# Herschel DP status and plans: the users point of view

**Bruno Merín**
Herschel Science Centre / ESAC

***Herschel Pre-Launch Data Processing Workshop***
*ESAC, 4-5 December 2008*

HERSCHEL SPACE OBSERVATORY

# Outline

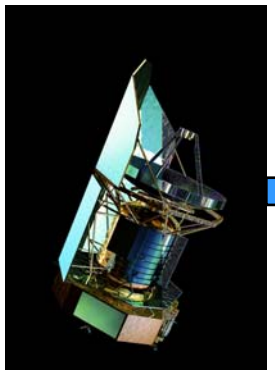
What do you need to know about..

- The Herschel data
- The Herschel Data Processing (DP) software
- Status and plans of the Herschel DP software
- HIPE: Herschel Interactive Processing Environment

HERSCHEL SPACE OBSERVATORY

# The Herschel Data: Dataflow

1. Herschel performs an observation following a user AOR
2. The spacecraft telemetry arrives at Mission Operations Control (MOC) centre packed in timelines and is sent to HSC



Telemetry Timelines
Level 0
MOC

Automatic Pipelines:
Level 0 to Level 2
HSC

Herschel Science Archive (HSA)
All levels
HSC

HIPE
- re-run Pipelines
- Scientific Analysis
Astronomer

Publication
Level 3
Back to HSA
Astronomer

3. The data is run automatically through the HSC data pipelines to remove all instrumental effects to produce level-2 products
4. Raw and processed data is placed in the HSA for the users
5. Astronomers will use HIPE to download, reprocess the data (if needed) and analyze it within just one environment

HERSCHEL SPACE OBSERVATORY

# The Herschel Data: Pipelines

## DP Pipelines: Standard Product Generation (SPG)

- DP pipelines (SPG) will be run routinely at the HSC and populate the Herschel Science Archive (HSA)
- AOT specific pipelines generate the Herschel standard products and can also be launched from HIPE
- In addition, it is also possible to request a on-demand reprocessing of the data with customized pipelines in the ESAC GRID through the Herschel Science Archive
- The S/W level of completeness of pipelines up to level-1 is > 95%. ICCs are currently working on their scientific validation and in the full integration of the different pipelines in HIPE as tasks.

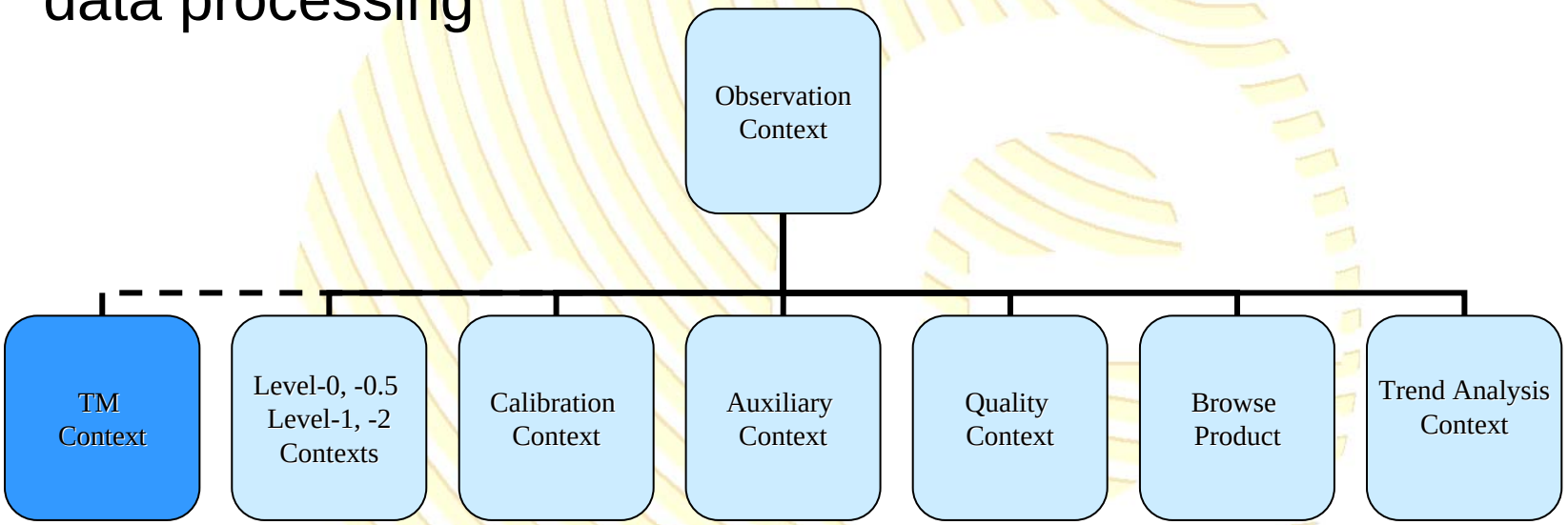
# The Herschel Data: Levels

- **Level 0 products:**
  - Raw science telemetry data in the form of (sub)instrument timelines
  - Equivalents: ISO's ERD (Edited Raw Data)
  - Examples: bolometer currents or chopper positions with time
- **Level 1 products:**
  - Science instrument timelines corrected for Herschel specifics
  - Equivalents: ISO's SPD (Standard Processed Data)
  - Examples: bolometer total fluxes in absolute physical units
- **Level 2 products:**
  - Fully calibrated science data ready for analysis
  - Equivalents: ISO's AAR (Auto-Analysis Results) and Spitzer's BCDs (Basic Calibrated Data)
  - Examples: flux-calibrated images with astrometrical information, base-line corrected spectra, spectral mapping data cubes
- **Level 3 products:**
  - Publishable results combining different observations. E.g. line fluxes or mosaics
  - Equivalents: ISO's HPDP (Highly Processed Data Products) and Spitzer post-BCDs

HERSCHEL SPACE OBSERVATORY

# The Herschel Data: Contexts

Data products are organized in “contexts”, which are wrappers of information with everything needed for the data processing

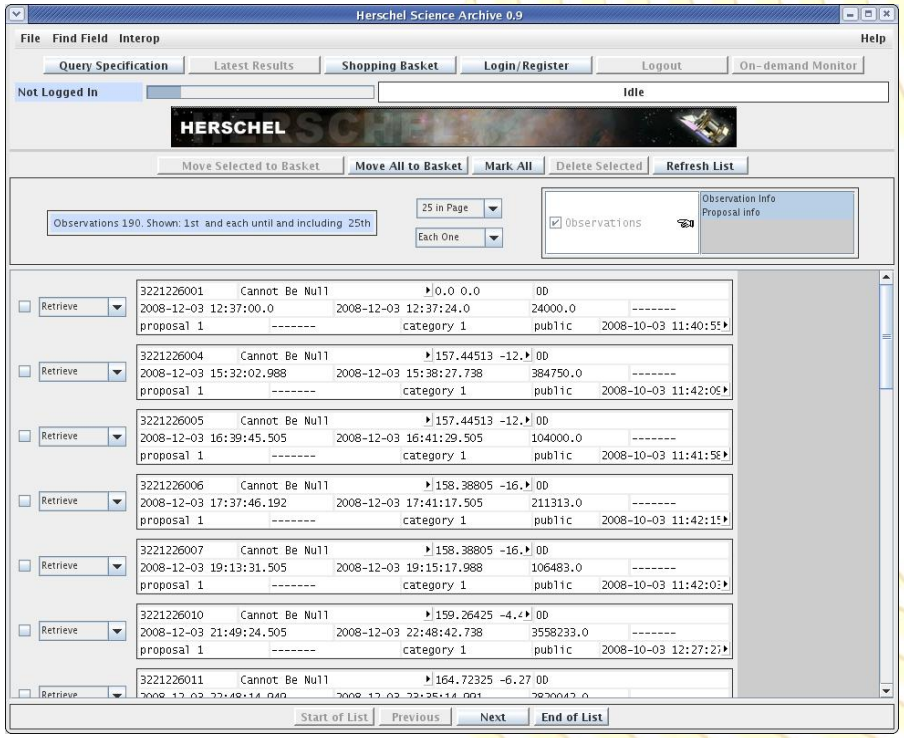



- The Observation Contexts contain all the calibration information, quality control, pointing data, onboard logs and science data that is needed to process any given observation.
- HIPE and the DP pipelines are programmed to handle and extract automatically the information needed from these products

# The Herschel Data: Classes

- **Any variable in the Herschel DP system belongs to a Class**
  - A class is just a high-level “type” of object or product
  - It uses the advantages of the Object Oriented programming, which allows dynamic association of tasks with the variables according to their classes.
  - All available classes are identified in Herschel DP and documented
- **Examples of classes:**
  - Numerical classes: Int1D, Bool2D, Double1D. They avoid syntax errors.
  - ObservationContext, level0, level1 and level2context classes: HIPE will know which pipeline tasks can be applied to them.
  - Timeline product, TableData product classes. Basic data assemblies typically of non-processed data.
  - simpleImage, simpleCube and simpleSpectrum: HIPE knows how to display them, which tasks are appropriate and how to export them to FITS

# The Herschel Data: Archive




## Herschel Science Archive

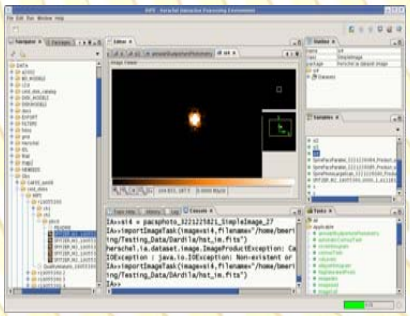
- Herschel data is stored in the archive after pipeline processing
- They can be retrieved:
  - In FITS format
  - Into a Herschel DP session for real-time browsing
- The data in the archive is reprocessed following each new major version of the pipeline
- Proprietary period: 1 Year
- It can contain user highly-processed data

HERSCHEL SPACE OBSERVATORY

# Data ingestion and storage

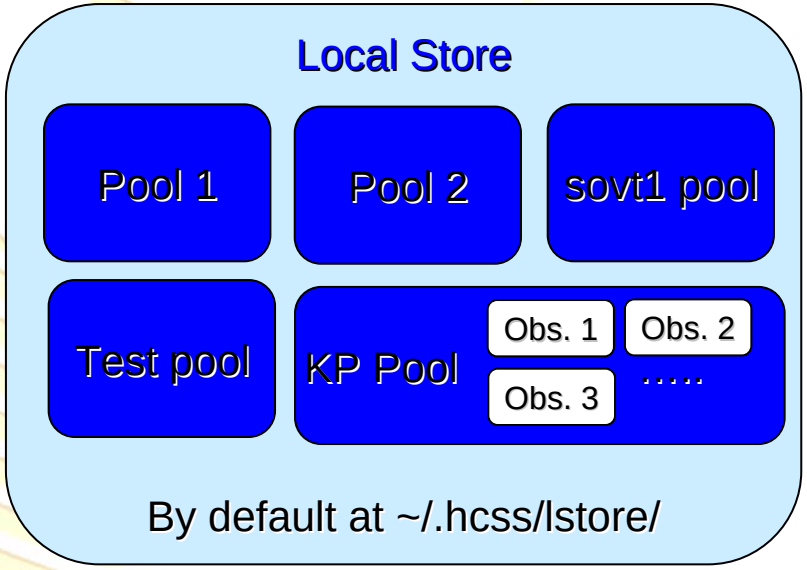
After data has been downloaded from the archive, they can be stored in a so-called local Store , which contains different Pools (mini data-bases) for different projects

HSA



Herschel DP session

1. Create a local store in a user's machine first time and then create pools in it afterwards




2. Register a pool from the Local store in a HIPE session, browse it with the Product Access Layer (PAL) to use its data, with:

```
lstore = LocalStoreFactory.getStore("sovt1")
store  = ProductStorage()store.register(lstore)
```

HERSCHEL SPACE OBSERVATORY

# HIPE software objectives

- **HIPE means Herschel Interactive Processing Environment**

- **HIPE is provided to the community with these objectives:**
  - To facilitate interaction at each pipeline step to optimise the scientific product generation
  - To provide interactive tools to generate products that capture all knowledge on spacecraft and instrument behaviour, calibration and Herschel data particularities
  - To provide a Herschel specific framework
    - For users to develop their own data reduction tools
    - To share data reduction tools with other users
  - Perform science analysis of the data, taking into account Herschel data characteristics

HERSCHEL SPACE OBSERVATORY

# HIPE

The welcome page brings you to all different areas of the environment

# The Herschel Data Processing software

- **Integrated framework for data retrieval, reduction and analysis with the latest packages for the removal of instrument and spacecraft artifacts**
  - Integrated and easy access to Herschel specific data structures
  - Similar FITS database and search capabilities as ESO's 'gasgano'
  - Similar variable-task association, GUI, automatic variable outline and built-in Help system as the 'IDL Development Environment' (IDLDE).
- **Programmed in Java and Jython (python over java)**
  - License free and platform independent
  - The python/jython provides a flexible scripting language
  - The java underneath provides powerful resources for highly demanding tasks. It also has many available numerical routines online

HERSCHEL SPACE OBSERVATORY

# The Herschel Data Processing software

- **Highly customizable and programmable software**
  - Flexible: allows interchangeable interactive (console) and batch-mode (script) processing. All HIPE sessions produce re-playable scripts
  - Highly extendable: new external functions easy integrated as tasks
  - Open source: all original code is provided with the installer
  - Highly customizable Graphical User Interface (GUI)
- **Compatible with all usual FITS-based processing software and integrated with the latest VO tools**
  - Capable of input/output of standard FITS files to IDL, IRAF, MIDAS, ds9 or any other FITS viewer compatible with the standard
  - Integrated with the VO tools TopCat, Aladin, VOSpec, SPLAT-VO, etc

HERSCHEL SPACE OBSERVATORY

# HIPE scientific tools and modules (i)

- **Pipeline scripts to be run interactively:**
  - Modification of processing parameters and routines
  - Data visualization at each step
  - Usage of different calibration products versions
  - Modification of the pipeline script and pipeline modules

- **Miscellaneous interactive analysis tools**
  - Defined by instrument specialists to improve data reduction and analysis in the specific cases when automation does not give good enough results
  - Particularly important in early phases of the mission (steep learning curve on instrumental effects understanding and calibration)

# HIPE scientific tools and modules (ii)

- **Advanced data processing modules**
  - Processing and analysis of level-2 data
  - Definition and development agreed by Herschel Science Centre, Instrument Control Centre's and NASA Herschel Science Center
  - When applicable to be incorporated in the pipeline for automatic processing
  - Currently in development:
    - Image visualisation, arithmetics and analysis
    - Cube data visualisation and analysis
    - Spectrum visualisation, arithmetics and analysis
    - Herschel Point Source Extraction
    - HIFI specific advanced HIPE modules
    - PACS specific advanced HIPE modules
    - SPIRE specific advanced HIPE modules

# HIPE deliveries

- **Immediately after this workshop a first set of Herschel DP documentation will be made available**
- **A beta version of HIPE will be made available to the KP consortia in March 2009**
  - KP observers to get familiar with HIPE and start preparing for KP data reduction
- **Keep in mind that pipelines, calibration files and data reduction algorithms are preliminary:**
  - Only after the Performance Verification phase we will know the behaviour of the spacecraft and instruments
  - AOTs, pipelines, data reduction algorithms, dataflows, documents, calibration and products will change
- **First HIPE delivery to the general community 5-6 months after launch**

# HIPE is today a beta version


- It shows already the main features and its design concept
- However, many details are still in development
- It is shown to the future users to allow fast acquaintance with the system and to get prompt feedback about the tool

HERSCHEL SPACE OBSERVATORY

# Summary

- Important things you should remember about..
- The Herschel data
  - The data products: structure, processing levels and file contents
  - Origin: Herschel Operations, dataflow and pipelines
  - End: the Herschel Science Archive and proprietary periods
- The Herschel Data Processing (DP) software
  - Integrated framework for data retrieval, reduction and analysis with latest software for instrument and spacecraft artifact removal
  - Next generation software: Flexible (interactive/batch-mode), extendable, fully integrated and highly customizable
  - Compatible with all usual FITS-based processing and latest VO-compliant software
  - Programmed in Java/Jython: easy scripting with strong numerics
  - Currently being used and developed by HSC, the ICCs and NHSC

HERSCHEL SPACE OBSERVATORY

# Summary

- Status and plans of the Herschel DP software
  - Current status of the development: beta version 0.6.6
  - Medium and Long-term plans and community involvement

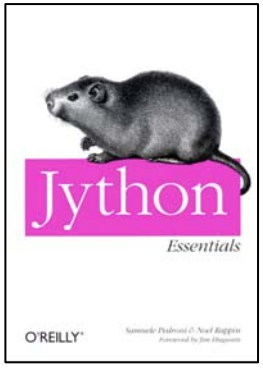
- HIPE: Herschel Interactive Processing Environment
  - HIPE: Herschel Interactive Processing Environment
  - Basic functionalities, windows and views
  - Launching customized pipelines from HIPE
  - Doing data analysis with HIPE
    - Image analysis: map making, photometry extraction
    - Spectral analysis: spectral line fitting, FFT
    - Spectral Cube analysis: spectra selection and averaging

HERSCHEL SPACE OBSERVATORY

# Supplementary material

# What can you do to get prepared?

- Read the documentation
  - Howto's
  - Herschel DP User Manual
  - Herschel DP Reference Manual (in order of specificity)
- Get some acquaintance with jython and java
  - Interesting sources of information:
    - *Jython Essentials*, by Samuelle Pedroni & Noel Rappin, O'Reilly
    - www.jython.org
    - www.python.org
    - www.sourceforge.org (for useful free java libraries)
    - Python for astronomers
      www.cfa.harvard.edu/~ebresser/python




- Buy new computers
  - If your Key Project has large mosaics or many observations

# Advanced DP HIFI modules

- **HIFI OTF map analysis tools**
  - Interactively extract a full spectrum from an OTF sample (including spatial averaging with weights)
  - Put adjusted spectra into an OTF data stream or into a cube
  - Status: Proto-typed
- **HIFI OTF to Cube tool**
  - Generate data cube from HIFI Spectra obtained during an HIFI OTF mode observation
  - Status: Proto-typed (Note: Hyperspectral Working Group has a proposal for a Herschel SpectralCube. HIFI Cube will be based on this)
- **HIFI Standing Wave Analysis**
  - Identify standing waves
  - Interactively mask spectral fragments (possibly containing spectral lines)
  - Convert standing wave frequencies to system lengths
  - Accept lengths and create appropriate standing wave
  - Optionally remove identified standing waves
  - Status: Conversion from ISO tool to JAVA underway and testing
- **HIFI Side-band de-convolution, visualisation and analysis tool**
  - Status: Tool in testing phase. Requires appropriate data.
- **HIFI Bands 6,7 Beam Synthesis**
  - Optimally combine small bands 6,7 maps into one final spectrum.
  - Status: Not yet started.

# Advanced DP PACS modules

- **PACS Integral Field Spectroscopy to Cube**
  - Combine spectra from line or range spectroscopy observations (25x16 spectra in time) into a set of 25 (spatial) spectra
  - Status: Definition of user requirements and discussion of algorithms. Prototypes for several steps (rebinning, clipping, projection)

- **PACS photometric observations to maps**
  - Combine detector level-1 products representing the set of scans or rasters across the sky and combine into a MAP product (plus a noise map) which gives flux as a function of sky position
  - Status: Definition of user requirements and discussion of algorithms. Prototype of MADmap implementation. Pipeline integration and testing under way.

# Advanced DP SPIRE modules

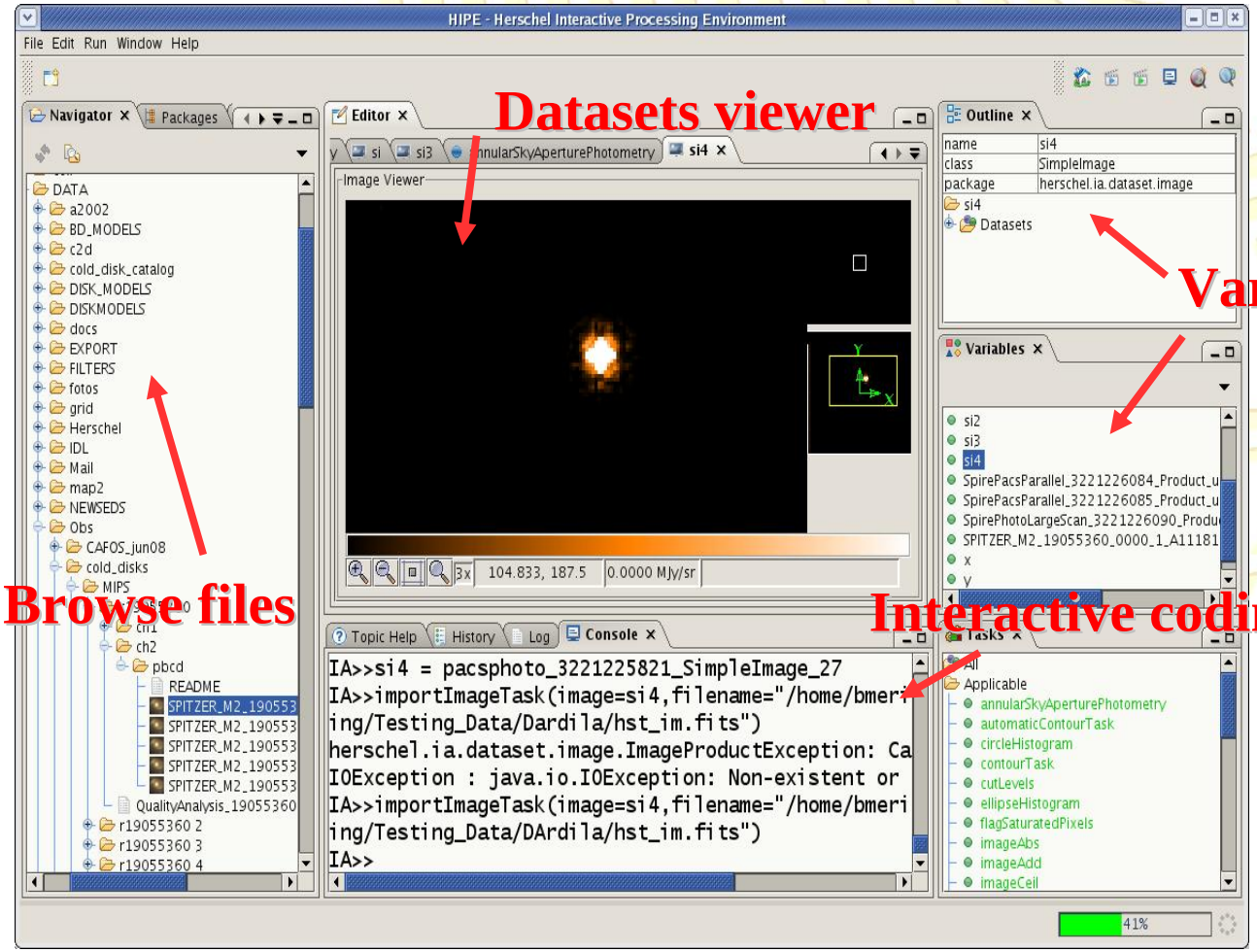
- **SPIRE scans to map**
  - Combine detector timelines representing the set of scans across the sky and combine them into a MAP product which gives flux as a function of sky position
  - SPIRE mapmaking algorithm, based on the CMB code MADmap. See papers
    - "Determining the optimum scan map strategy for Herschel-SPIRE using the SPIRE photometer simulator", 2007, Waskett et al., MNRAS 381, p. 1583
    - Chanial et al., PASP submitted, Mapmaking Algorithm Selection for Herschel Spectral and Photometric Imaging Receiver (SPIRE) Scan Maps
  - Status: Module completed and tested, mapmaking pipeline integration and testing underway. Also naive mapmaker complete and available for scan maps.
- **SPIRE jiggle to map**
  - Provision of a SPIRE data processing module to take detector timelines and combine them into a MAP product which gives flux as a function of sky position
  - Status: Naive mapmaking module for jiggle maps complete and tested, jiggle-mapmaking pipeline integration and testing underway.
- **SPIRE FTS to cube**
  - Provision of a data processing module to combine spectra from a spectrometer observation into a spectral data cube
  - Status: Discussion of requirements

HERSCHEL SPACE OBSERVATORY

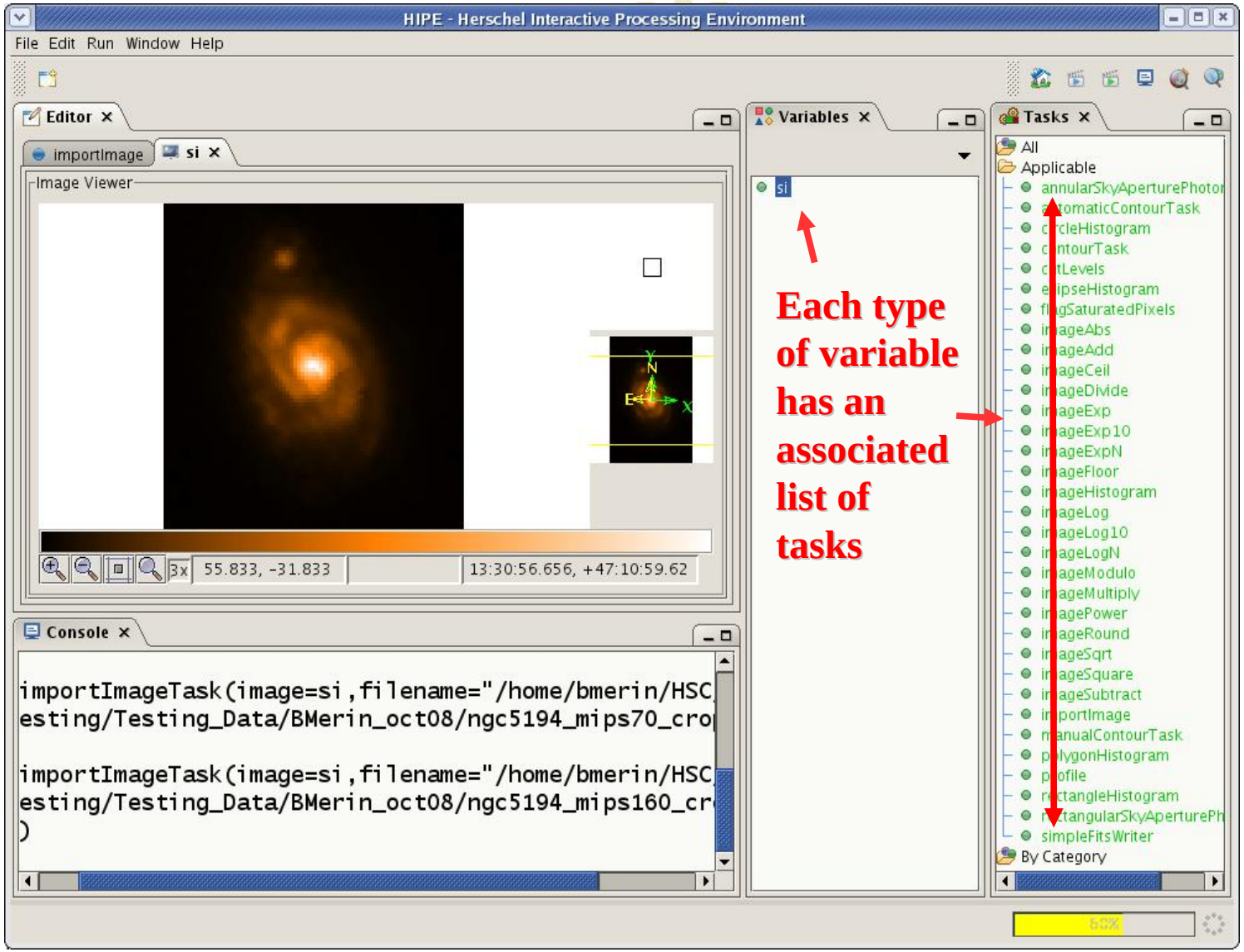
# HIPE tour

# HIPE: The Workbench

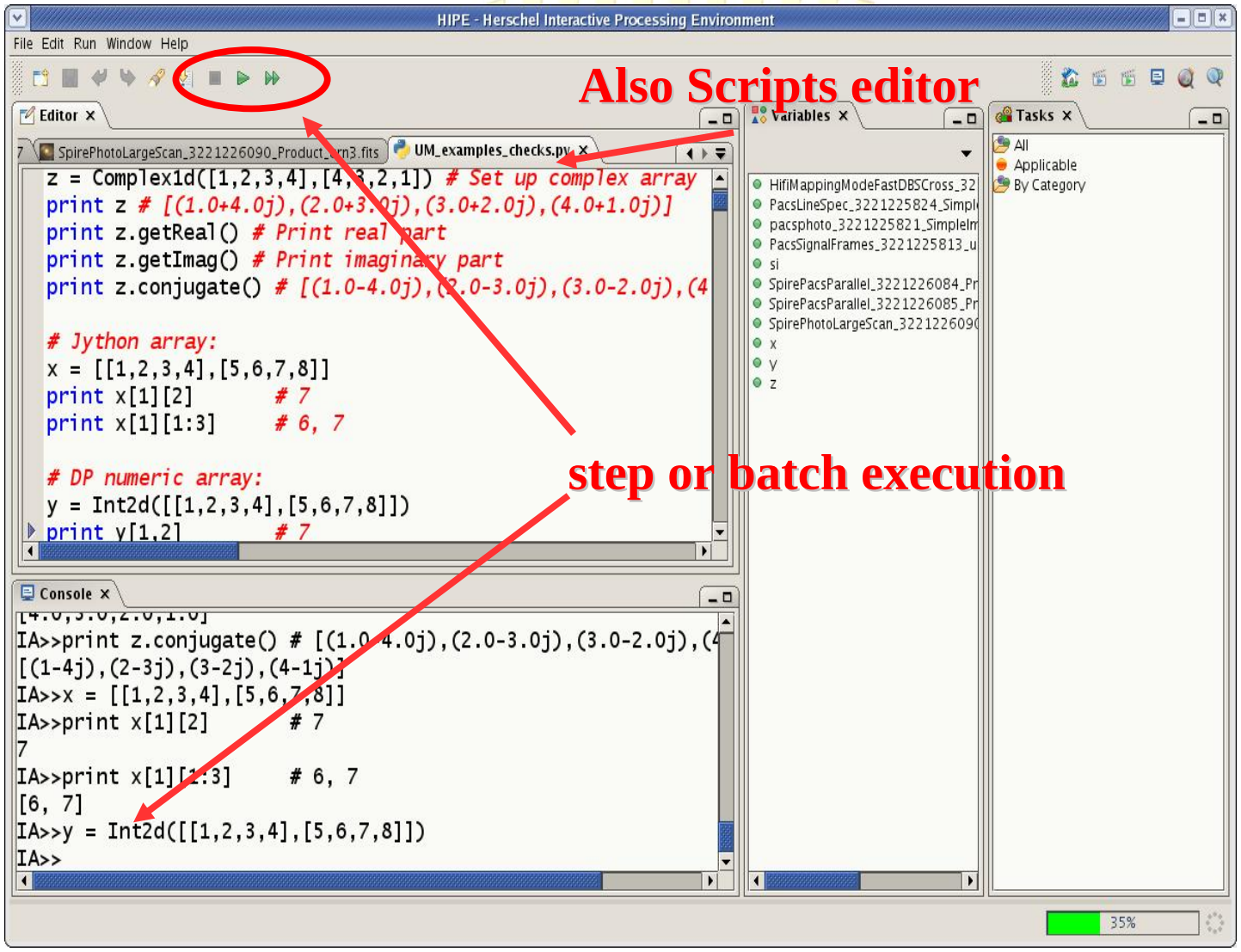
Integrated framework for jython/java coding, data visualization and launching of customized instrument pipelines



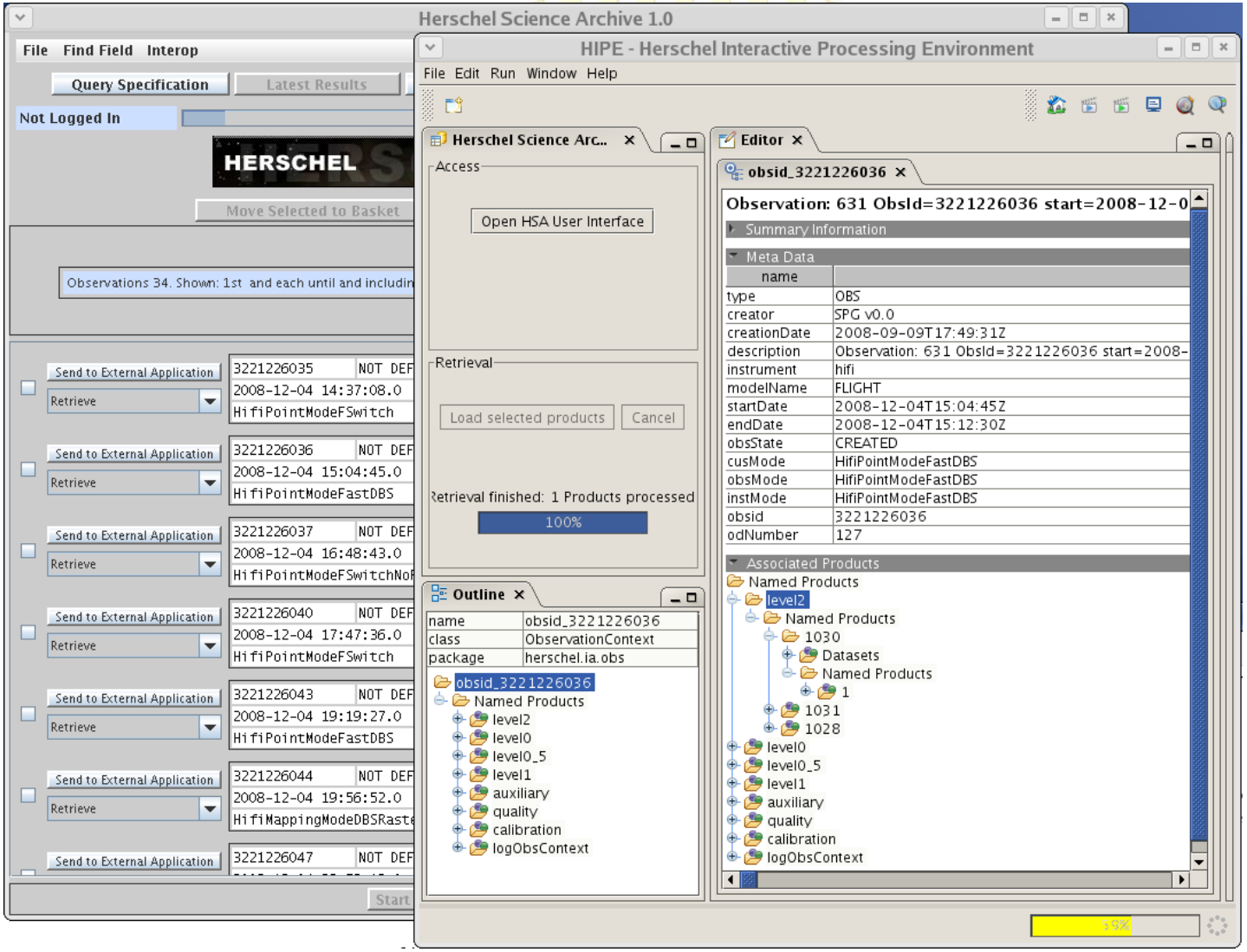
# Dynamic association of variables and tasks/functions



# Powerful editing and running of jython scripts

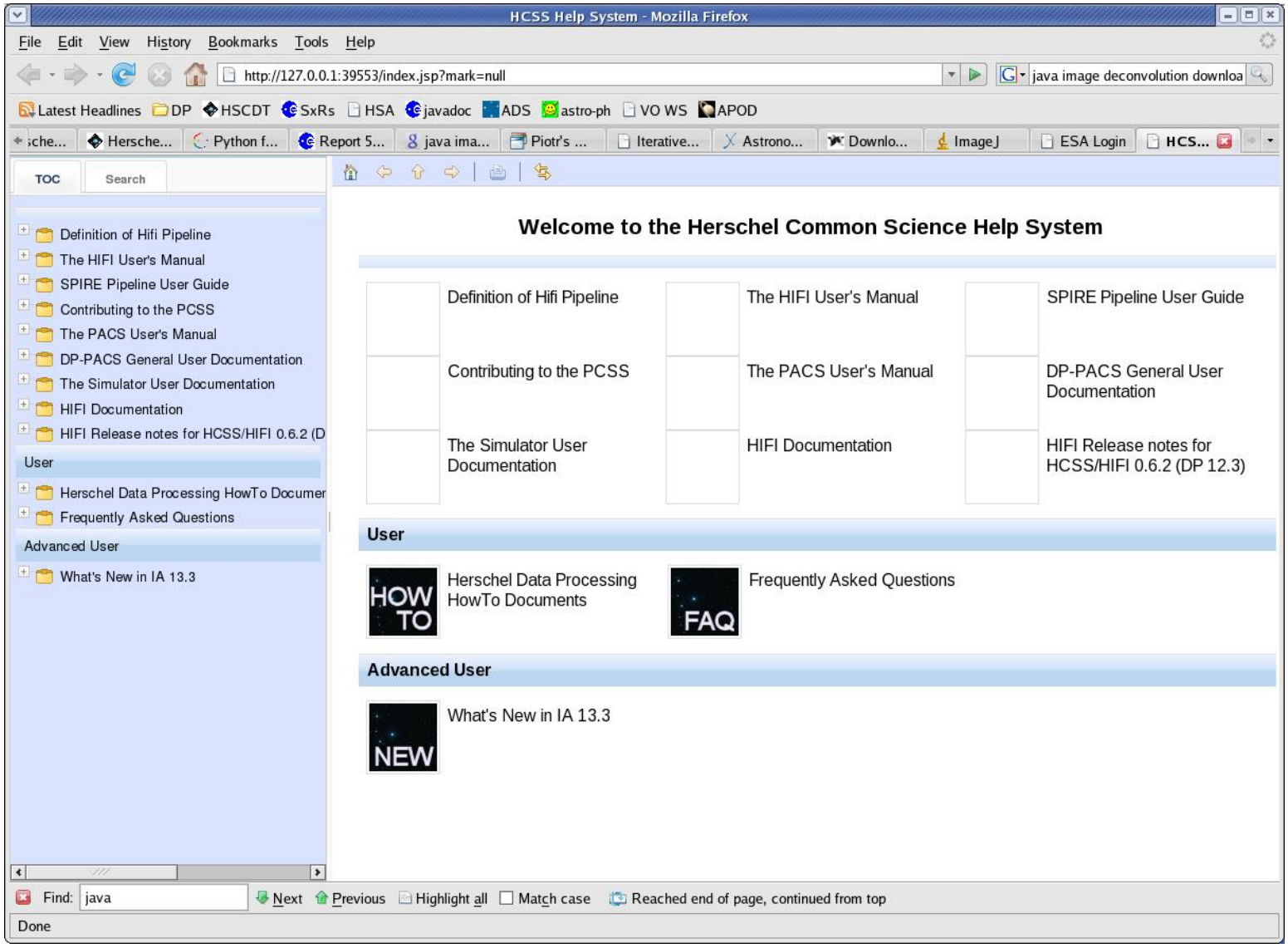


# HIPE: Access to the Archive



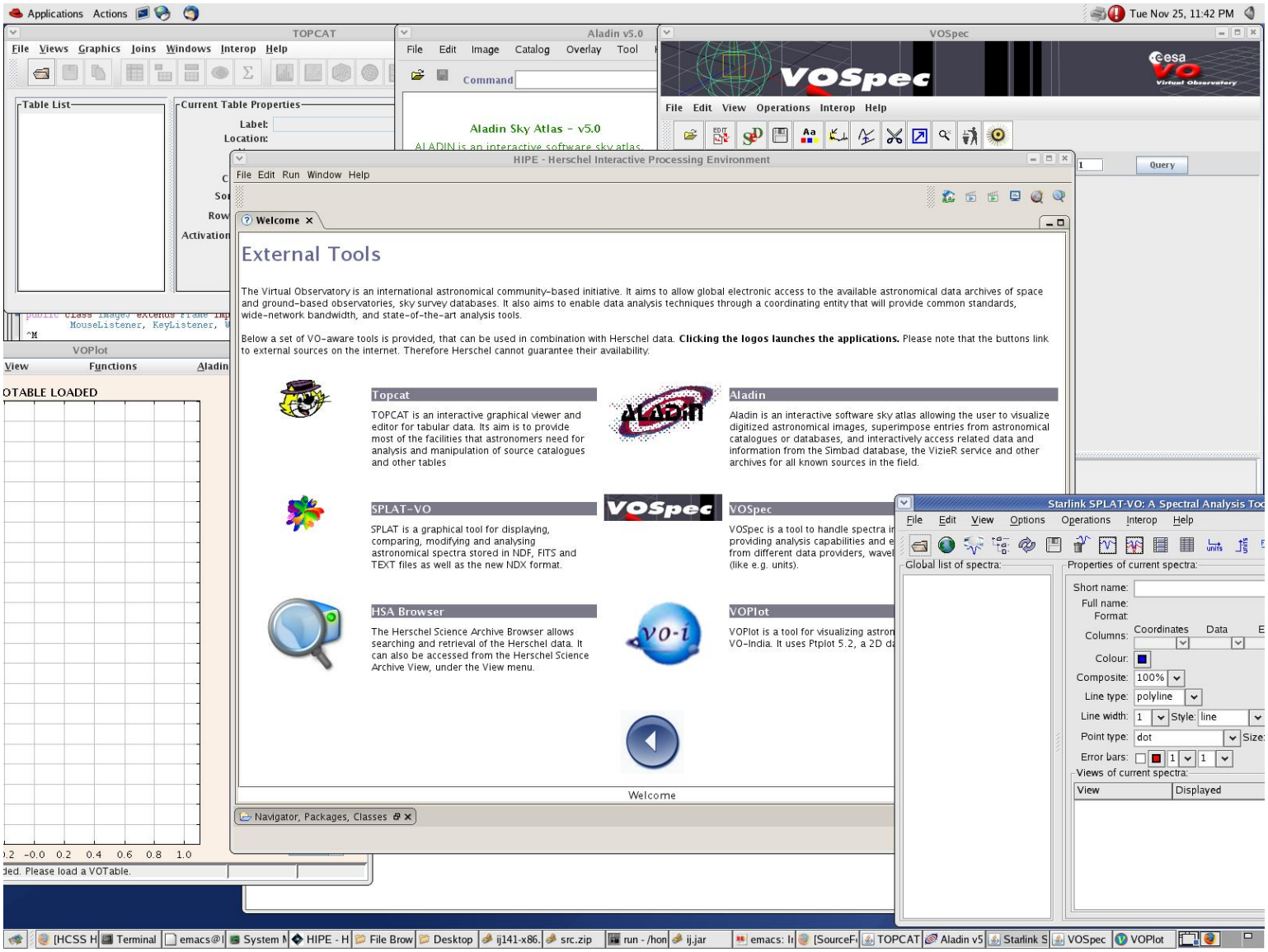

HERSCHEL SPACE OBSERVATORY

# HIPE: Online help




HERSCHEL SPACE OBSERVATORY

# HIPE: external VO-tools




HERSCHEL SPACE OBSERVATORY