The Herschel Data Processing System: History, Status and Plans

Stephan Ott

Herschel Science Data Processing Development Manager
Herschel Science Data Processing Coordinator

Herschel Science Centre Development Team, ESA/ESTEC

on behalf of the Herschel Data Processing Development Team

HERSCHEL SPACE OBSERVATORY

Outline

- Preface
- History
- Scope and Methodology of Herschel Data Processing
- Aspects of Herschel Data Processing
- Plans
- Summary

HERSCHEL SPACE OBSERVATORY

Preface
Pre-Launch Data Processing Workshop and SOVT

Data Processing Workshop and SOVT-2 (1)

- Originally this Pre-Launch Data Processing workshop was scheduled for December as at this time the spacecraft was expected to be shipped to Kourou
- Due to various problems with the cryogenic system the original schedule could not be kept
- Up to one month ago, System Operational Verification Test 2 (SOVT-2) was planned for January 2009
- SOVTs are *the* major test campaigns for the Herschel Ground Segment, with activities for all parties at peak intensity

HERSCHEL SPACE OBSERVATORY

Data Processing Workshop and SOVT-2 (2)

- By advancing the test sequence by one month, Herschel Project could accommodate our long-standing wish to perform SOVT-2 in conditions that are close to in-flight conditions
- Among other benefits this means that that we will receive data that are much more realistic compared to those of previous test campaigns, a major advantage to derive better calibration files and to tune the data analysis algorithms
- The schedule for SOVT-2 coincided with the Pre-Launch Data Processing workshop

HERSCHEL SPACE OBSERVATORY

Data Processing Workshop and SOVT-2 (3)

- Support to the Herschel project for spacecraft testing has the highest priority, and the same experts who now need to support SOVT-2 on a compressed schedule were expected to prepare for the Pre-Launch Data Processing workshop
- Weighting all options, the workshop agenda was adapted to ensure that we can support simultaneously an effective workshop, and the vital SOVT-2 exercise. Everyone put in extra efforts to support these two major events as well as possible

We are certain that this workshop will be very interesting and fruitful for all participants. We have a wealth of software to present and demonstrate. A cordial welcome on behalf of the Herschel Data Processing Development Team!

HERSCHEL SPACE OBSERVATORY

History of Herschel
Data Processing

Herschel Data Processing History (1)

- When ESA's Science Programme Committee approved the Science Management Plan in 1997 it was considered sufficient to provide astronomers with raw data and software tools to carry out the data processing, and no 'data products' were to be generated and delivered. The strategy adopted was to offer the means for the individual observer to generate any desired product him/herself.
- This also meant that the Herschel archive was not to be populated (except for with raw data) until in the post-operations phase, severely restricting the usefulness of Herschel data for the wider community during the Herschel mission.

HERSCHEL SPACE OBSERVATORY

Herschel Data Processing History (2)

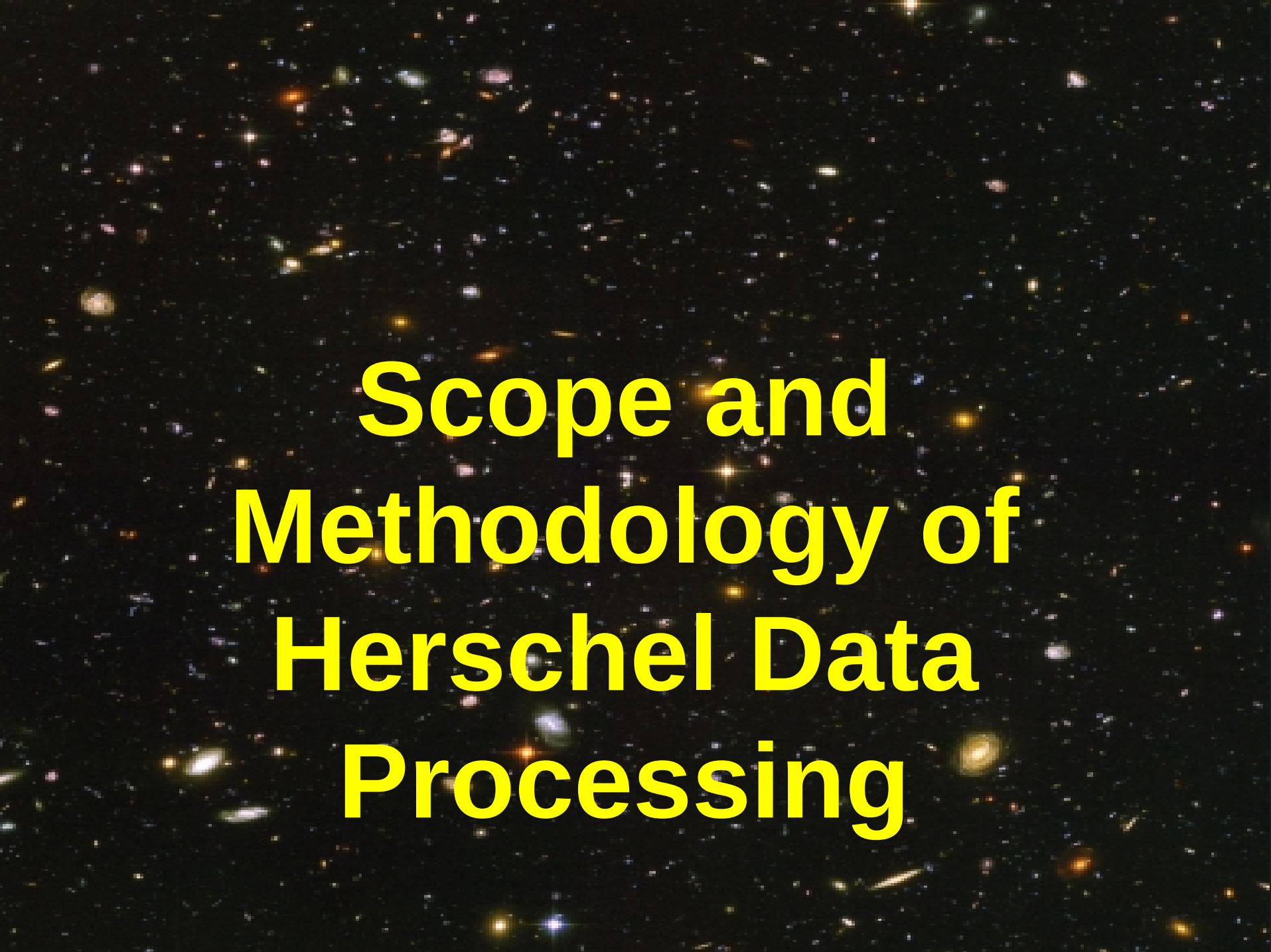
- 2002: Development of the common Herschel Data Processing framework started to support the data analysis for Instrument Level Tests
- At the same time it was realised that expectations of the astronomical community on what to expect from observatories and their data processing systems, data products, and archives have evolved; what was once considered acceptable is no longer considered palatable
- 2004: Management gave the final go-ahead to continue with the development of the common Herschel Data Processing framework

HERSCHEL SPACE OBSERVATORY

Herschel Data Processing History (3)

- 2005: ESA's Science Programme Committee approved additional funding for an "extended Data Processing" development, recognising that Herschel must offer its users, the astronomical community, comprehensive means, beyond observing opportunities, to do science. This permitted to award contracts to the Instrument Control Centres and to hire developers for the Herschel Science Centre to work on software for data products, algorithms, pipeline and observers' framework and support
- 2006: Recruitment of HSC staff and negotiations with ICCs started
- 2007: Contracts with ICCs were signed
- – Summer 2007: Development focussed to meet instrument experts' needs
- Summer 2007 – : Development started to address end-users' needs
- Summer 2008: First set of deliveries from ICCs received

HERSCHEL SPACE OBSERVATORY


Scope and
Methodology of
Herschel Data
Processing

Management and development methodology (1)

- **Herschel Science Centre (ESA), the Instrument Control Centres (HIFI, PACS and SPIRE) and NHSC jointly manage and contribute to the Herschel Data Processing system**

Bill Latter (NHSC)
Stephan Ott (ESA)
Data Processing
Calibration Scientist
Coordinator
Manager

Management and development methodology (1)

- **Herschel Science Centre (ESA), the Instrument Control Centres (HIFI, PACS and SPIRE) and NHSC jointly manage and contribute to the Herschel Data Processing system**

- **Very distributed development – we work 24/5 (and often 24/7) on Herschel Data processing development**

HERSCHEL SPACE OBSERVATORY

Management and development methodology (2)

Development of the Herschel Data Processing System is a significant effort

- HSC, ICCs and NHSC nominally contribute to the Herschel Data Processing framework
  - HIFI 1.0 f.t.e for coding efforts
  - HSC 6.5 f.t.e for coding efforts
  - NHSC 1.5 f.t.e for coding efforts
  - PACS 1.0 f.t.e. for coding efforts
  - SPIRE 1.0 f.t.e. for coding efforts
- Additionally, ICCs provide algorithms to remove instrumental and satellite effects and tools to calibrate data
- Additionally, ESA funds ICCs 3*10 staff years to provide user friendly functionality (advanced data processing algorithms and interactive tools)

HERSCHEL SPACE OBSERVATORY

Management and development methodology (3)

Current main responsibilities for Data Processing framework

- **HIFI**
  documentation framework; spectrum and line list dataset/product; dataflow
- **HSC**
  GUI framework (HIPE, JIDE); pipeline and quality control framework; low level I/O; dataset/product infrastructure; generation of DP-AM, including installer and updater framework
- **NHSC**
  numerics
- **PACS**
  display; contour plot and history; image, cube, stack and point source dataset/product; explorer infrastructure and default plug ins; binstruct
- **SPIRE**
  plot; product access layer; access

In addition to their framework contributions, the ICCs are responsible for development of their pipelines and instrument calibration

HERSCHEL SPACE OBSERVATORY

Management and development methodology (4)

- The Data Processing system is part of Herschel Common Science System, and therefore coded in Java/Jython for platform independence
- Java with scripting/prototyping in Jython gives full Java power to developers and expert users while maintaining user friendliness via GUIs and abbreviated Jython commands
- The general astronomical community that is using the system only occasionally is shielded from the object oriented nature of Java and Jython by the GUI based / data driven data access
- Legacy programs or VO services can be called using VO protocols

HERSCHEL SPACE OBSERVATORY

Management and development methodology (5)

- All tools for data reduction and analysis, e.g. also the “Instrument Calibration” and “Trend Analysis” systems are part of the general Data Processing system
- The pipeline system that automatically generates the data products is a subset of the Interactive Analysis system
- The system is extensively used by instrument specialists during instrument development and testing
- The Data Processing software will be distributed as source code, and users can write and contribute their own routines

HERSCHEL SPACE OBSERVATORY

Management and development methodology (6)

- Only high level requirements and plans are drawn up and maintained
- An iterative development approach is followed to adapt to changing and unforeseen user needs
- Prioritisation (and re-prioritisation) is jointly performed by all partners contributing to the Herschel Data Processing system
- The wishes of the key program users, and later the whole community, will be taken into consideration

HERSCHEL SPACE OBSERVATORY

(Some) Software
Aspects of Herschel
Data Processing

Scope of Herschel Data Processing (1)

Data Processing (DP) = Herschel Interactive Processing Environment (HIPE) + Standard Product Generation (SPG) + Quality Control Processing (QCP) + supporting tools

- Integrated, easy to use, license free, well tested and documented Data Processing system available free of charge to the astronomical community
- Basic functionality must be available off-line
- Tools to remove instrumental and satellite effects from Herschel Data (generation of basic level 1 products)
- Capability to further process basic products to a higher level so that scientific analysis can be performed (generation of advanced level 2 products)
- Higher level products must be VO compliant

HERSCHEL SPACE OBSERVATORY

Scope of Herschel Data Processing (2)

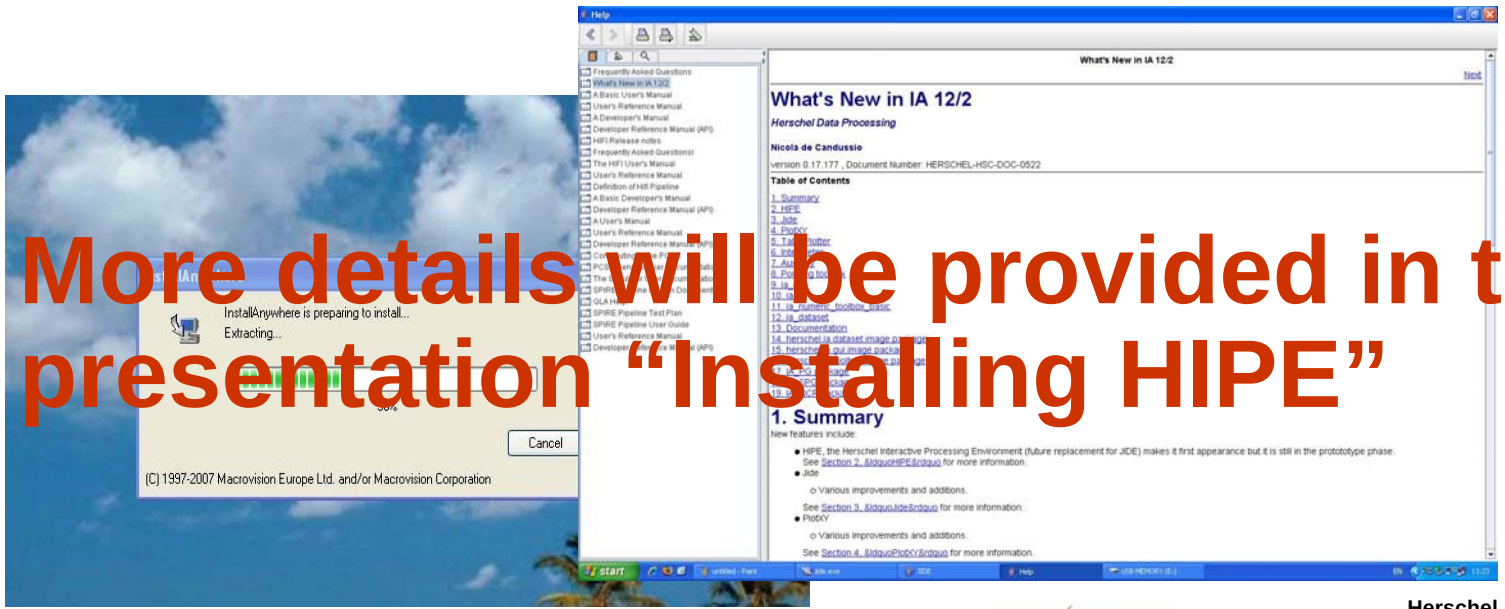
Data Processing (DP) = Herschel Interactive Processing Environment (HIPE) + Standard Product Generation (SPG) + Quality Control Processing (QCP) + supporting tools

- Systematic pipeline processing, quality control and distribution of all generated data
- Pipelines are scripted HIPE tasks that are executed on the ESAC grid
  - User can choose most suitable pipeline and adjust parameters for on-demand reprocessing
  - Pipelines can also be run interactively at users' computer
- Tools for support of operations (calibration, trend analysis, instrument monitoring)
- Tools to enable astronomers to perform Herschel specific science
- Export to FITS for ingesting in other Data Processing systems

HERSCHEL SPACE OBSERVATORY

Software aspects – DP ALL (1)

- Software that was independently developed by the four centres (3 ICCs and HSC) had to be re-integrated and bundled into one distributable and easily installable package (DP-ALL)
- The installer is a single downloadable and executable file.
- The user clicks on it and the installation begins.
- The installers are available for Windows XP, Linux, Mac OS X

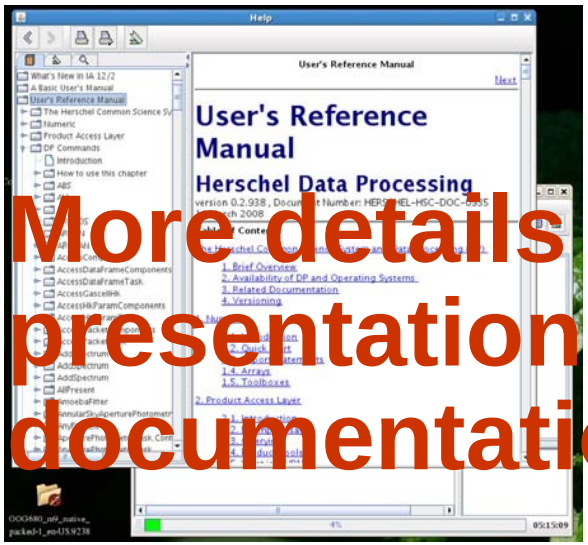


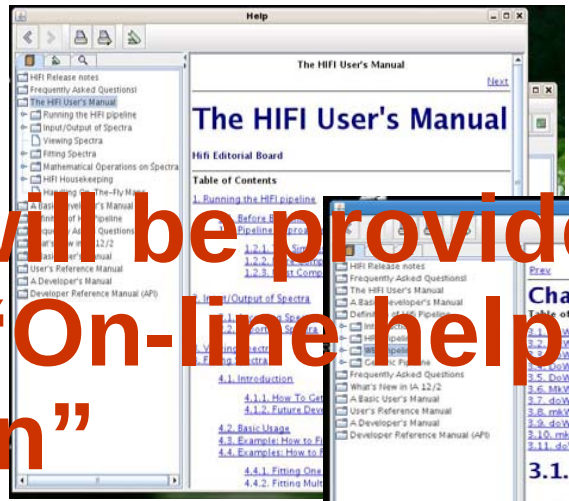
More details will be provided in the presentation "Installing HIPE"

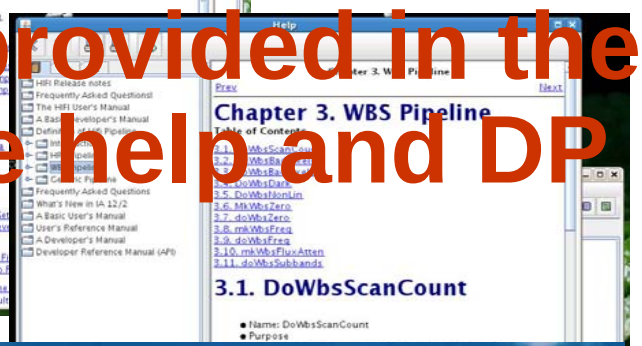
HERSCHEL SPACE OBSERVATORY

Software aspects – DP ALL (2)

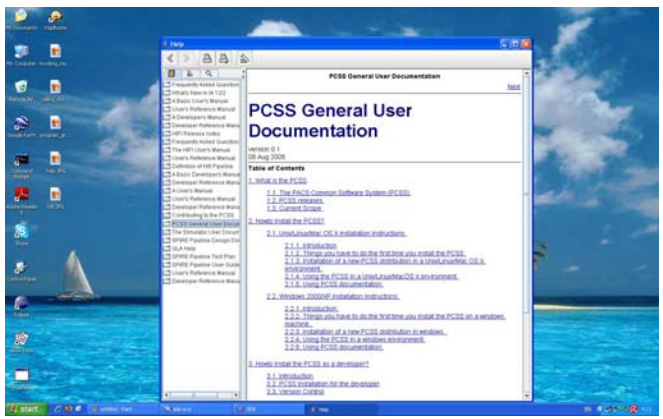
- **Documentation for the framework and all instruments available on-line**

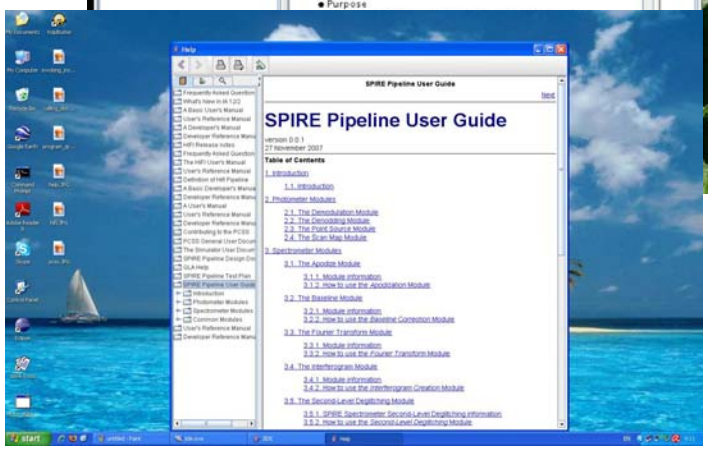






More details will be provided in the presentation "On-line help and DP documentation"

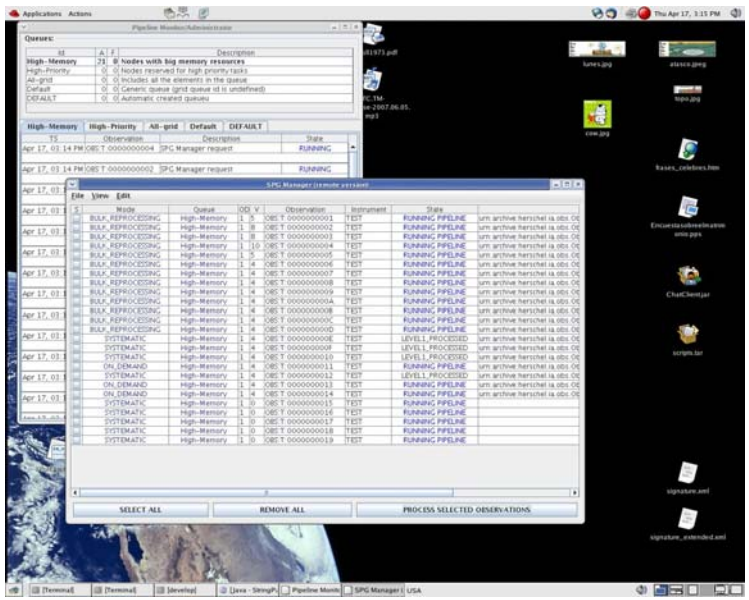


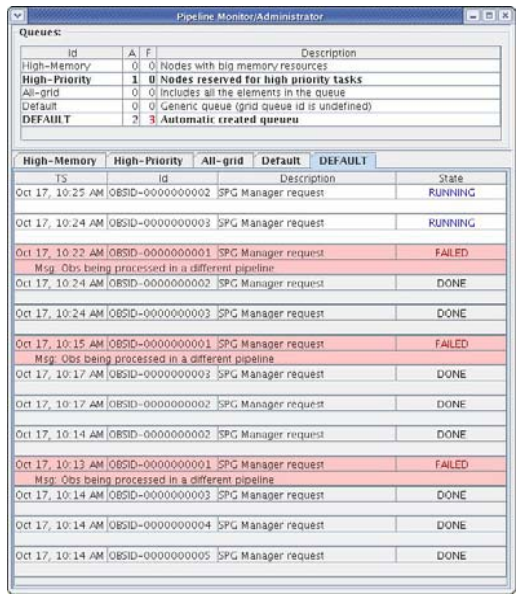


HERSCHEL SPACE OBSERVATORY

Software aspects – Pipeline (1)

- Pipeline is executed on the ESAC Grid to produce Herschel Products to different reduction levels
  - Level 0 raw data
  - Level 1 instrumental and satellite effects removed
  - Level 2 scientific analysis can be performed
- At the end of the pipeline processing, products are ingested into the Herschel Science Archive for immediate availability

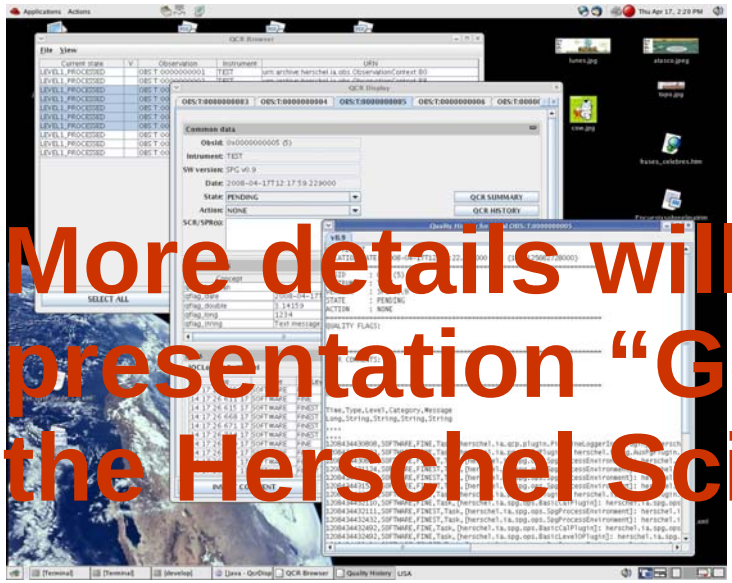


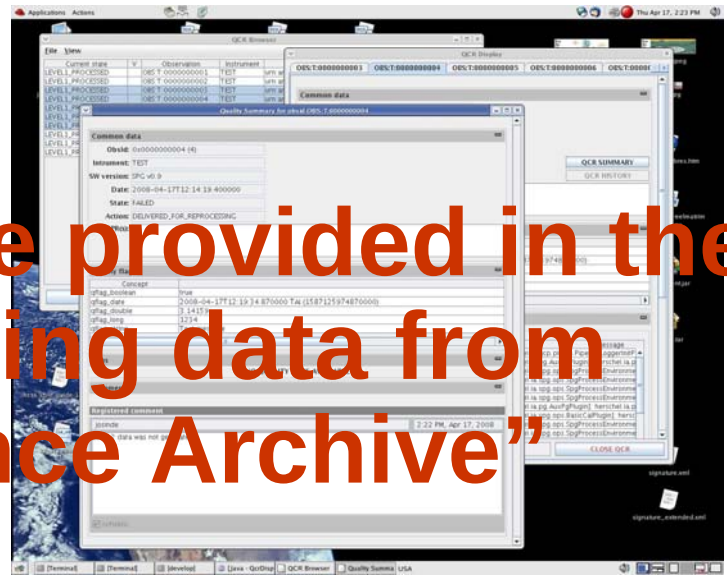


HERSCHEL SPACE OBSERVATORY

Software aspects – Pipeline (2)

- Data quality control is a combination of automatic screening and manual inspection
- Each observation is manually inspected at least once
- Quality control reports are electronically distributed to experts
- After completion of the quality assessment, quality control information is made available to the users through the HSA





More details will be provided in the presentation "Getting data from the Herschel Science Archive"

HERSCHEL SPACE OBSERVATORY

Software aspects – Herschel Interactive Processing Environment HIPE (1)

From a developer / expert environment

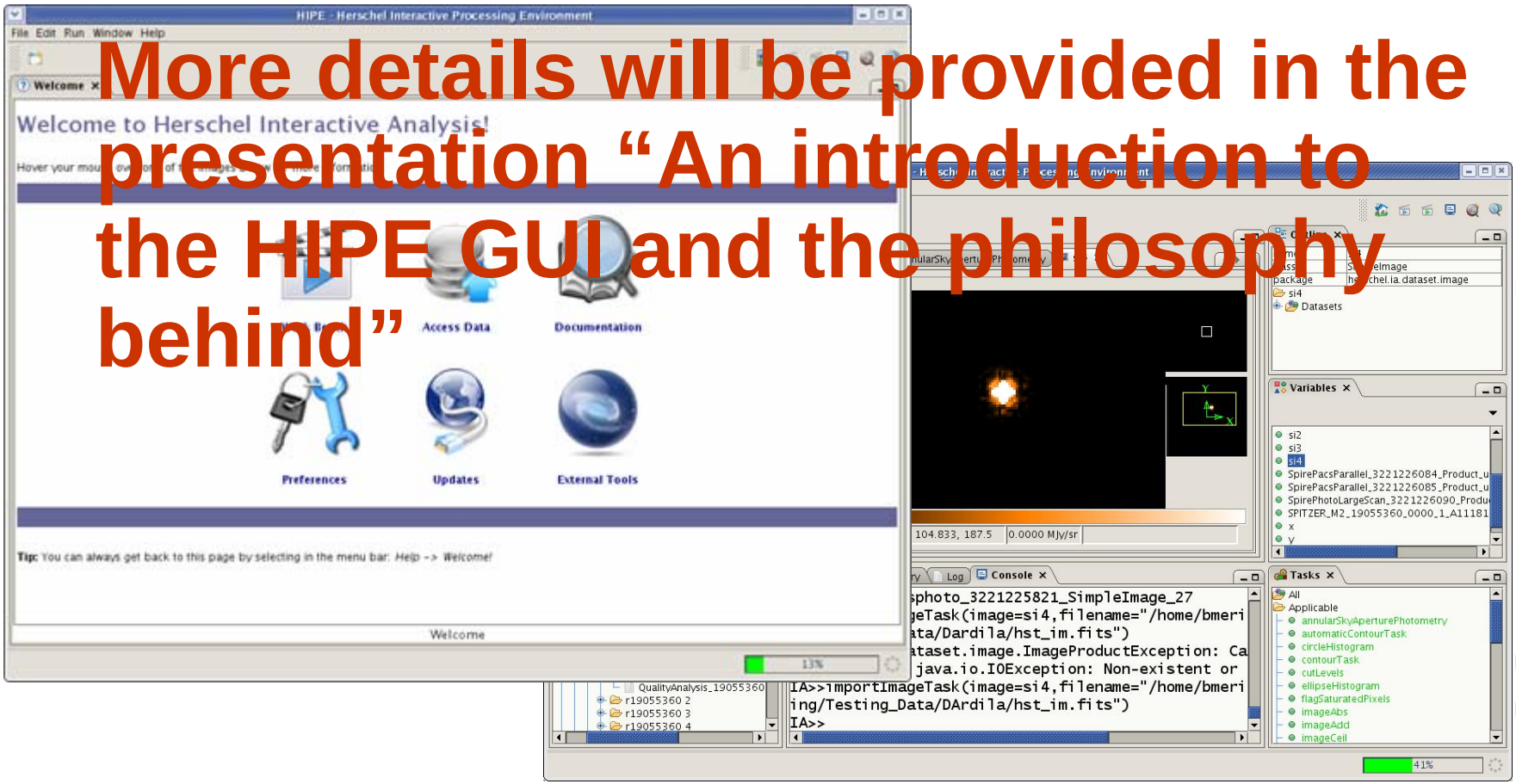
- Script driven
- Command line based (requires good knowledge of the system)
- Tools oriented (loose graphical integration)
- Stand alone documentation tool

to an end user oriented environment:

- GUI based
  - Data view (acquisition, inspection, edition)
  - Function view (library of functions)
  - Editor view (editing of data and calling of functions)
  - Navigator view (files, script, classes, libraries)
- GUI/Script driven
  - Start up guide, help at glance
  - Selection at fingertip (filtering on applicable options based on data selection)
  - Script generation from user actions

HERSCHEL SPACE OBSERVATORY

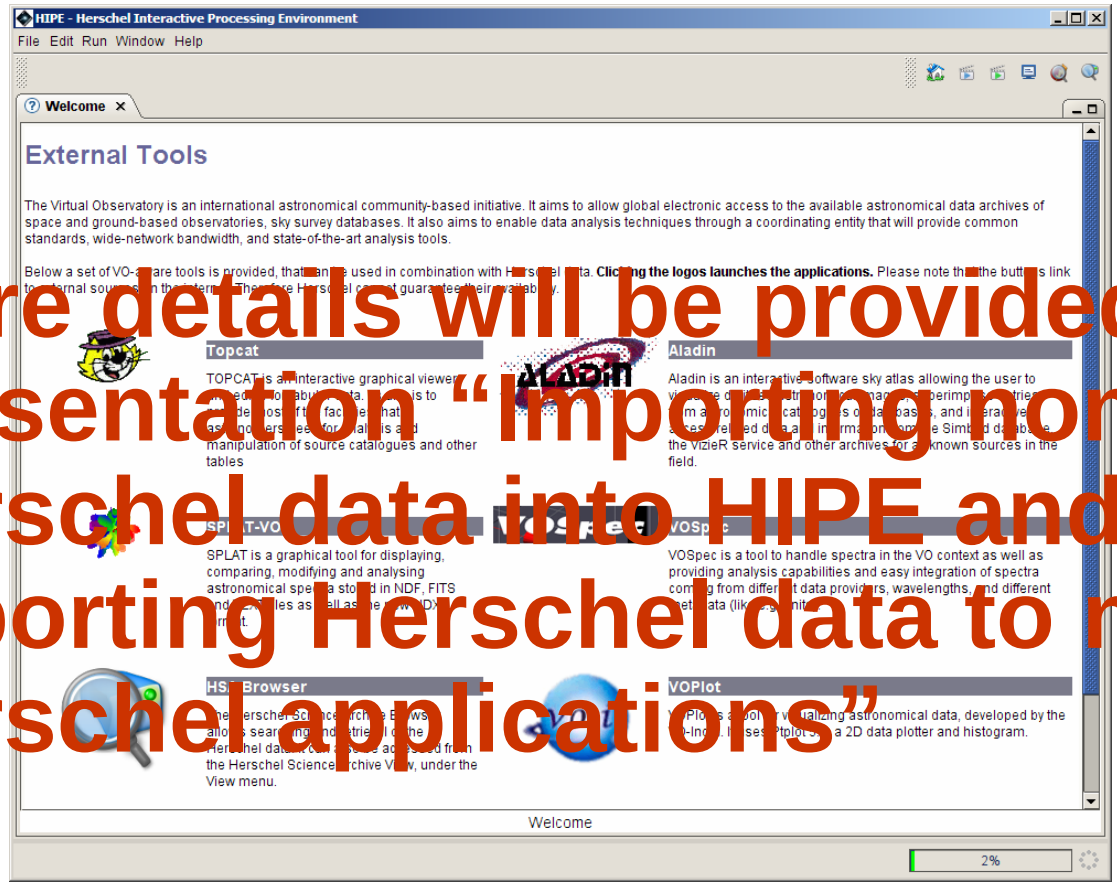
Software aspects – Herschel Interactive Processing Environment HIPE (2)



More details will be provided in the presentation "An introduction to the HIPE GUI and the philosophy behind"

HERSCHEL SPACE OBSERVATORY

Software aspects – Herschel Interactive Processing Environment HIPE (3)

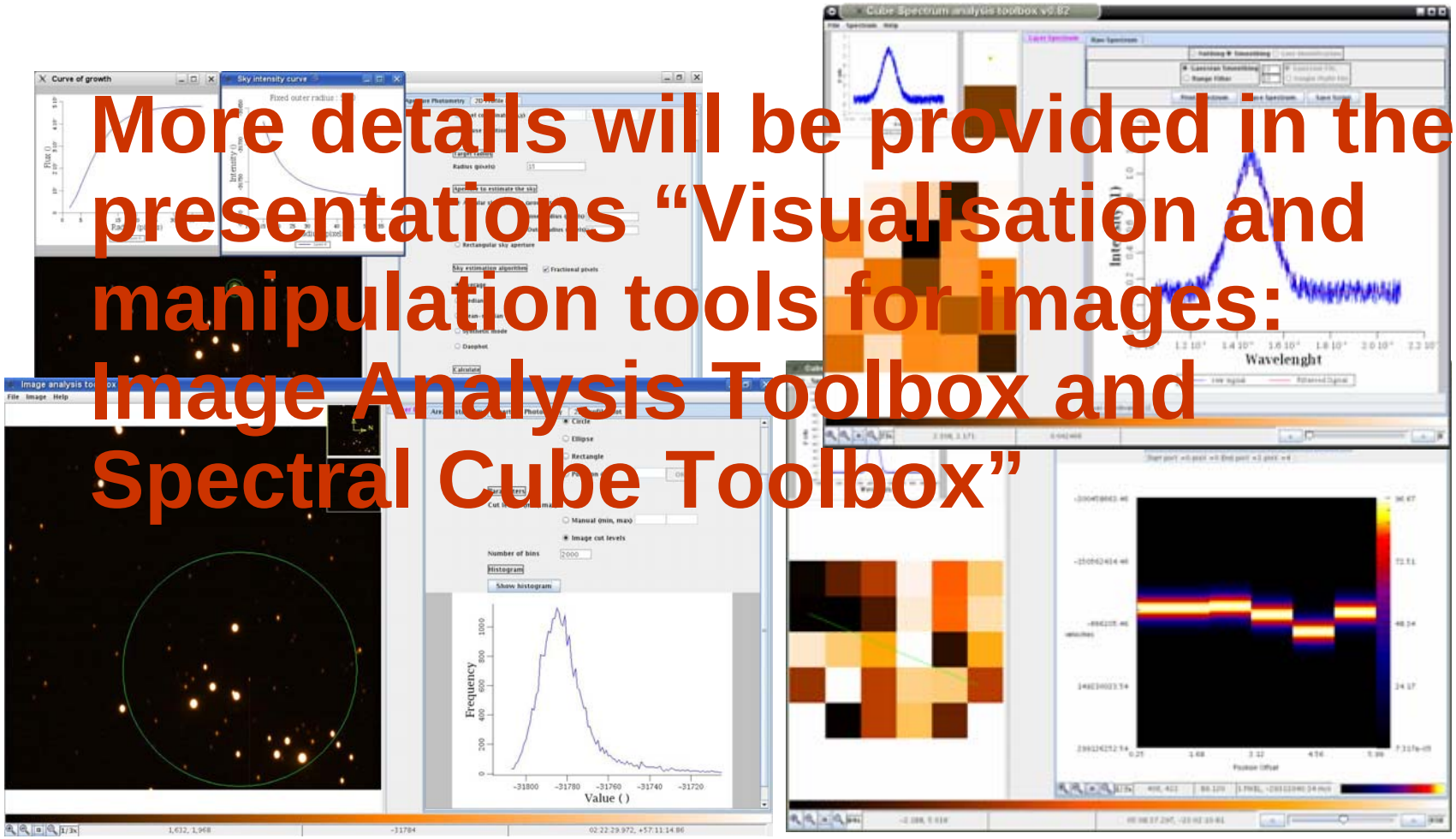


More details will be provided in the presentation "importing non-Herschel data into HIPE and exporting Herschel data to non-Herschel applications"

Virtual Observatory tools can be launched from HIPE

Software aspects – user friendly tools (1)

Tools to inspect images and data cubes



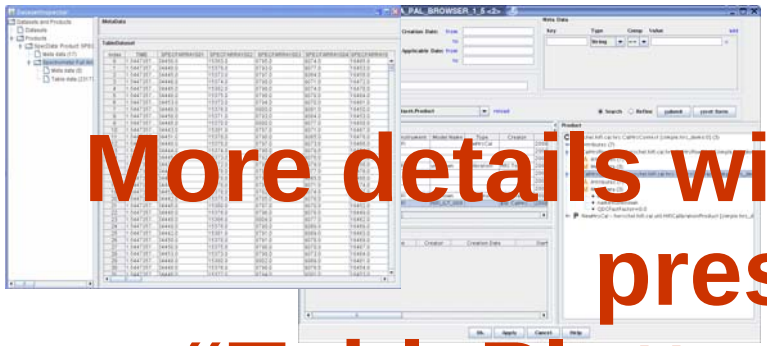
More details will be provided in the presentations "Visualisation and manipulation tools for images: Image Analysis Toolbox and Spectral Cube Toolbox"

HERSCHEL SPACE OBSERVATORY

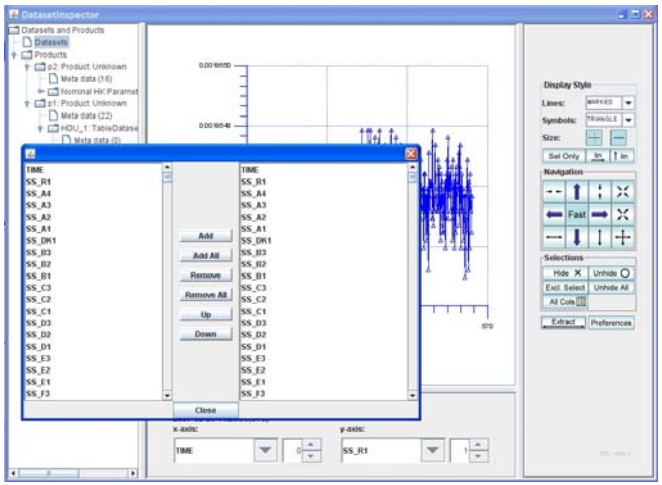
Software aspects – user friendly tools (2)

Table Plotter: Tool to display and manipulate tables

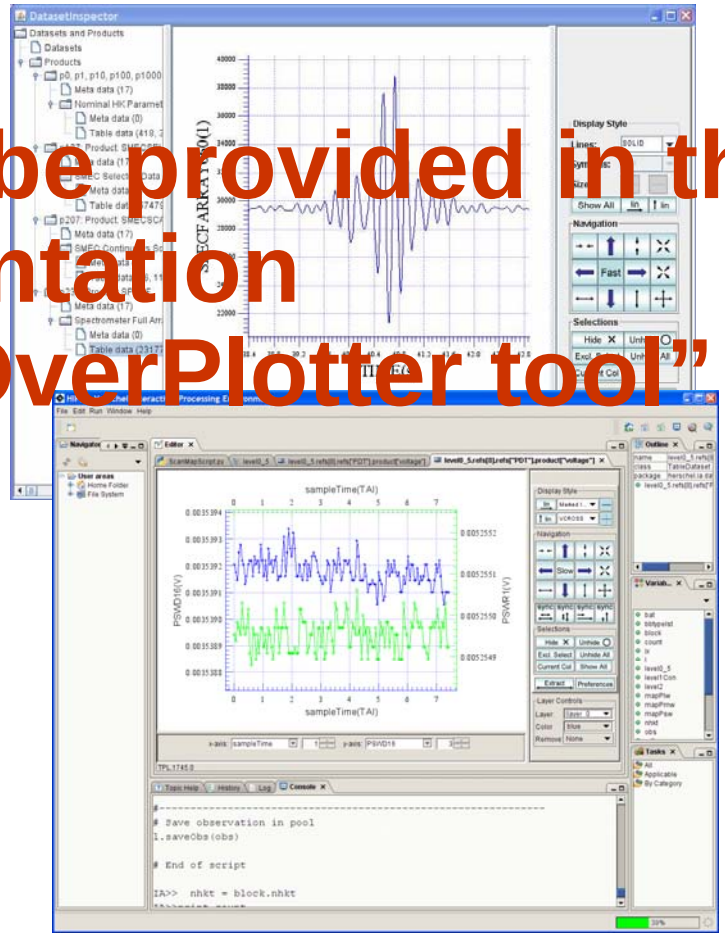
More details will be provided in the presentation "TablePlotter/OverPlotter tool"



Select observation via product browser



Manipulate data

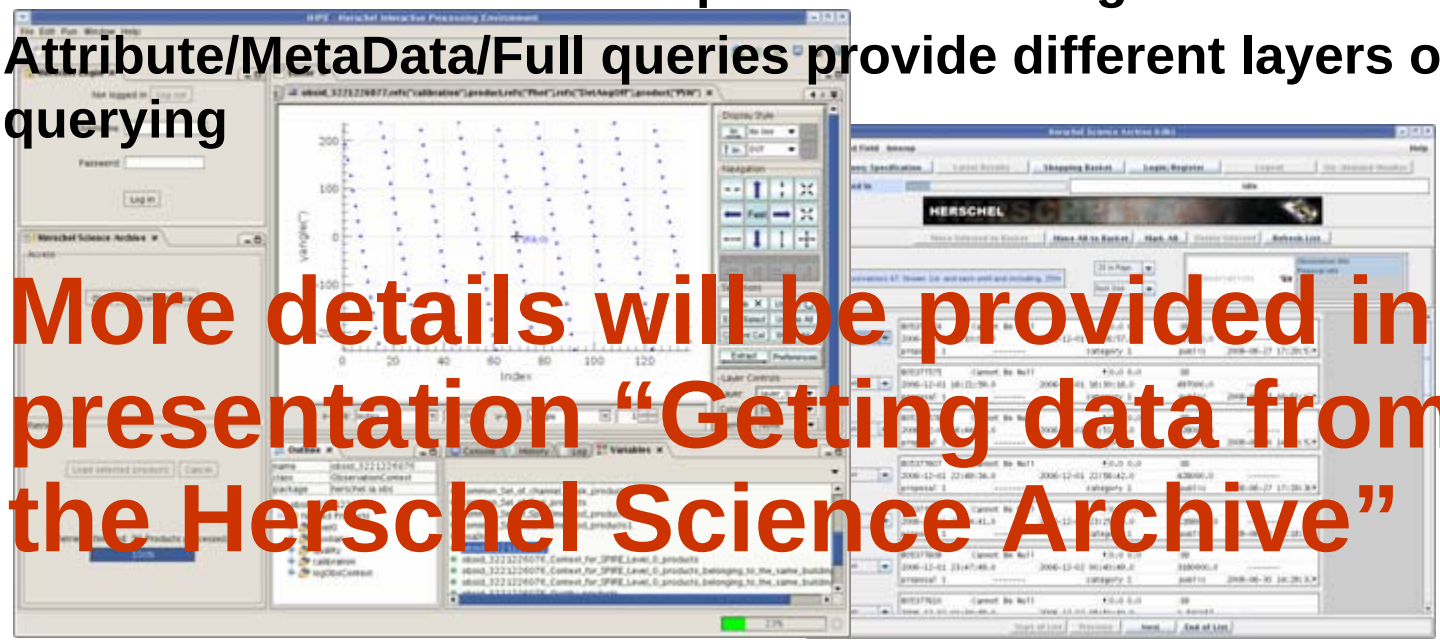


Display data

HERSCHEL SPACE OBSERVATORY

Software aspects – Data Access

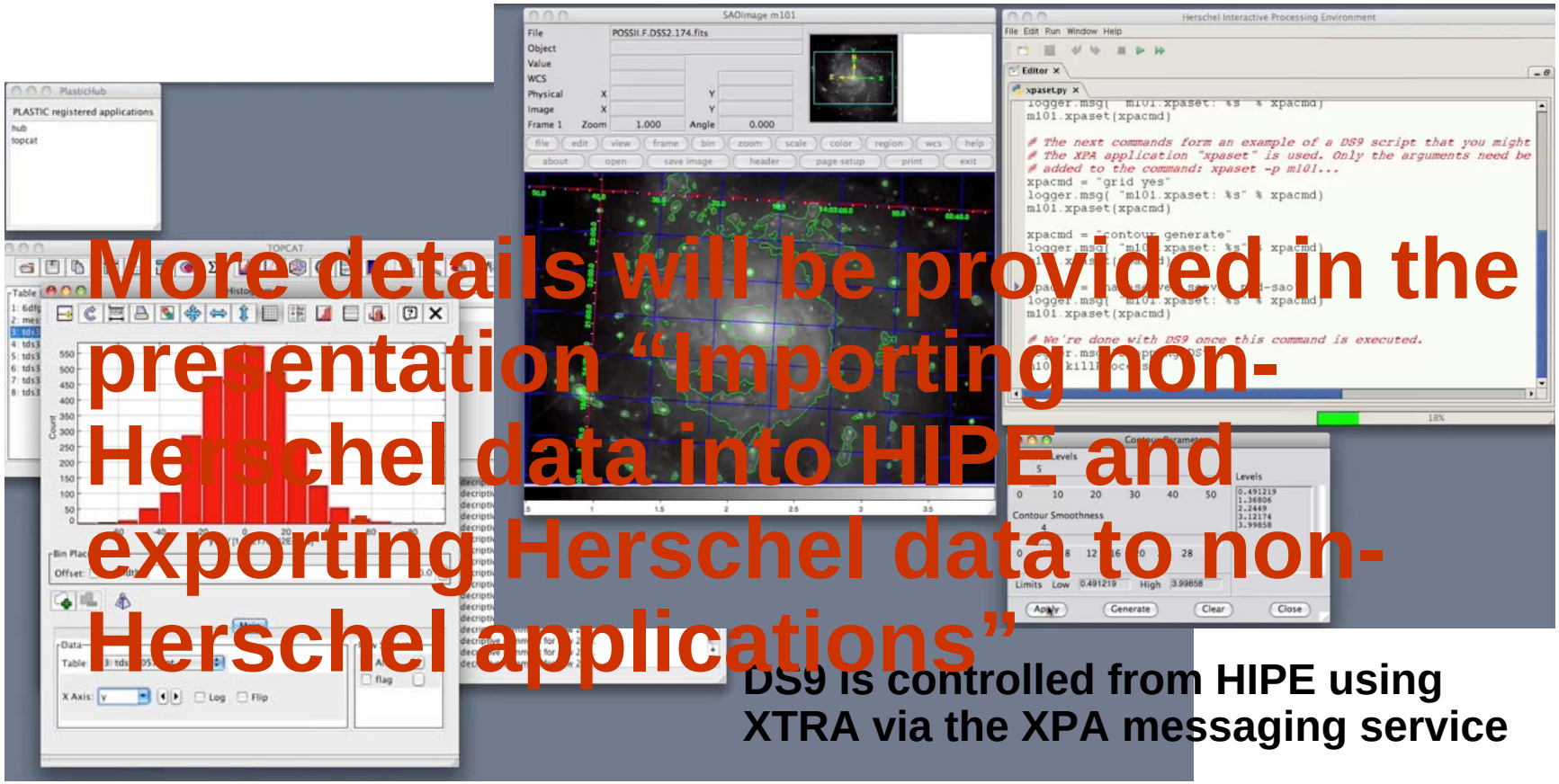
- HIPE is connected to the Herschel Science Archive, and data can be retrieved directly into an Interactive Analysis Session
- The Product Access Layer (PAL) provides an implementation-independent interface for astronomers, calibration scientists and developers alike to read, query, write and save Products to persistent storage locations
- Attribute/MetaData/Full queries provide different layers of querying



More details will be provided in the presentation "Getting data from the Herschel Science Archive"

HERSCHEL SPACE OBSERVATORY

Software aspects – Interoperability
Access to other programs via XTRA



DS9 is controlled from HIPE using XTRA via the XPA messaging service

VO TopCat displaying data from Products within HIPE via PLASTIC hub

HERSCHEL SPACE OBSERVATORY

Plans (and some
words of caution)

Words of caution first …

- While very advanced in view of the project life cycle, the Herschel Data Processing system is in beta state, and will remain so until we have seen first light
- Following in-flight experience the Data Processing system might undergo major revision. Be prepared to change your programs and scripts during the first year(s) of the mission
- Only limited manpower is available. Usually we have an oversubscription factor of three for wishes to be implemented
- We are depending on third party libraries like Java, Jython or Macrovision. Therefore we are not in complete control of the timescale for implementation of new features, or the resolution of problems

HERSCHEL SPACE OBSERVATORY

Currently ongoing major activities

- Preparation for SOVT-2
- Finalisation of system clean-up
- Finalisation of sliced dataset
- Implementation of history
- Vista support
- Implementation of user friendly property system
- Finalisation of definition of high-level products quality control products, quality control flags and reports
- Refinement of instrument pipelines and calibration
- Development of advanced pipeline algorithms and interactive tools
- Increasing the user friendliness
- Updates to documentation

We have still quite some work in front of us - prelaunch, and during operations

HERSCHEL SPACE OBSERVATORY

Plans

- Mid March 2009: A beta version of HIPE will be made available to the KP consortia
- March/April 2009: Series of topical hands-on workshops in ESAC and in NHSC
- April 2009: Launch of Herschel
- Launch + ~4 months: Planning scope and schedule of extended data processing efforts for following two years
- Launch + ~6+ months: HIPE V1.0 will be made available. Future versions will be provided in regular intervals
- Launch + ~7+ months: Workshops in ESAC and NHSC to brief community about updated pipelines and any calibration/data reduction issues after in-flight experience

Prediction is very difficult, especially about the future

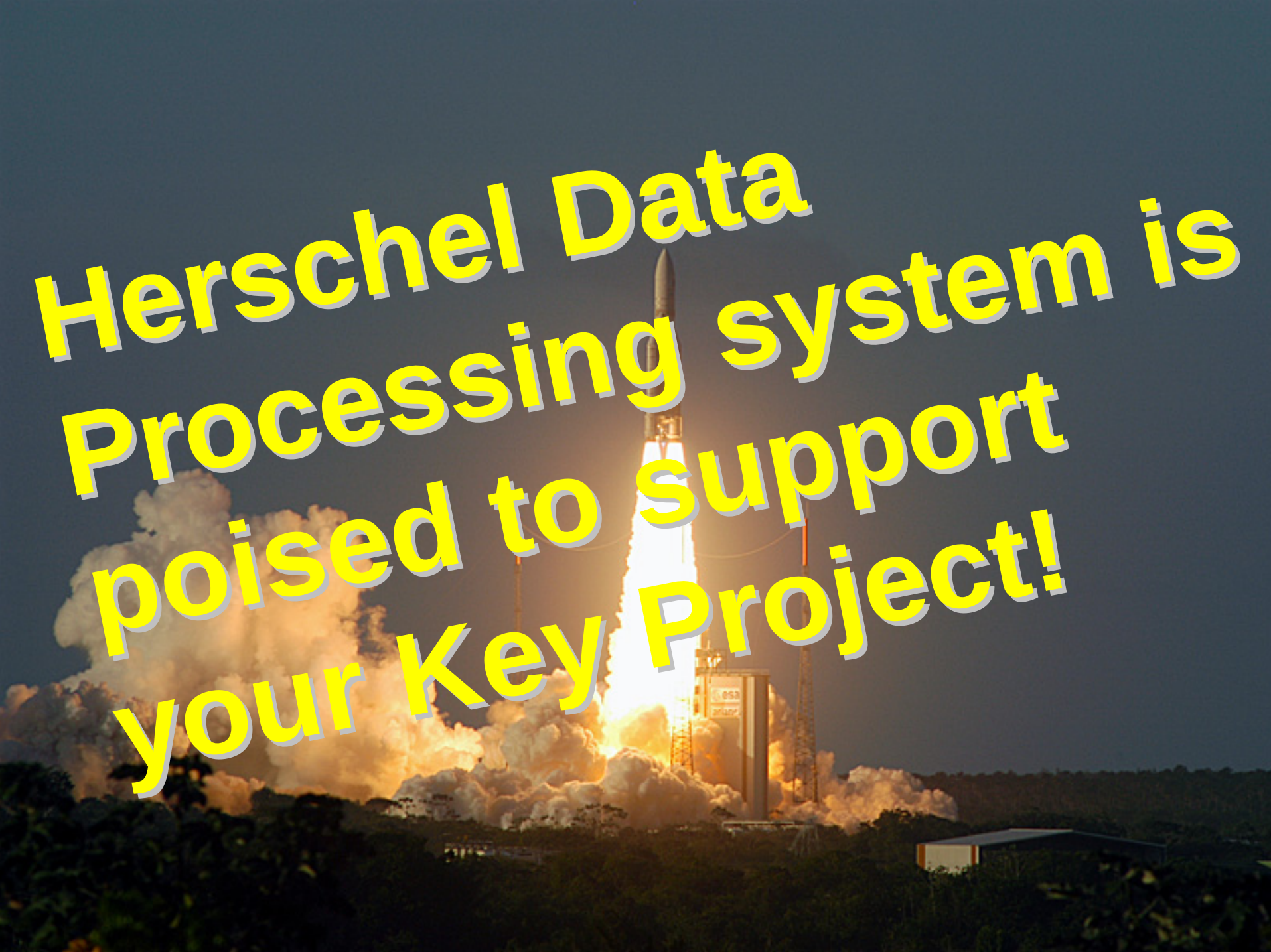
HERSCHEL SPACE OBSERVATORY

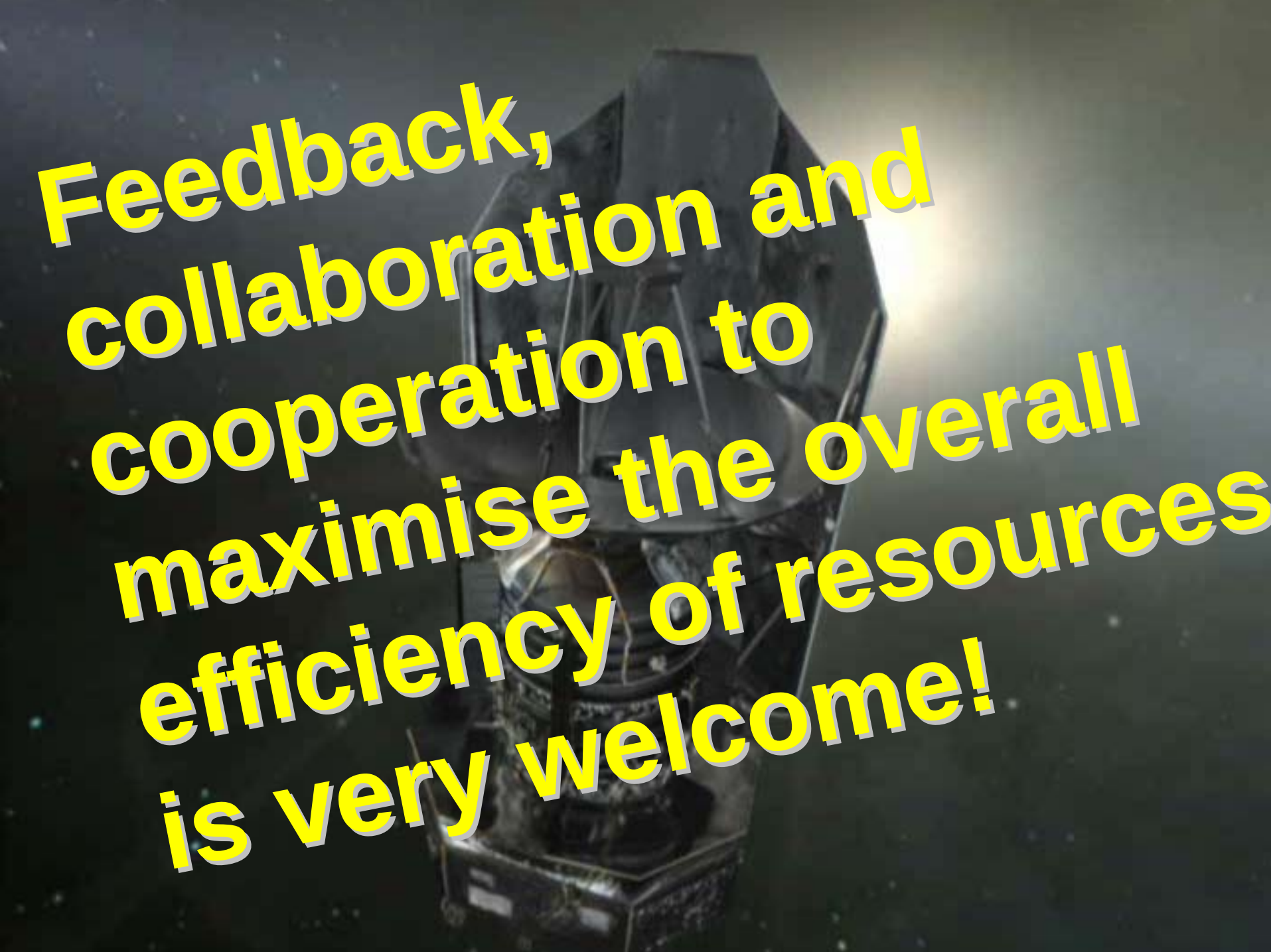
Summary

Summary

- **Herschel Data Processing system supported data analysis by the Instrument Teams for Instrument Level Tests**
- **Herschel Data Processing system supported the first System Operational Verification Test and is ready to support the second**
- **In March 2009 more topical hands-on workshops on data reduction will follow, together with a first beta release of software and data to the Key Programme community**
- **For a mission ~ 4 months before launch, the Herschel Data Processing system is in a well advanced state, containing an impressive suite of functionality and documentation**

HERSCHEL SPACE OBSERVATORY


Herschel Data
Processing system is
poised to support
your Key Project!


Feedback, collaboration and cooperation to maximise the overall efficiency of resources is very welcome!