ESO Phase 3 Operations & the 3D Data Standard

Jörg Retzlaff

Archive Science Group,

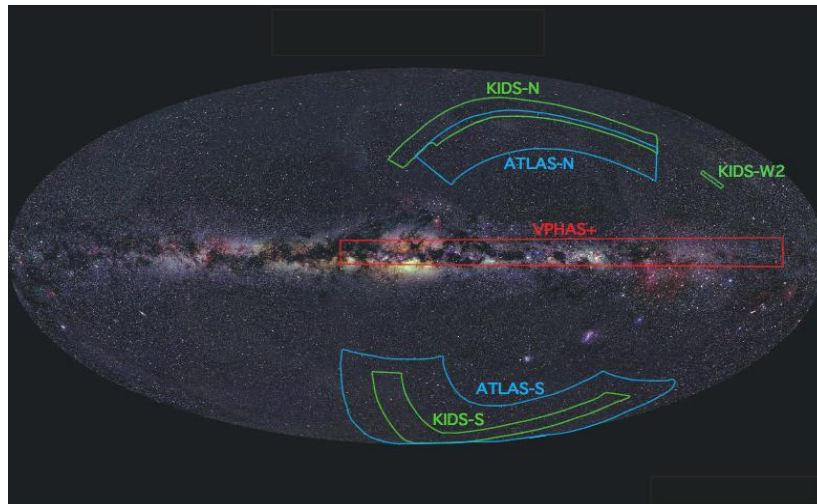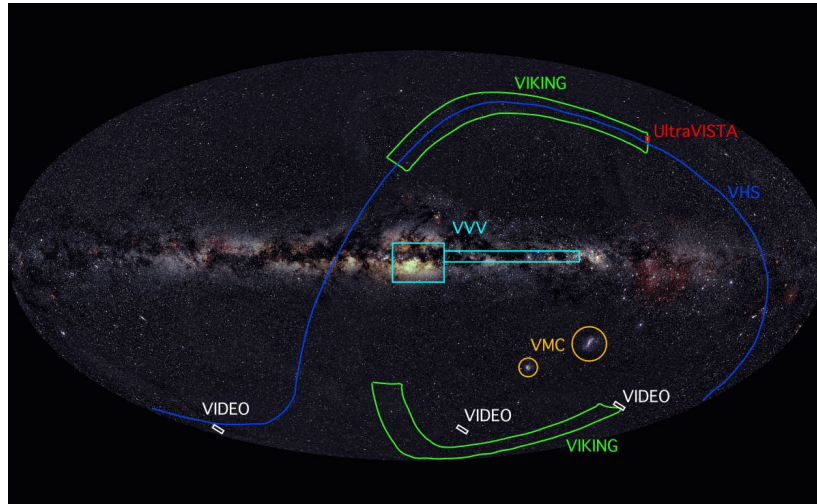ESO Back-end Operations (DMO)

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 deg^2 YJHKs (Ks<20 AB)

- VIKING - 1500 deg^2 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 deg^2 in bulge+plane plus multi-color map

- VMC - Magellanic Survey

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

- VST-Atlas – 4500 deg^2 UVRIz, like SDSS

- KIDS – 1500 deg^2 UVRI, 2.5 mag deeper than SDSS

- VPHAS+ 1800 deg^2 UVHaRI in the Southern Galactic Plane
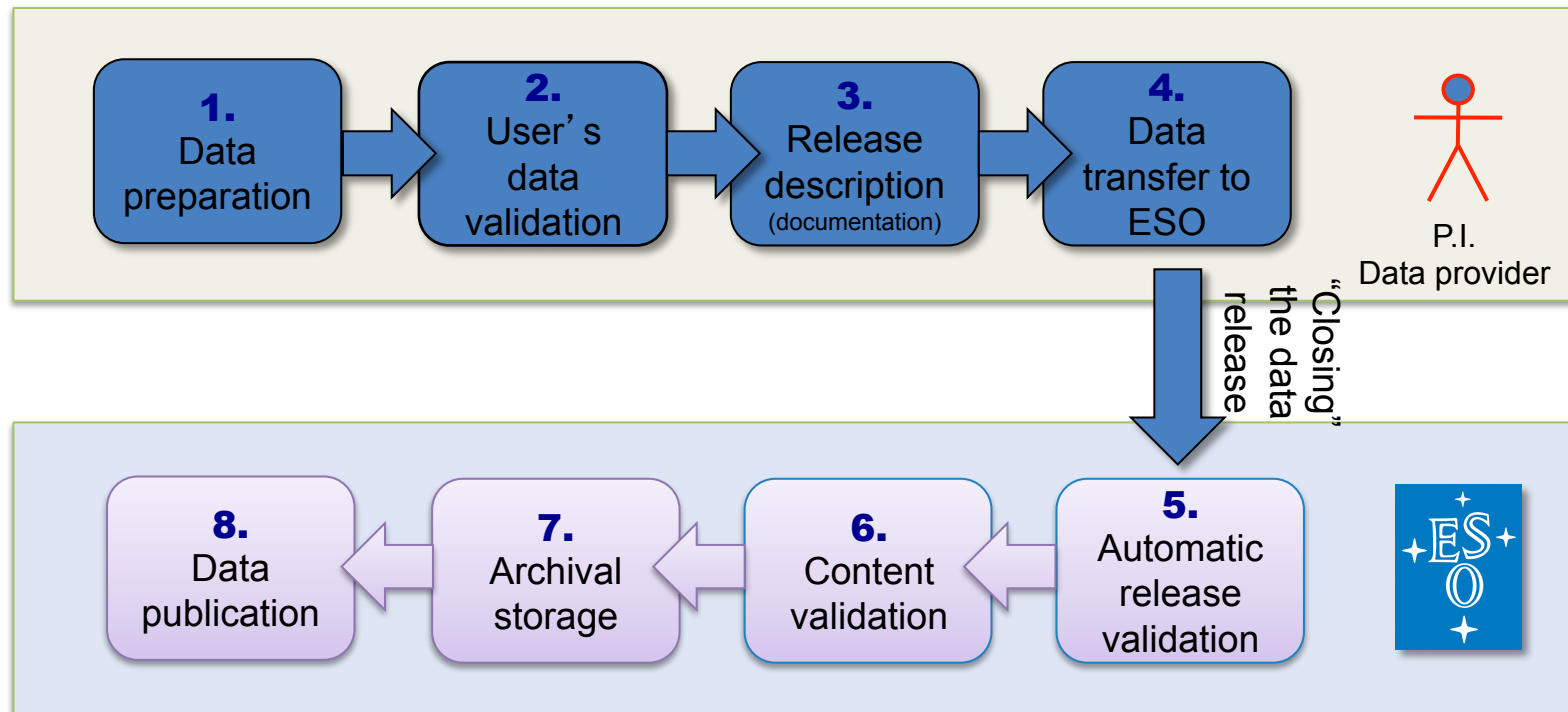
Spectroscopic Surveys: started 1 Jan 2012

- 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.
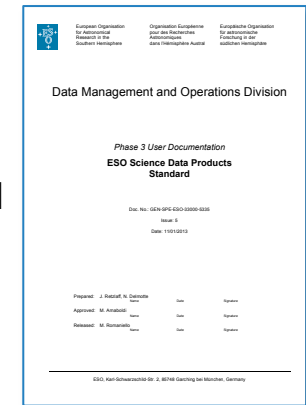
Data Preparation
ESO Science Data Product Standard

■ 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

Phase 3 Key Features

- **Data Validation:** check of format, completeness and quality of the metadata prior to archive ingestion.

- **Data documentation** associated to each Phase 3 release ("release-description")

  - ➢ based on ESO template, providing uniform structure
  - ➢ PDF document linked to each data product

- Support for incremental and superseding data releases.

- Phase 3 User Delegation supported

- Data release announcement via ESO Science Newsletter, Archive web etc.

Phase 3 Release Manager
http://www.eso.org/rm

Phase 3 FTP server

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

Release Overview

Phase 3 data releases organized by

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

Links to the associated data documentation

Links for Data Access

Note: Legacy data sets already published prior to the start of Phase 3 were migrated (e.g. the ESO-GOODS data collection)

ASTERICS DADI

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
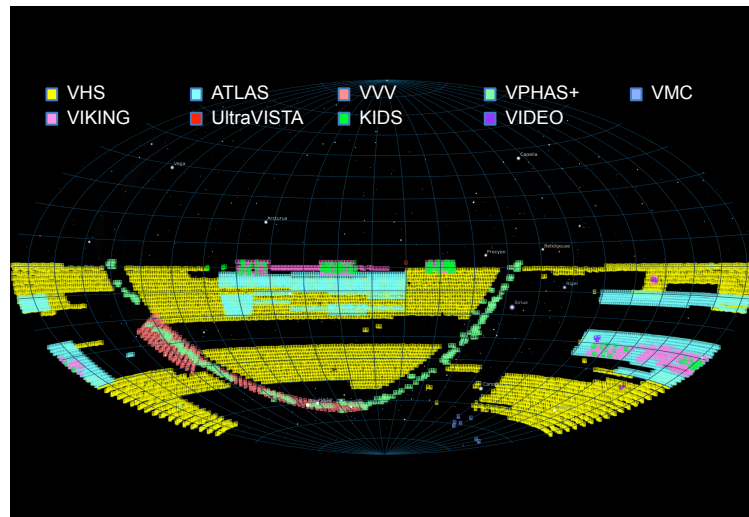
Reduced data from 11 ESO public surveys

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

- ■ VHS
- ■ VIKING
- ■ ATLAS
- ■ UltraVISTA
- ■ VVV
- ■ KIDS
- ■ VPHAS+
- ■ VIDEO
- ■ VMC

Science data processed by ESO

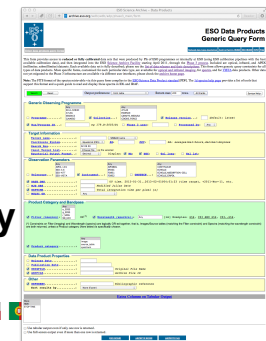|	No. of spectra	GB
UVES-Echelle	107,469	280
X-Shooter	47,919	878
HARPS	237,062	2,362
Giraffe-Medusa	1,142,908	5,308

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

- 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.**

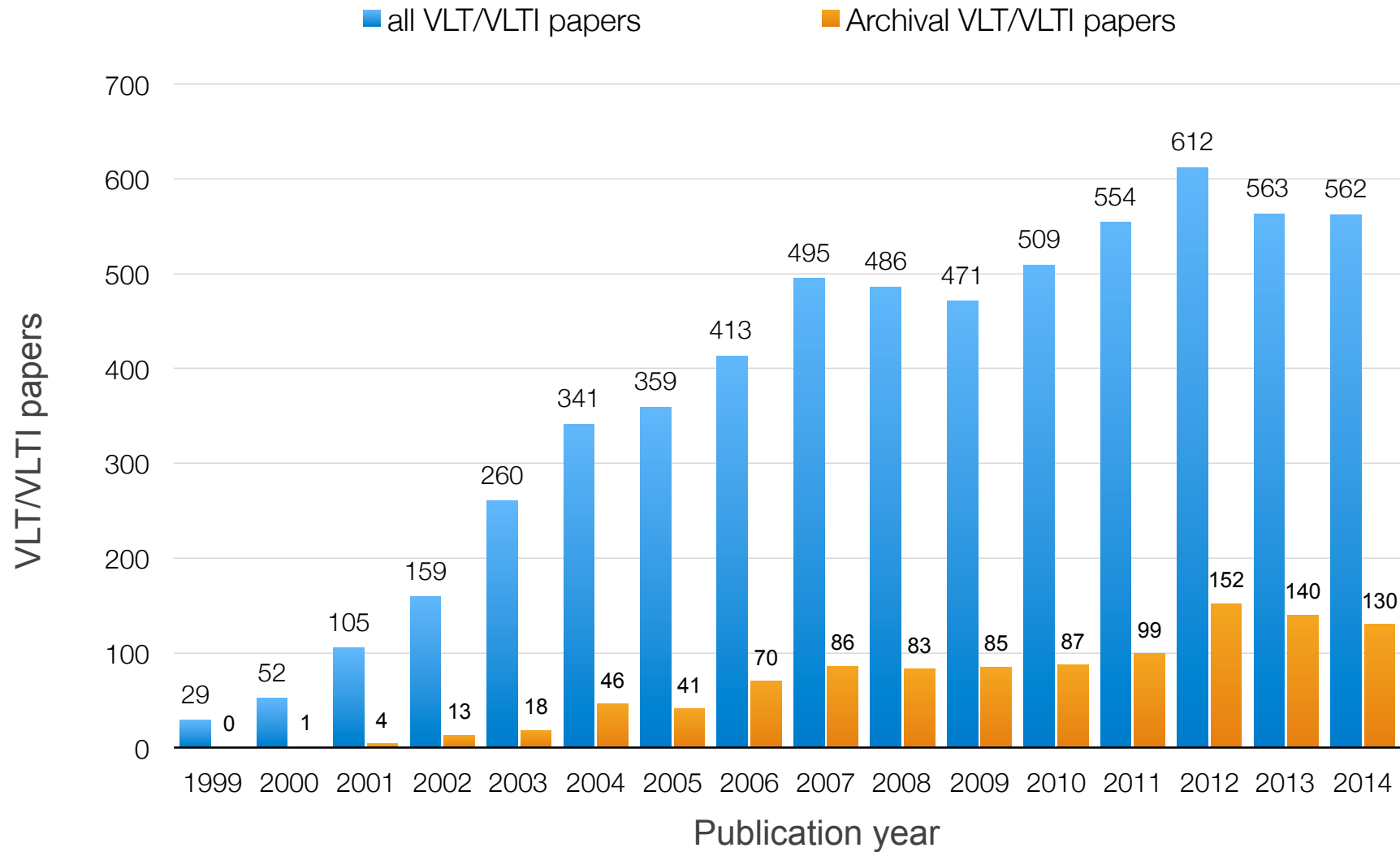# ARCHIVE SCIENCE AND DATA ACCESS

Archive Science

Example: VVV ESO public survey

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

Publications based on ESO Data

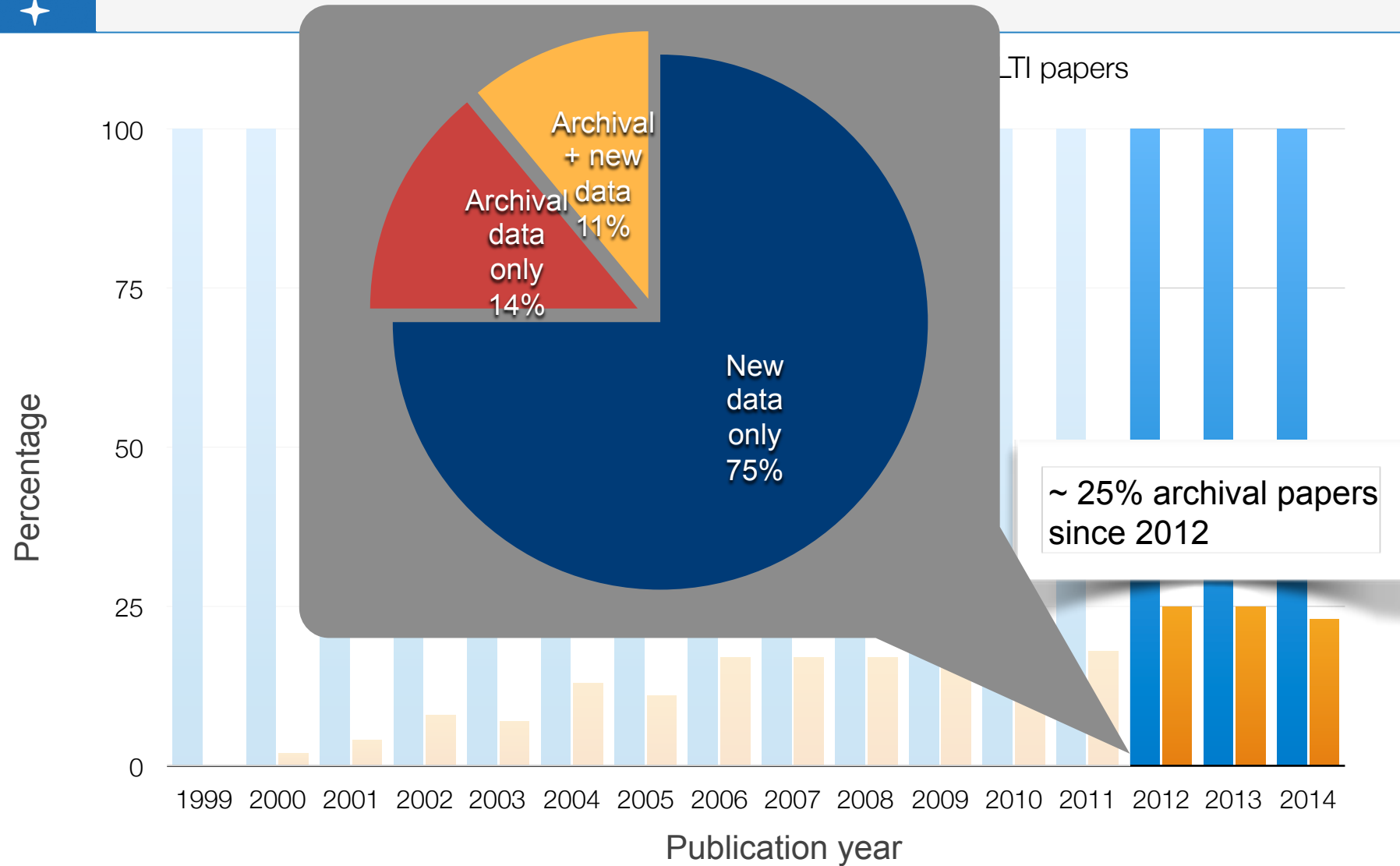- all VLT/VLTI papers
- Archival VLT/VLTI papers

Courtesy: U. Grothkopf/ESO Library

Archival papers - normalized

■ all VLT/VLTI papers   ■ Archival VLT/VLTI papers

~ 25% archival papers since 2012

Percentage

Publication year

Courtesy: U. Grothkopf/ESO Library

Archival papers - normalized

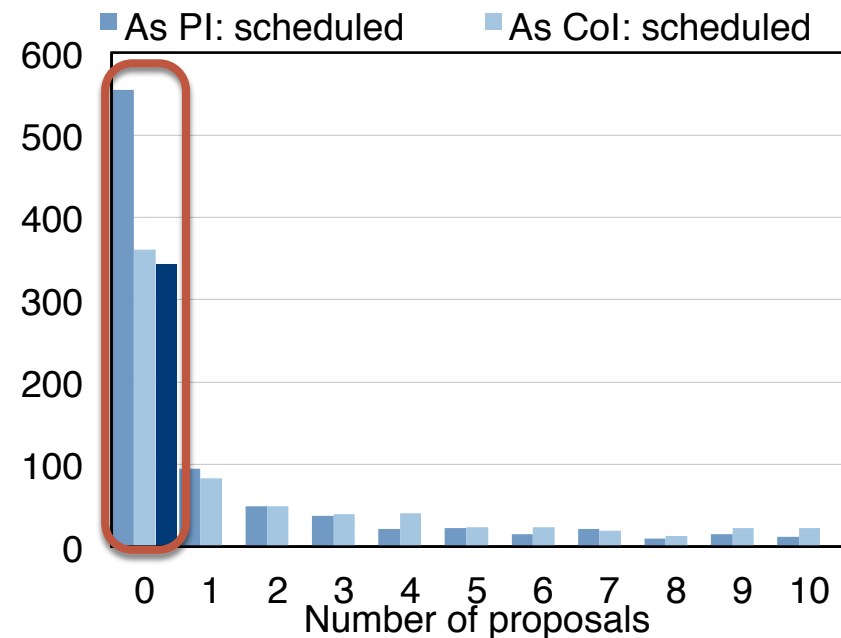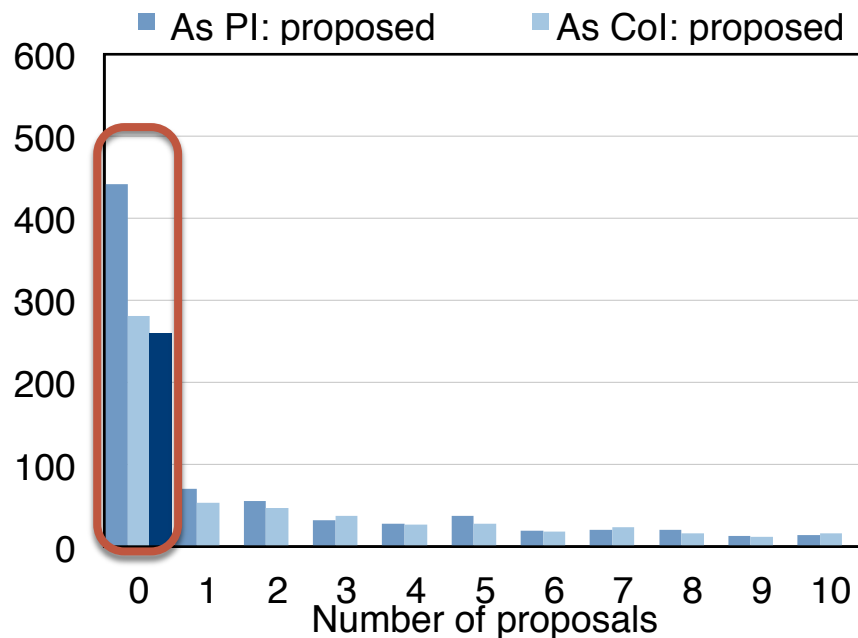# SAF Data Product Downloads

Phase 1 habits of archive users
Data products

- 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

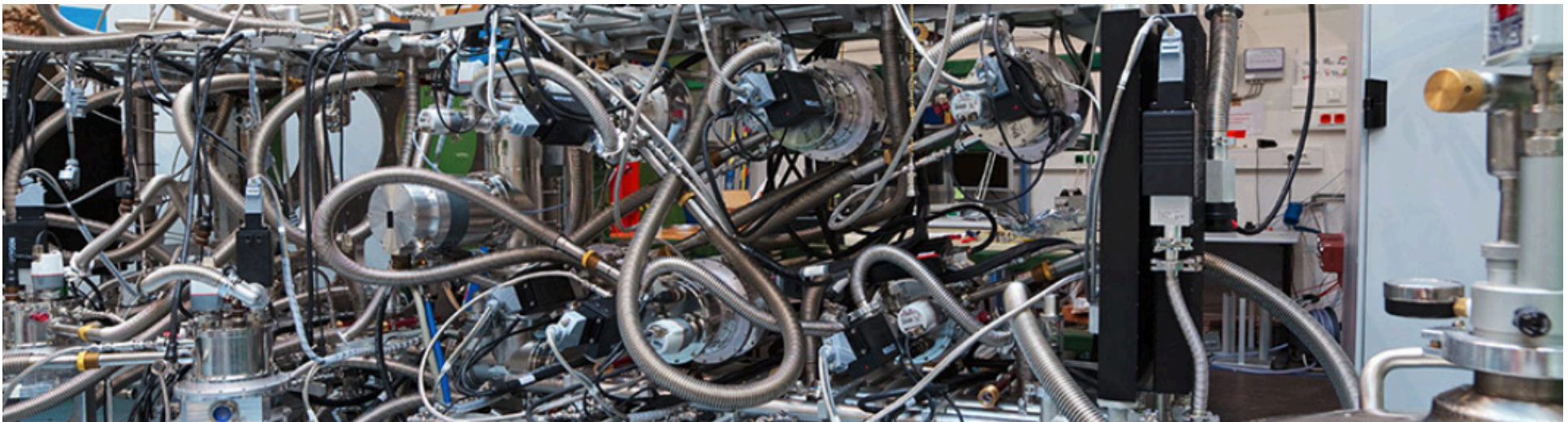 ➤ For comparison, 1/3 of the Phase 1 PIs have never got time
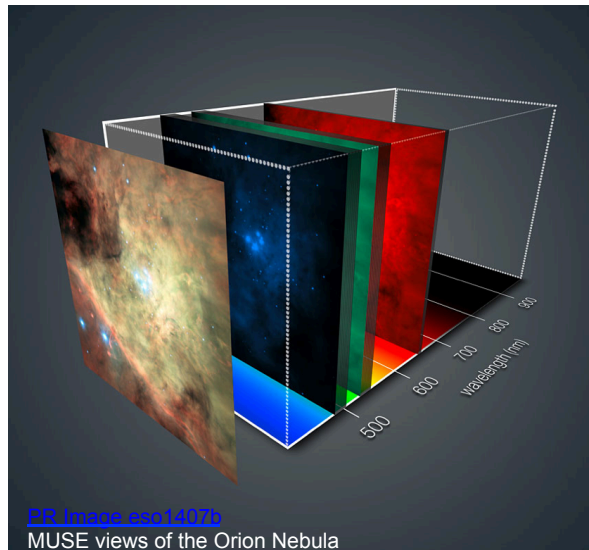
THE 3D DATA STANDARD

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

PR Image eso1407b
MUSE views of the Orion Nebula

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)

- 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.