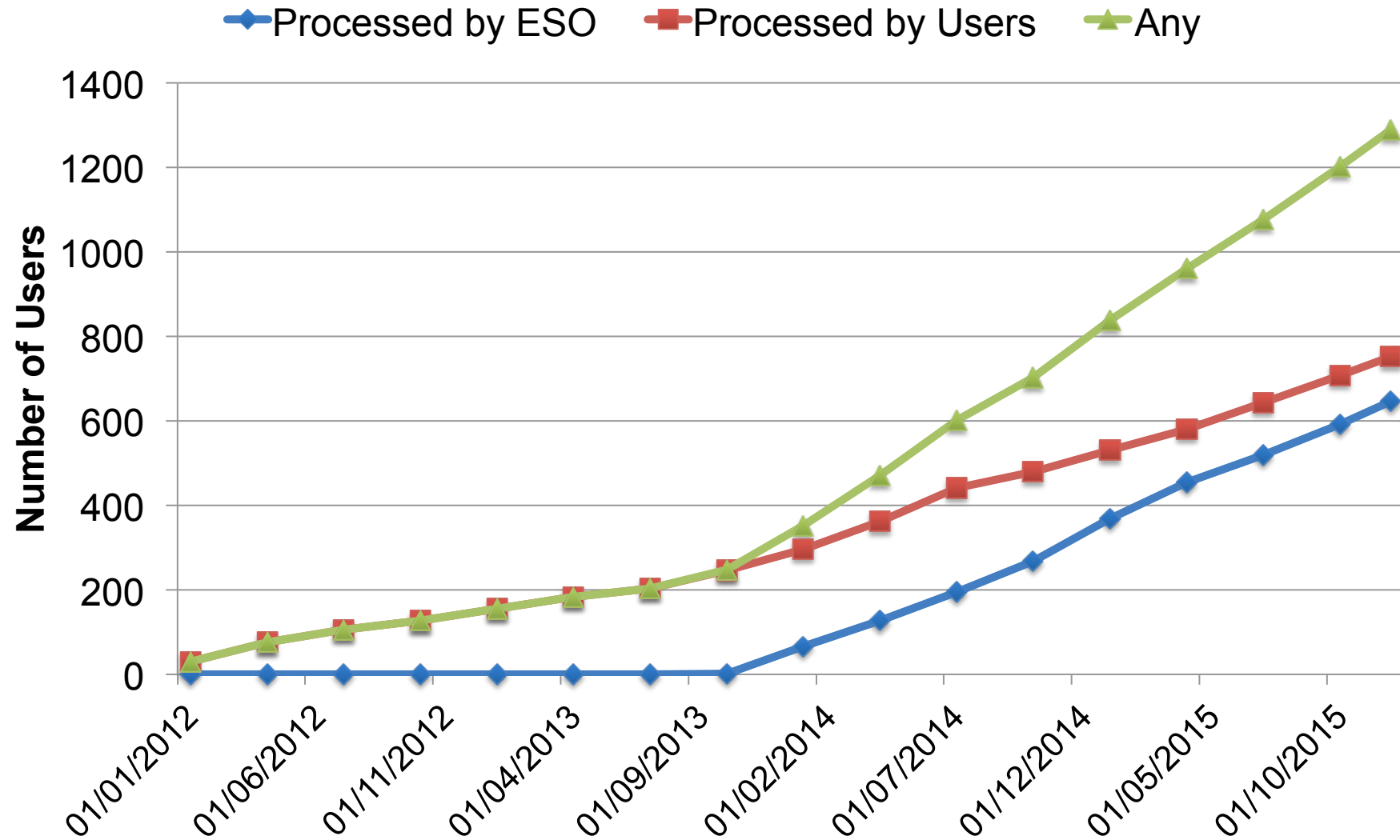
Archival papers - normalized



Courtesy: U. Grothkopf/ESO Library

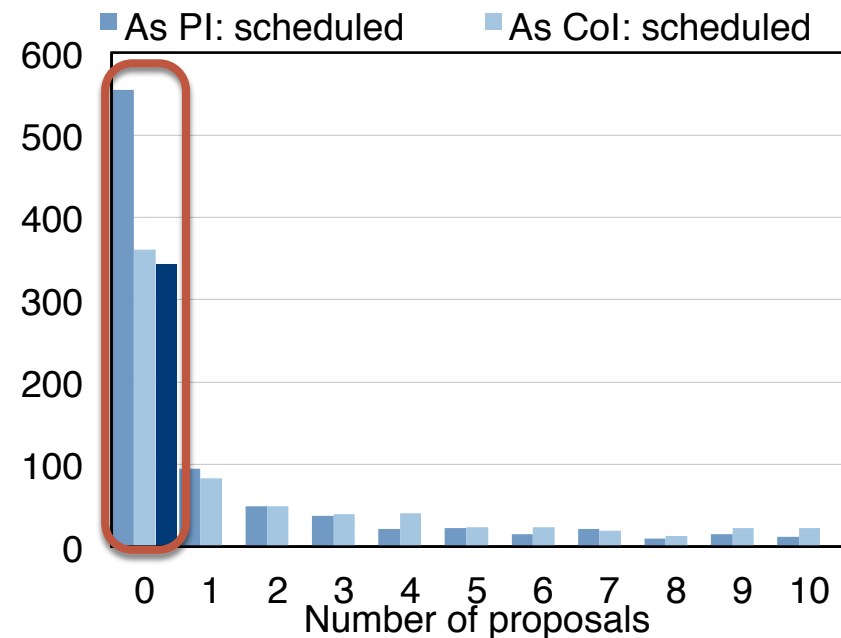
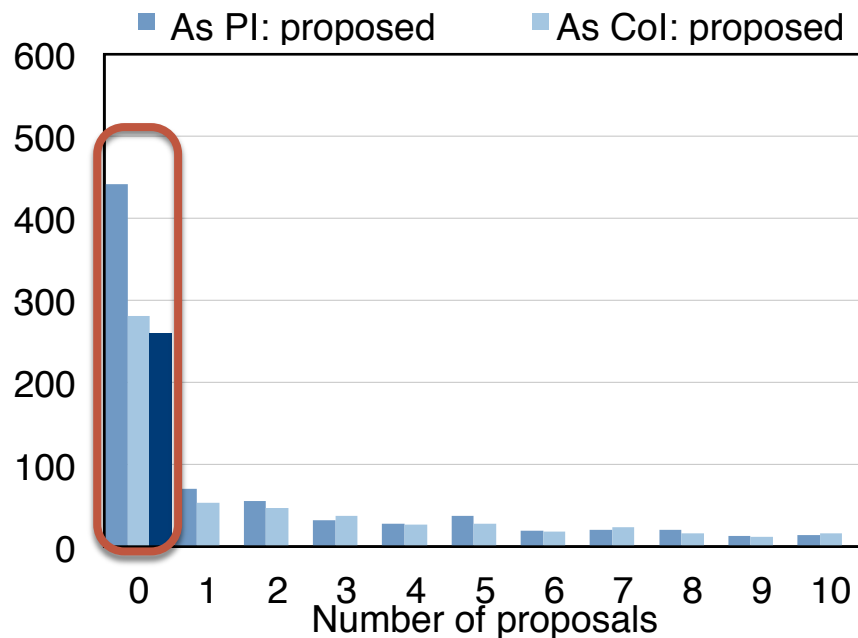


SAF Data Product Downloads



Phase 1 habits of archive users

Data products



Source: ESO Back-end Operations

- SAF data products users: 28% (260) have never applied for time, neither as PIs nor co-Is; 37% (344) have never gotten time, neither as PIs nor co-Is
 - For comparison, 1/3 of the Phase 1 PIs have never got time

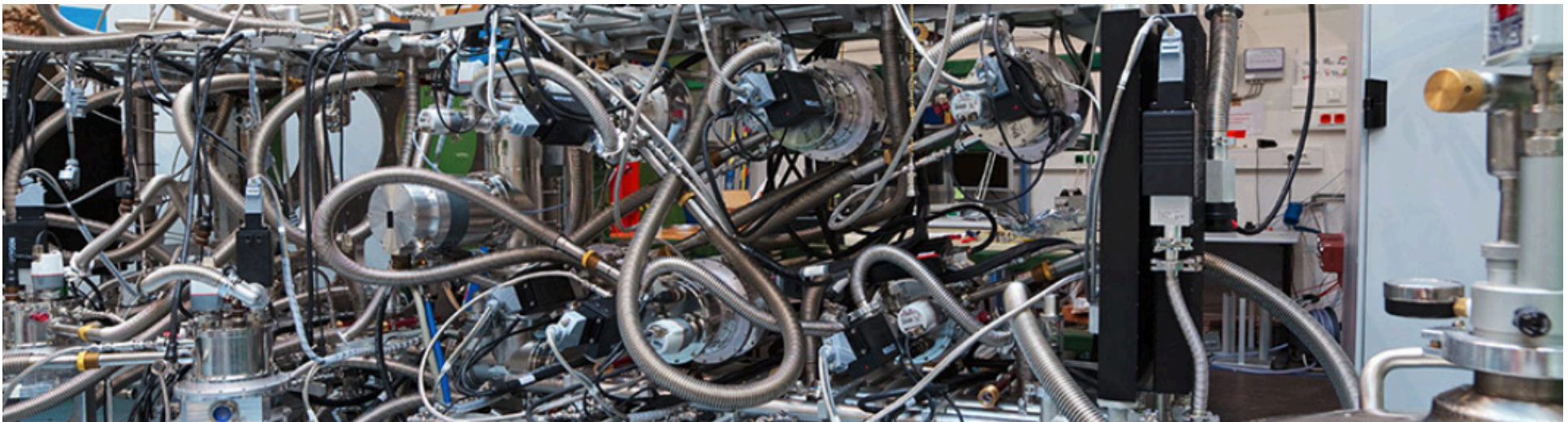


THE 3D DATA STANDARD

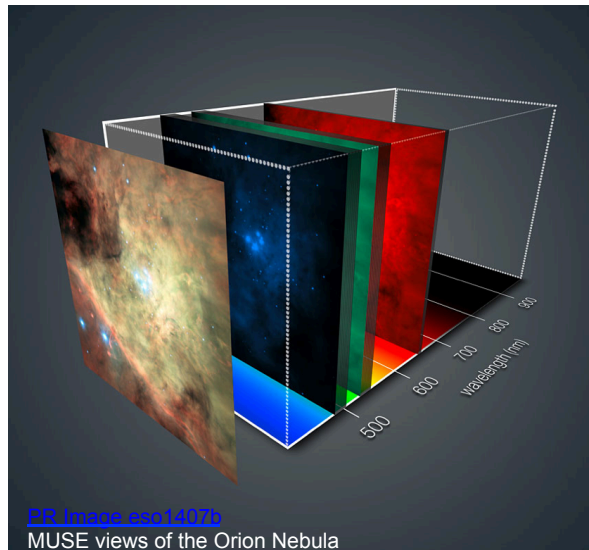


Multi Unit Spectroscopic Explorer

- Integral field spectrograph mounted on the Nasmyth focus of UT4
- Belongs to the suite of VLT 2nd generation instruments
- Wavelength coverage: 465 – 930 nm
- Spatial resolution: Seeing limited, with 0.2-arcsecond pixels. With the GALACSI Adaptive Optics system: diffraction limited with 0.025-arcsecond pixels.
- Spectral resolution: From 1700 in the blue to 3400 in the red.
- First light date: March 2014



3D Data Cube Format



MUSE reduced data:
(re)sampled 3D array
(300 x 300 x 3700) with
error and DQ information,
~2.6 GB per cube

- Immediate Goal: publication of reduced MUSE data through the ESO/SAF
 - Motivation: relatively demanding system requirements for MUSE data reduction
 - Support for Phase 3 data submissions from ESO large programmes (Note: refined policies as of P95)
- ESO Science Data Products Standard has been extended to cover 3D data cubes format for Integral Field Spectroscopy
 - Available from the ESO Phase 3 web
<http://www.eso.org/sci/observing/phase3>
- Implemented in the MUSE pipeline reduction software
- Timeline for publication of first products: 2016Q1
 - Next: K-band Multi-Object Spectrograph (KMOS)



3D Data Cube Format (2)

```

SIMPLE      = .....7 / file does conform to FITS standard
BITPIX     = .....8 / number of bits per data extension
NAXIS      = .....0 / number of data axes
EXTEND     = .....7 / FITS dataset may contain extensions
DATE       = '2015-05-20T10:20:35.' / file creation date (YYYY-MM-DDThh:mm:ss UT)
ORIGIN     = 'ESO-PARANAL' / European Southern Observatory
TELESCOP   = 'ESO-VLT-U4' / ESO Telescope
INSTRUME   = 'MUSE' / ESO Instrument name
RA         = .....183.46028 / [deg] Image centre (J2000.0)
DEC        = .....7.20120 / [deg] Image centre (J2000.0)
EQUINOX    = .....2000. / Standard FK5
RADECSYS   = 'FK5' / Coordinate system
EXPTIME    = .....2520.0 / Total integration time per pixel
TEXTTIME   = .....2520.0 / Total integration time all exposures
NCOMBINE   = .....3 / # of combined raw science data files
MJD_OBS    = .....57126.04953770 / 2015-04-14T01:11:20.1
MJD-END     = .....57126.08556882 / 2015-04-14T02:03:13.2
DATE-OBS   = '2015-04-14T01:11:20.057' / Observing date
OBJECT     = 'NGC 4191' / Target designation
OBJID     = .....1164690 / Observation Block ID
PROC ID    = '095.B-068(A)' / ESO programme identification code
PROVID    = 'MUSE.2015-04-14T01:11:20.057.fits' / Original science file

```

- 3-dimensional FITS image with two spatial and one spectral axis (FITS WCS convention)
- Pixel-by-pixel **error** and **data quality** information to be stored in separate extensions
 - Compatibility with CASA viewer
- Association of “white-light” image
- Processing provenance to trace back to the original raw data
- Phase 3 keywords:
 - PRODCATG= 'SCIENCE.CUBE.IFS'
 - WAVELMIN, WAVELMAX, ABMAGLIM, PIXNOISE, SPEC_RES, SKY_RES, SKY_RERR
- See: <http://www.eso.org/sci/observing/phase3>





Summary

- ✓ The ESO Science Archive Facility offers a wealth of science data products including calibrated images, spectra, and catalogues (total 56 TB, 2.5 Mio. files reduced data).
- ✓ Data being reduced both by ESO users (public surveys etc.) and by ESO in-house are available to the community in a seamless way.
- ✓ Download statistics demonstrate a strong interest by the user community.
- ✓ Upcoming: Phase 3 publication of data covering further instruments and modes (**MUSE**, HAWKI, VIMOS imaging, PIONIER)
- ✓ Upcoming public survey data product submissions for VISTA (DR4), VST (DR3), spectroscopic public surveys.