



# (a quick introduction to) The ESO Science Archive Facility

Martino Romaniello

Head, Back-end Operations Department

ESO HQ - Garching bei München



# The ESO Data Flow System

- The science operations of the La Silla Paranal Observatory are embedded in an end-to-end data flow system that encompasses the entire lifecycle of scientific data
  - From the preparation of observing proposals to telescope scheduling, from the detailed definition of observing strategies to their execution at the telescope, from data processing to archival exploitation of the data
- The goal is to deliver science data to users specs and feed the Science Archive Facility
- It is one of the pillars of the success of the VLT
- Engraved in the “VLT/MLTI Science Operations Policy” document by the ESO Council



# Data access through the SAF

- The Science Archive Facility is **the** access point to ESO data
  - Online self service
    - Media only for veeeery large requests (400 GB and above)
  - Time critical data access: Target of Opportunity/Rapid Response Mode, pre-imaging, transients, planets, ...
  - Proprietary access to PIs and delegates
  - Subscription service to notify users of observation execution
  - Generic and data specific query interfaces
  - (Limited) programmatic access

<http://archive.eso.org>



# Data access through the SAF

Category	Query Forms	Data collection	Data Type	Instruments
<b>LPO Raw Data</b>	<b>Raw data query form</b> (all instruments) <b>Instrument specific query forms</b> <b>Direct retrieval of raw data by file name</b>	All ESO raw data	Various	Many La Silla Paranal instruments
<b>LPO Data Products</b>	<b>Phase 3 main query form</b> <b>Phase 3 imaging query form</b> <b>Phase 3 instrument specific query form</b>	Phase 3 Data Products (ESO Public Surveys)	Currently, Imaging	Currently, VISTA/VIRCAM
	<b>Catalogue Facility query interface</b>	Phase 3 Catalogues [ESO User Portal authentication required also when browsing]	Catalogues	Currently, VISTA/VIRCAM
	<b>Advanced Data Products query form</b>	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
		UVES reprocessed	Spectroscopy	UVES
		Time-domain survey of NGC 2547 (S.Aigrain)	Imaging	FEROS
	<b>FEROS/HARPS pipeline processed data query form</b>	FEROS/HARPS pipeline processed data	Spectroscopy	FEROS, HARPS
<b>Science Verification, Commissioning, EIS, etc.</b>	Full list of available data packages	Various	Many	
<b>APEX Quick Look Products</b>	<b>APEX query form</b>	APEX	Heterodyne, Bolometer	APEX-2A, LABOCA, SABOCA, SHeFI
<b>ALMA Data</b>	<b>ALMA Science Archive</b>	All ALMA data	Cube	ALMA

<http://archive.eso.org>

# CalSelector

- Archive service to associate to raw science files all the raw and static calibrations needed for processing, plus ancillary files (e.g. acquisition frames), night log excerpts and a description of the association itself (xml format)
  - Delivers quality certified calibrations
  - Reproduces the Calibration Plans (and works with Reflex)
- Complete coverage from 2009
  - Effort to extend as far back in time as possible: the look-back time depends on instrument and mode
- Currently raw-to-raw, raw-to-masters under study
- <http://www.eso.org/sci/archive/calselectorInfo.html>



# CalSelector

Select	Dataset	File (Category)	Size	Access
<input type="checkbox"/>	<input type="checkbox"/> ✓ SAF+UVES.2008-04-28T06:43:54.974			
<input type="checkbox"/>		UVES.2008-04-28T06:43:54.974.fits.Z (SCI_POINT_RED)	16.2MB	✓
<input type="checkbox"/>		M.UVES.2012-03-07T17:01:05.530.tfits (LINE_REFER_TABLE)	33.8KB	✓
<input type="checkbox"/>		M.UVES.2012-03-07T17:02:30.714.tfits (LINE_INTMON_TABLE)	8.4KB	✓
<input type="checkbox"/>		M.UVES.2012-03-07T17:04:35.088.tfits (EXTCOEFF_TABLE)	8.4KB	✓
<input type="checkbox"/>		UVES.2008-04-28T06:43:46.615.fits.Z (ACQ_ECH)	317.0KB	✓
<input type="checkbox"/>		UVES.2008-04-28T10:27:16.099.fits.Z (BIAS_RED)	8.5MB	✓
<input type="checkbox"/>		UVES.2008-04-28T10:28:02.813.fits.Z (BIAS_RED)	8.5MB	✓
<input type="checkbox"/>		UVES.2008-04-28T10:28:49.547.fits.Z (BIAS_RED)	8.5MB	✓
<input type="checkbox"/>		UVES.2008-04-28T10:29:36.261.fits.Z (BIAS_RED)	8.5MB	✓
<input type="checkbox"/>		UVES.2008-04-28T10:30:22.975.fits.Z (BIAS_RED)	8.5MB	✓
<input type="checkbox"/>		UVES.2008-04-28T12:23:04.986.fits.Z (EFLAT_RED)	30.0MB	✓
<input type="checkbox"/>		UVES.2008-04-28T12:24:26.823.fits.Z (EFLAT_RED)	29.9MB	✓
<input type="checkbox"/>		UVES.2008-04-28T12:25:48.881.fits.Z (EFLAT_RED)	30.0MB	✓
<input type="checkbox"/>		UVES.2008-04-28T12:27:10.808.fits.Z (EFLAT_RED)	30.0MB	✓
<input type="checkbox"/>		UVES.2008-04-28T12:28:32.795.fits.Z (EFLAT_RED)	30.1MB	✓
<input type="checkbox"/>		UVES.2008-04-28T12:30:02.093.fits.Z (ECH_ARC_LAMP_RED)	21.3MB	✓
<input type="checkbox"/>		UVES.2008-04-28T12:31:47.652.fits.Z (ORDER_FLAT_RED)	16.2MB	✓
<input type="checkbox"/>		UVES.2008-04-28T12:33:24.621.fits.Z (ECH_ARC_LAMP_FORM_RED)	11.6MB	✓
<input type="checkbox"/>		UVES.2008-04-28T06:43:54.974.NL.txt (NIGHTLOG_INFO)	332B	✓
<input type="checkbox"/>		UVES.2008-04-28T06:43:54.974.xml (ASSOCIATION_TREE)	7.6KB	✓

# The SAF's content

- All raw data from the La Silla Paranal Observatory (including APEX)
  - ALMA European copy also stored and operated at ESO
    - (mostly) joint operations, service integration under study
- Selected data products
  - Contributed by the community
  - Generated in-house
- Total holdings: about 1/3 of a Petabyte
  - Monthly inflow ~7-8 TB
  - Two copies on spinning disks for safety and redundant access
  - Custom storage and handling softwares; Sybase databases

# External Data Products (Phase 3)

- Phase 3 – PIs of observing programmes return data products to ESO for publication to the scientific community
  - Phase 3 is mandatory for ESO Public Surveys and for ESO Large Programmes since period 75; available also for other ESO observations
- Each Phase 3 programme is focussed on a specific science case, *together* they cover a broad range of cases (wide, deep, Galactic, extragalactic, imaging, spectroscopy, transients, monitoring, ...)
- Recent releases
  - Data Release 2 from VISTA Public Surveys (IR imaging, including source lists)
  - Data Release 1 from VST Public Surveys (optical imaging, including source lists)
  - Stellar parameters from MATISSE@Observatoire de la Côte d’Azur
  - ...
- Upcoming releases
  - Source Catalogues from VST, VISTA Public Surveys (Ultra-VISTA already available)
  - 1D calibrated spectra from Spectroscopic Public Surveys
  - **Your** contribution, perhaps?



# Internal Data Products

- Science grade data products generated in-house by running the corresponding instrument pipelines
  - Driven by data, rather than by a specific science goal
  - Uniform processing with a standard set of processing parameters
- Seamless archive experience with External Data Products, e.g. from Public Surveys, Large Programmes, etc.
- Timeline
  - Publication of **UVES** Echelle data as we speak (since start of operations+stream of new data)
  - Then (*tentative*): X-Shooter-Echelle, FLAMES-MEDUSA, HAWK-I and VIMOS imaging (UK in-kind contribution), KMOS, MUSE

# Digression: user science processing

- Should the data products from the Science Archive Facility not suit your needs, you can go back and reduce the data!
  - Data reduction pipelines available for virtually all of Paranal instrument modes
- Reflex: an environment for easy and flexible execution of VLT/I data reduction modules
  - Fully automatic data organization (full calibration cascade supported)
  - Support for multi OB processing
  - Conditional branches, loops and conditional stops
  - Interfaces to Python (hence IRAF and MIDAS) and IDL
- ESO releases fully functional workflows
  - Mixture of fully scientific validated and interactive workflows (UVES, X-Shooter, KMOS) and simpler ones to speed-up data organization and basic reduction (VIMOS spectroscopy, FORS2 MXU)



# Digression: user science processing



## X-shooter Workflow for Physical Mode Slit Data Reduction (v. 2.2.0)

Workflow Instructions	Setup Directories	Global Parameters
-----------------------	-------------------	-------------------

To run this workflow on the demo data:

- Turn on highlighting. Choose "Tools" -> "Animate at Runtime" from top menu and set it to "1".
- Press the "Run" button OR cntrl-R to start the workflow.

To run on a different data set:

- Click on ROOT\_DATA\_DIR and set as appropriate. All subdirectories of RAWDATA\_DIR will be searched for data.
- If desired, change END\_PRODUCTS\_DIR.
- Press the "Run" button OR cntrl-R to start the workflow.

To monitor the progress of the workflow in more detail:

- Open "Window" -> "Runtime Window" in top menu before starting the workflow.

**Input:**

- ROOT\_DATA\_DIR: /Users/mromanie/Reflex/Kepler/RELEASES/v2.4/data\_wkf
- RAWDATA\_DIR: \$ROOT\_DATA\_DIR/reflex\_input/Xshooter
- CALIB\_DATA\_DIR: /Users/mromanie/Reflex/Kepler/RELEASES/v2.4/install/calib/xsh-2.2.0

**Working Directories:**

- BOOKKEEPING\_DIR: \$ROOT\_DATA\_DIR/reflex\_book\_keeping/Xshooter
- LOGS\_DIR: \$ROOT\_DATA\_DIR/reflex\_logs/Xshooter
- TMP\_PRODUCTS\_DIR: \$ROOT\_DATA\_DIR/reflex\_tmp\_products/Xshooter

**Output:**

- END\_PRODUCTS\_DIR: \$ROOT\_DATA\_DIR/reflex\_end\_products

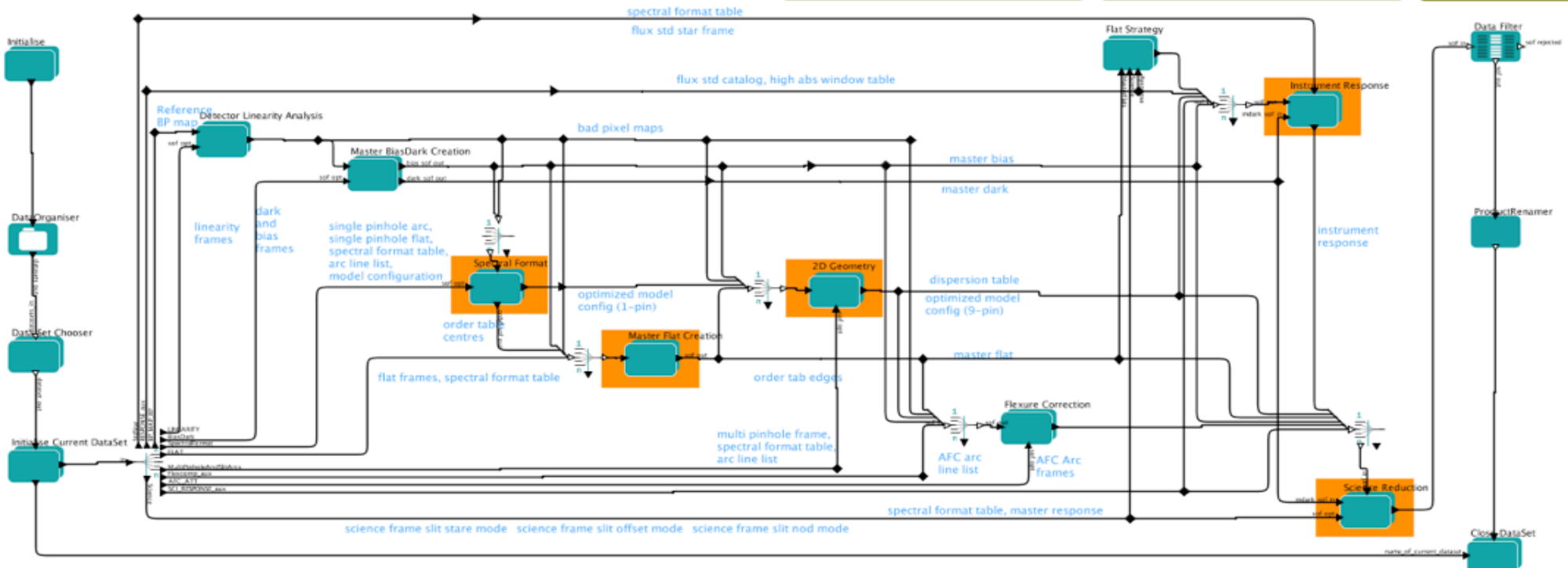
● RecipeFailureMode: Ask  
● EraseDirs: false  
● FITS\_VIEWER: fv

Global parameter for the behaviour when a recipe fails. 'Ask' means that each time a recipe fails, the choice to continue or stop will be presented. 'Continue' means that the workflow will ignore errors and continue, 'Stop' means it will stop.

Change "EraseDirs" to true to erase BOOKKEEPING\_DIR, TMP\_PRODUCTS\_DIR and LOGS\_DIR each time the workflow is run (Lazy Mode will not work anymore)

fits viewer to use for the inspection of input/output products

The X-shooter workflow tutorial and demo data and the pipeline user manual can be found here: [http://www.eso.org/sci/software/pipelines/#reflex\\_workflows](http://www.eso.org/sci/software/pipelines/#reflex_workflows)





# Digression: user science processing



## X-shooter Workflow for Physical Mode Slit Data Reduction (v. 2.2.0)

**Workflow Instructions**      **Setup Directories**      **Global Parameters**      = actor with interactive option

To run this workflow on the demo data:

- Turn on highlighting. Choose "Tools" from top menu and set it to "1".
- Press the "Run" button OR cntrl-R to

To run on a different data set:

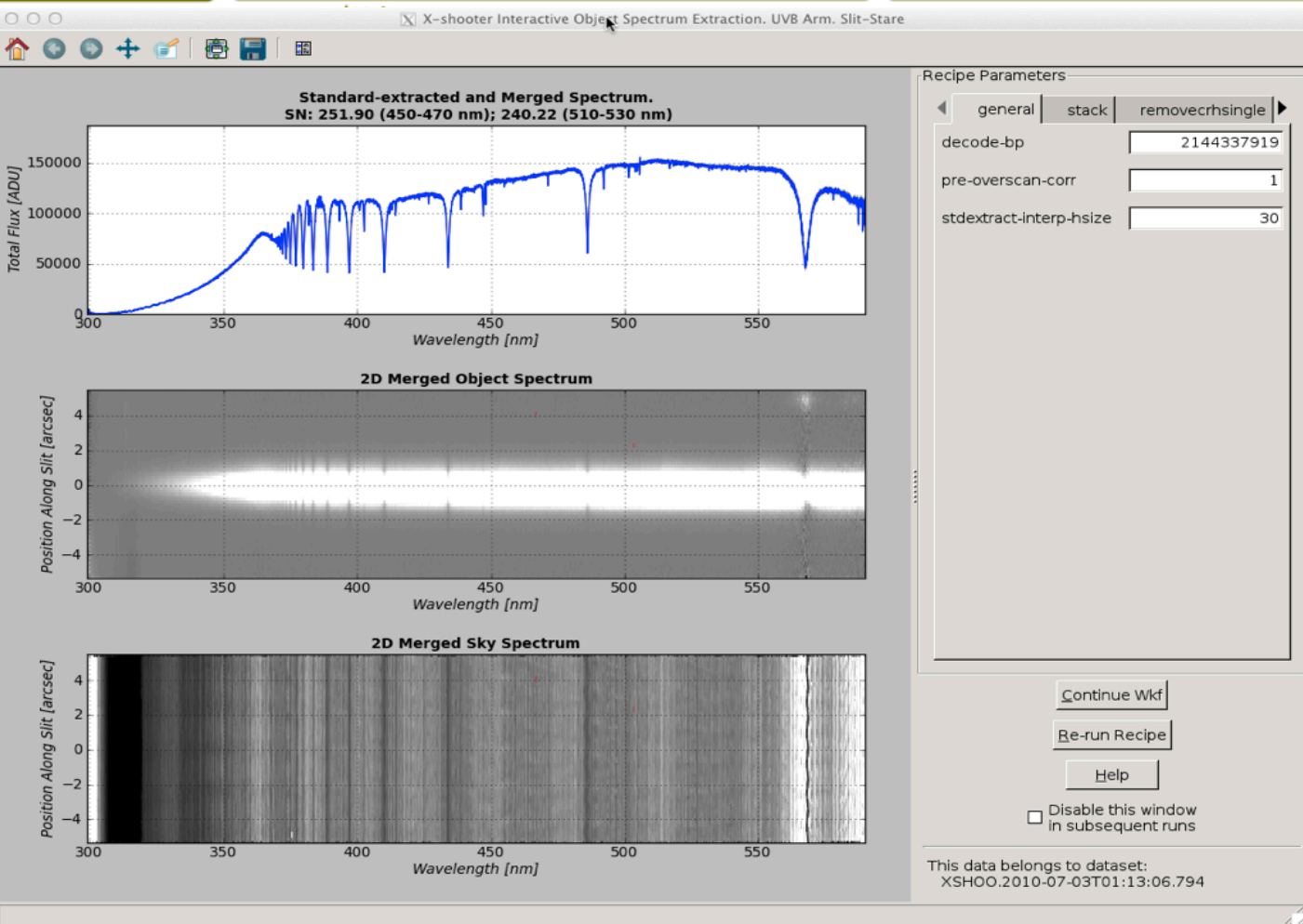
- Click on ROOT\_DATA\_DIR and set as All subdirectories of RAWDATA\_DIR w
- If desired, change END\_PRODUCTS\_DI
- Press the "Run" button OR cntrl-R to

To monitor the progress of the workflow

- Open "Window" -> "Runtime Window starting the workflow.

The X-shooter workflow tutorial and dem <http://www.eso.org/sci/software/pipeline>

**Step 1: Data Organisation and Selection**

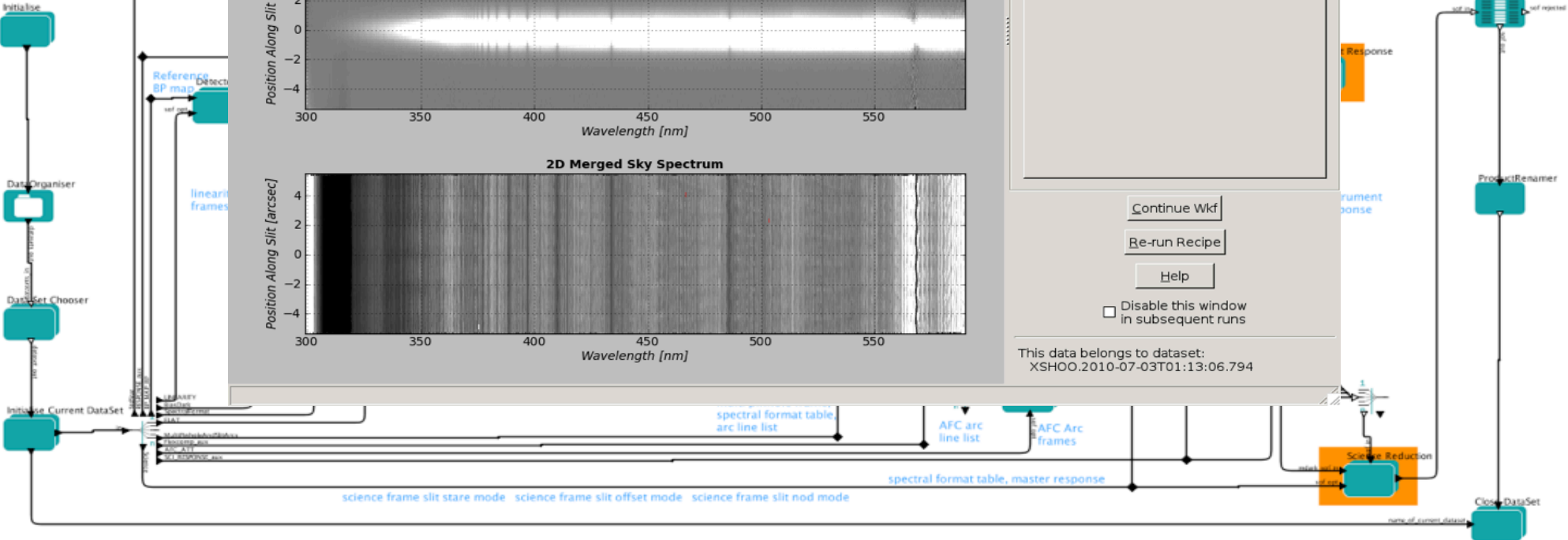


When a recipe fails, the choice to 'Continue' means that the workflow will continue, 'Stop' means the workflow will stop.

Use 'Continue' to continue the workflow. 'Stop' means the workflow will stop. 'Continue' means that the workflow will continue, 'Stop' means the workflow will stop.

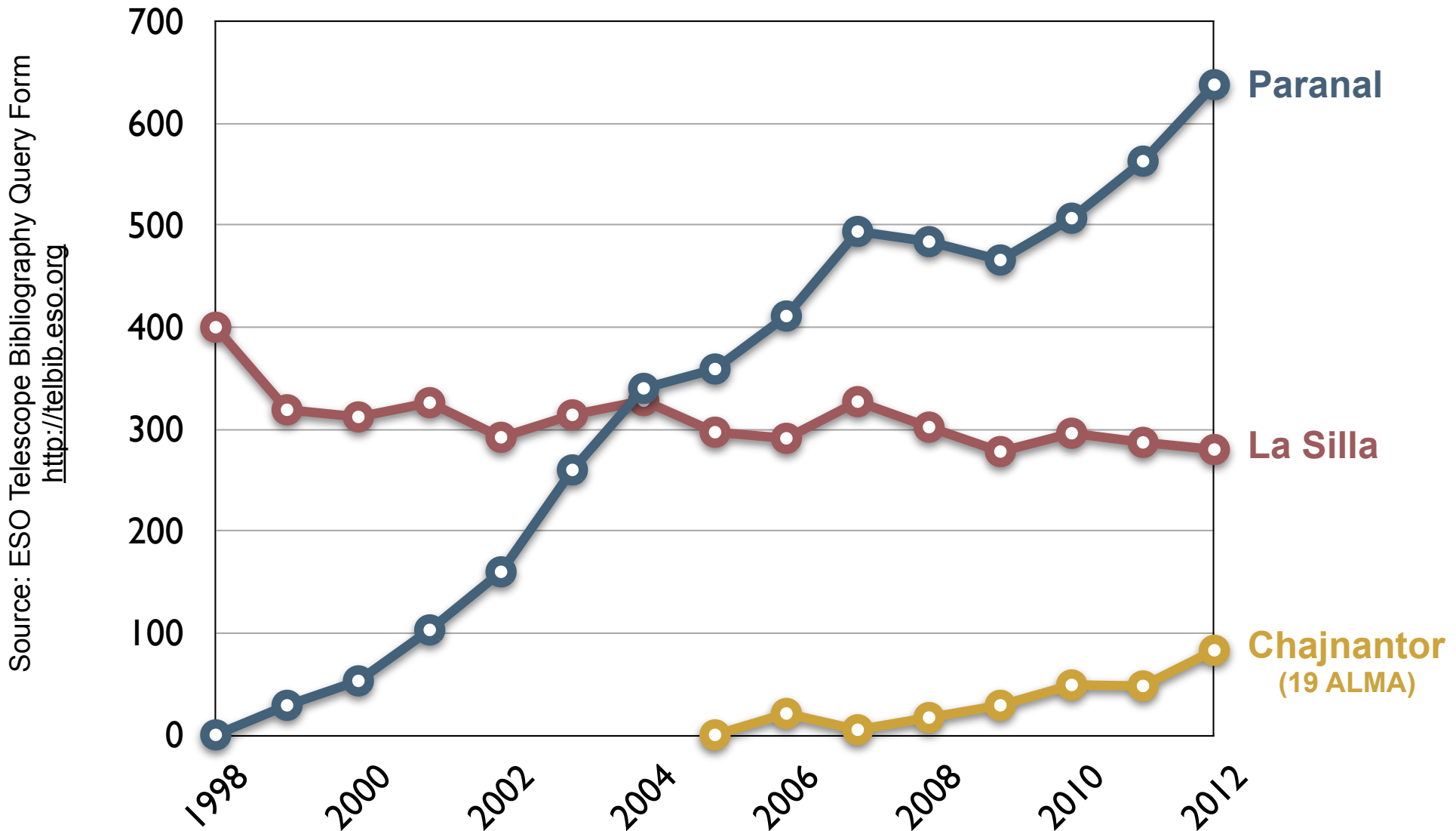
Use 'Continue' to continue the workflow. 'Stop' means the workflow will stop. 'Continue' means that the workflow will continue, 'Stop' means the workflow will stop.

**Step 5: Output Organisation**



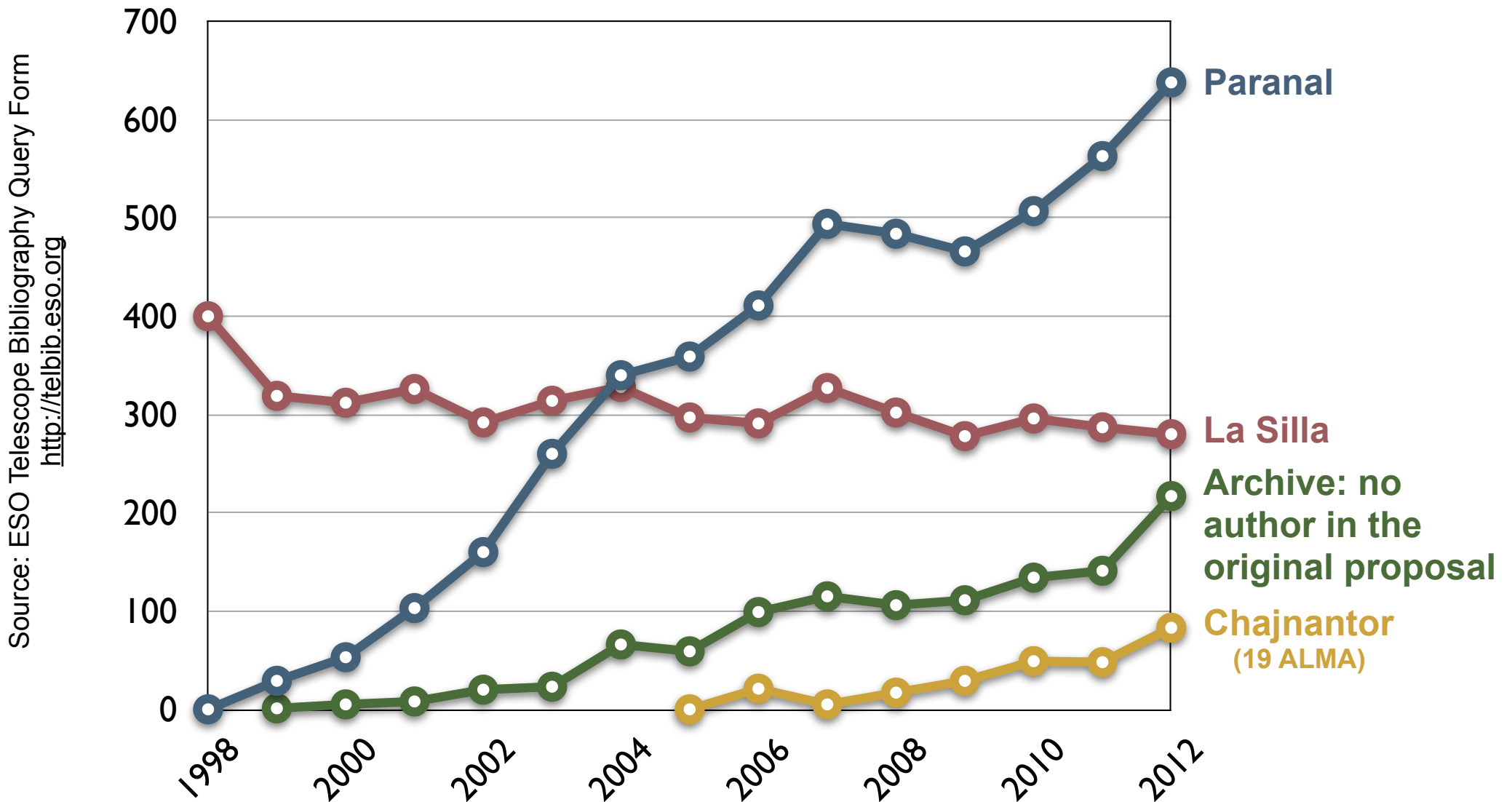
# The Archive as Science Resource

## ■ Refereed publications from ESO facilities



# The Archive as Science Resource

## ■ Refereed publications from ESO facilities



# Evolution of archive services

- The inclusion of Internal Data Products in Phase 3 marks the completion of the science data product handling infrastructure
  - Filling of the SAF as “routine” operations
  - Reuse/adaptation of mature data standards
- Next is to keep building archive services to exploit the content
  - We are exploring classic archive services (e.g. cutouts, datasets, etc.), as well as novel ways to establish a dialogue between the science user and the SAF (e.g. faceted search)
  - We are also actively cooperating and seeking synergies with other data centers and archives, e.g. the CDS in Strasbourg, the Canadian CADC and ALMA

What would **you** like to have??