(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/VLTI 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
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 instrument specific** query form	Phase 3 Data Products (ESO Public Surveys)	Currently, Imaging	Currently, VISTA/VIRCAM
	Catalogue Facility query interface	Phase 3 Catalogues [ESO User Portal authentication required also when browsing]	Catalogues	Currently, VISTA/VIRCAM
	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
		UVES reprocessed	Spectroscopy	UVES
		Time-domain survey of NGC 2547 (S.Aigrain)	Imaging	FEROS
	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
ALMA Data	**ALMA Science Archive**	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
☐	⊟ ✔ SAF+UVES.2008-04-28T06:43:54.974			
☐		UVES.2008-04-28T06:43:54.974.fits.Z *(SCI_POINT_RED)*	16.2MB	✔
☐		M.UVES.2012-03-07T17:01:05.530.tfits *(LINE_REFER_TABLE)*	33.8KB	✔
☐		M.UVES.2012-03-07T17:02:30.714.tfits *(LINE_INTMON_TABLE)*	8.4KB	✔
☐		M.UVES.2012-03-07T17:04:35.088.tfits *(EXTCOEFF_TABLE)*	8.4KB	✔
☐		UVES.2008-04-28T06:43:46.615.fits.Z *(ACQ_ECH)*	317.0KB	✔
☐		UVES.2008-04-28T10:27:16.099.fits.Z *(BIAS_RED)*	8.5MB	✔
☐		UVES.2008-04-28T10:28:02.813.fits.Z *(BIAS_RED)*	8.5MB	✔
☐		UVES.2008-04-28T10:28:49.547.fits.Z *(BIAS_RED)*	8.5MB	✔
☐		UVES.2008-04-28T10:29:36.261.fits.Z *(BIAS_RED)*	8.5MB	✔
☐		UVES.2008-04-28T10:30:22.975.fits.Z *(BIAS_RED)*	8.5MB	✔
☐		UVES.2008-04-28T12:23:04.986.fits.Z *(EFLAT_RED)*	30.0MB	✔
☐		UVES.2008-04-28T12:24:26.823.fits.Z *(EFLAT_RED)*	29.9MB	✔
☐		UVES.2008-04-28T12:25:48.881.fits.Z *(EFLAT_RED)*	30.0MB	✔
☐		UVES.2008-04-28T12:27:10.808.fits.Z *(EFLAT_RED)*	30.0MB	✔
☐		UVES.2008-04-28T12:28:32.795.fits.Z *(EFLAT_RED)*	30.1MB	✔
☐		UVES.2008-04-28T12:30:02.093.fits.Z *(ECH_ARC_LAMP_RED)*	21.3MB	✔
☐		UVES.2008-04-28T12:31:47.652.fits.Z *(ORDER_FLAT_RED)*	16.2MB	✔
☐		UVES.2008-04-28T12:33:24.621.fits.Z *(ECH_ARC_LAMP_FORM_RED)*	11.6MB	✔
☐		UVES.2008-04-28T06:43:54.974.NL.txt *(NIGHTLOG_INFO)*	332B	✔
☐		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**?**

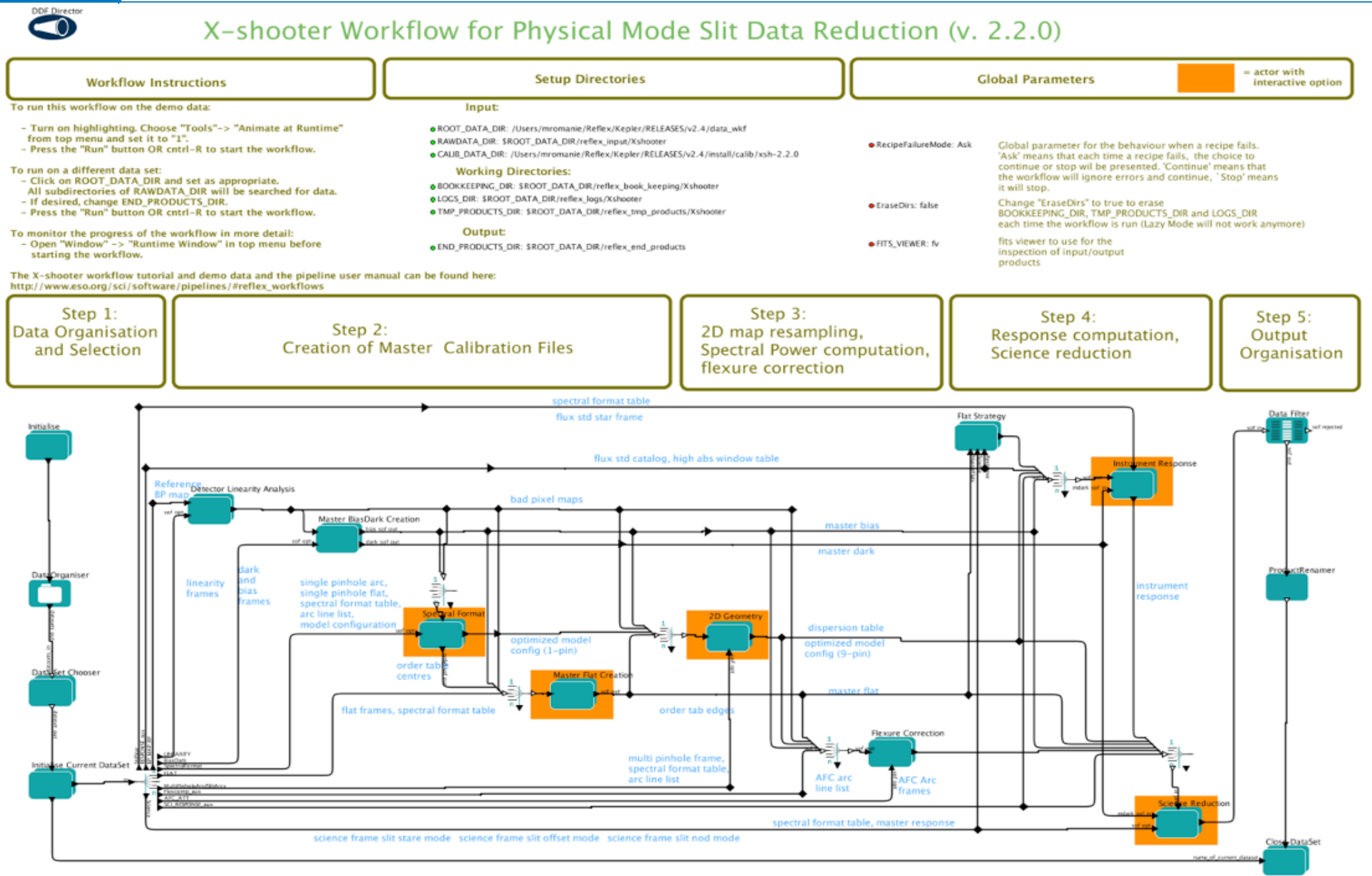
Internal Data Products

- Science grade data products generated in-house by running the corresponding instrument pipelines
  - Driven by data, rather then 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 **U**VES 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

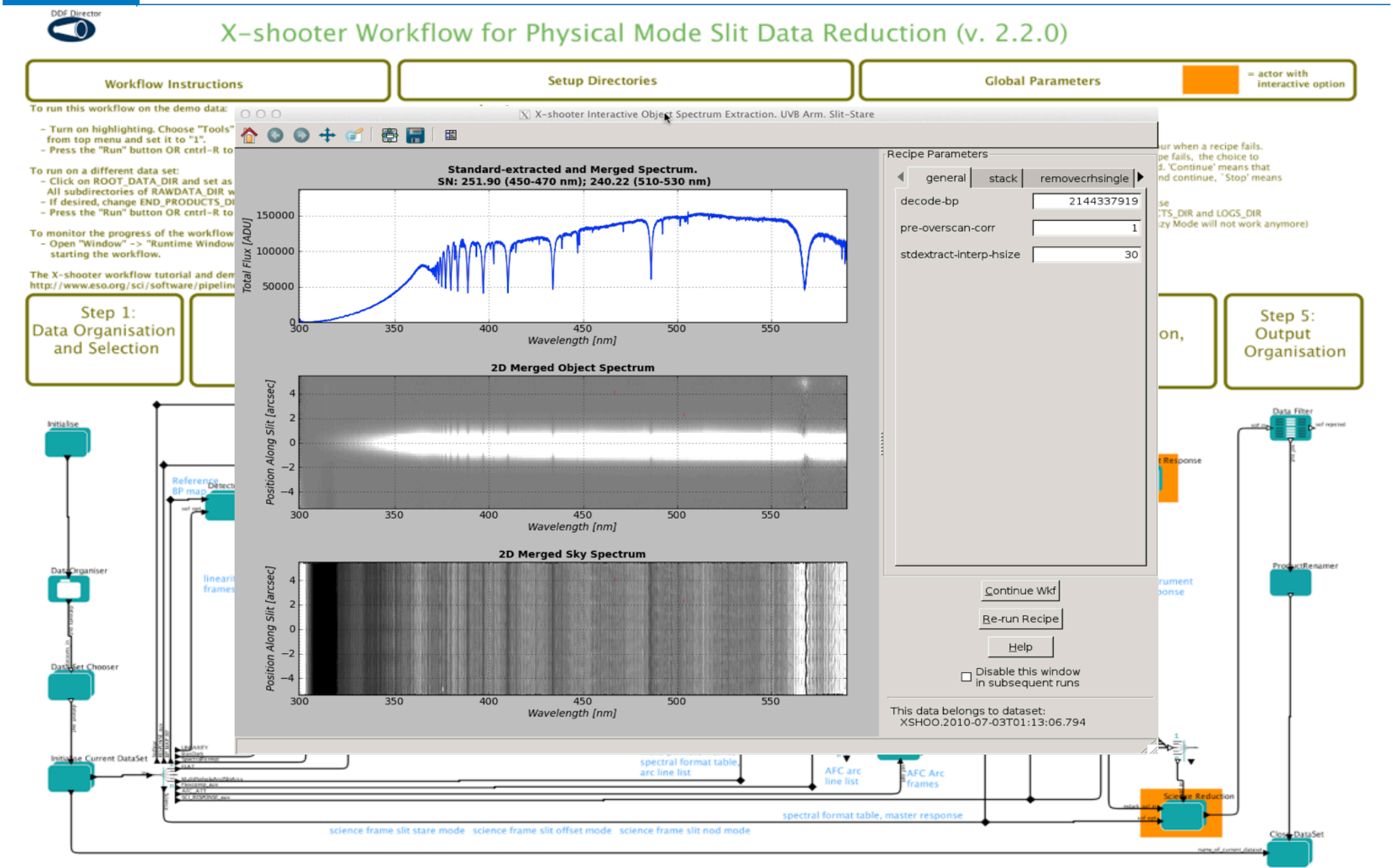
Digression: user science processing

- 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

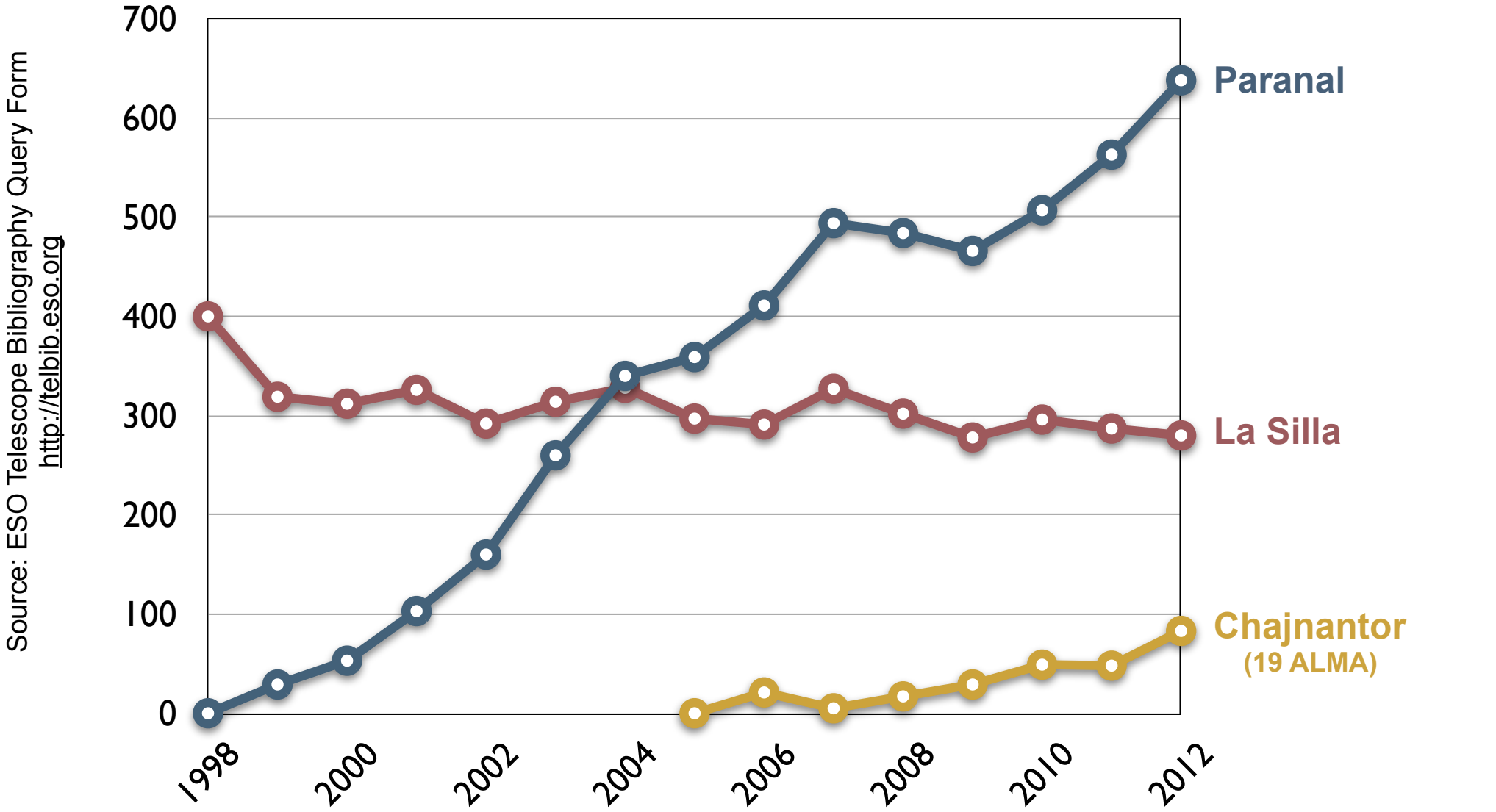


Digression: user science processing



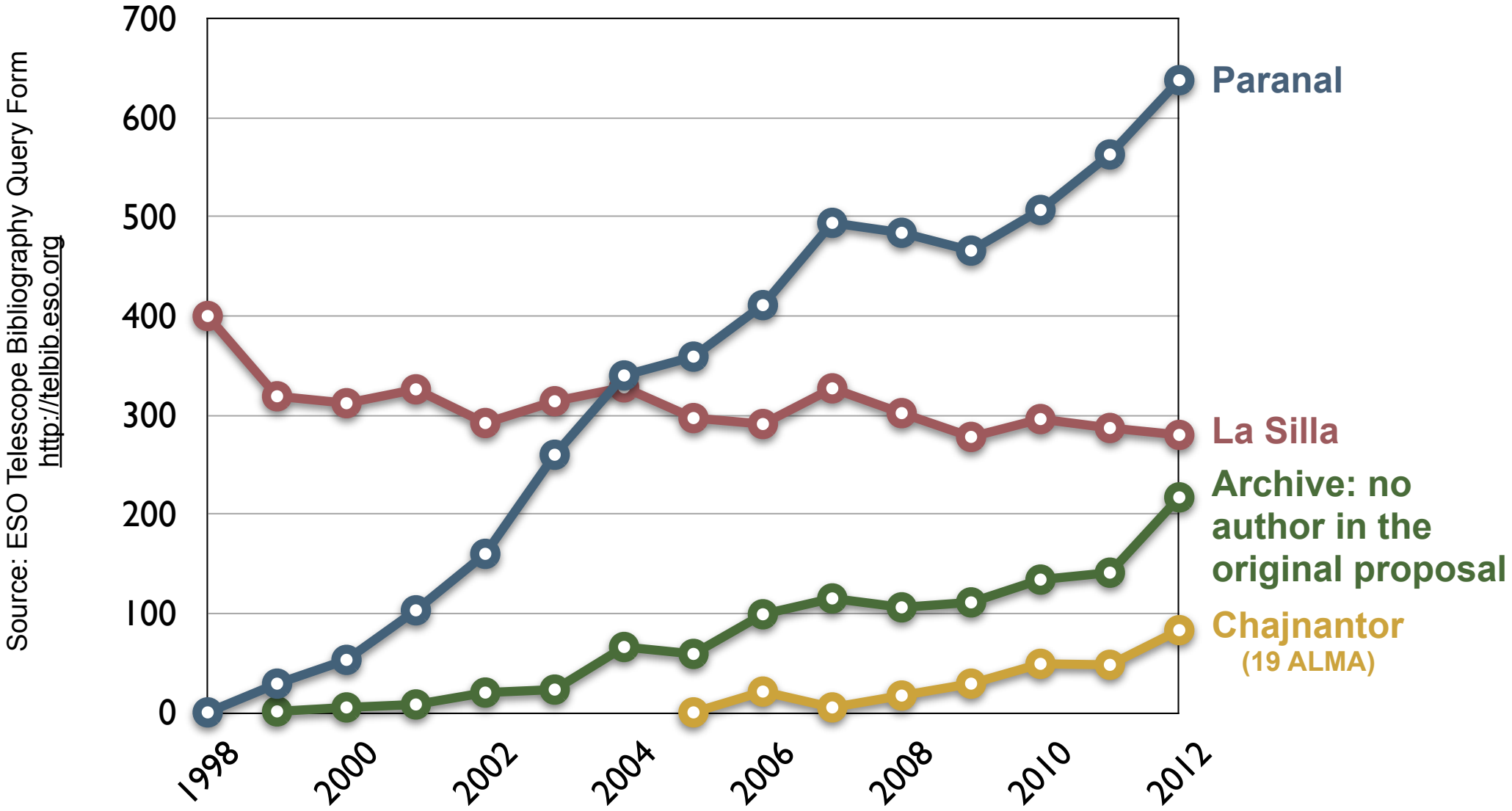
The Archive as Science Resource

- 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??