SOFIA Data Processing and Archive

Tom Stephens

USRA Information Systems Development Manager

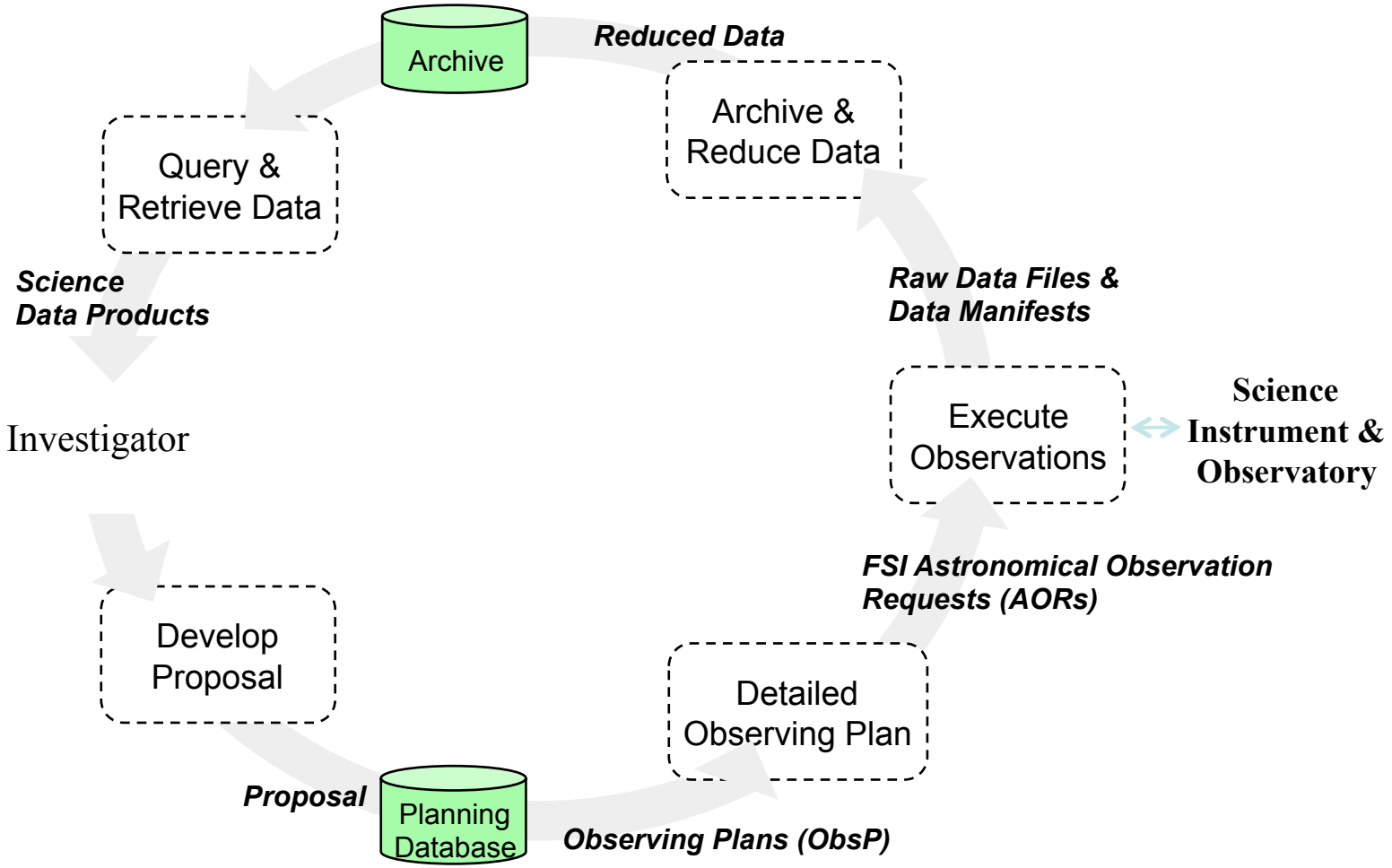
AAS SOFIA Splinter Session

Jan 4, 2010

Outline

- Data Processing Pipelines
- Data in the Archive
- Archive Access
- Help Desk support

The SOFIA Data Cycle



Data Processing Pipelines and Archive

- Purpose
  - To provide reduced and calibrated data usable by the general astronomical community without the need for specialized knowledge about the instruments

Pipeline Operational Plan

- All data will be processed by the current pipeline upon ingestion into the archive
- Archive will serve the processed data directly
- After an upgrade/improvement to the processing pipeline, older data will be reprocessed “on-demand” with the new pipeline as it requested by the users.
  - The reprocessed data becomes the new “default” data in the archive so if requested again reprocessing isn’t rerun
  - Older versions are still stored in the archive
  - *In light of experiences with other missions, the option of doing bulk reprocessing is being evaluated.*

Timescale Requirements

- Physical Timescales
  - Raw data – available in the archive 24 hours after landing
  - Fully processed data – available 14 days after flight
  - Archive queries – requested data to be available for download no more than 6 hours after request submitted
- Exclusive Data Access Periods
  - Early Science – Data will be publicly available 9 months after observations
  - Normal Operations – Data will be publicly available 12 months after observations

Pipeline Output/Archive Content

- All data in the archive will be stored in its native format
  - This includes calibration and housekeeping data
  - For all science and calibration data the native format is currently FITS
- The exact data products produced by the pipeline depend on the instruments

Disclaimer

- The following slides represent proposed contents for the archive and pipeline processing
- However, not everything can be provided given current resources.
- We'd like your input on what data products are needed, what's missing and what should have higher priority

Imagers

- Raw images (science, calibration, flats, darks, etc)
- Basic reduced images (non-linearity corrections, flat fielded and dark corrected)
- Calibrated data
- Co-added images (multiple chop/nod beams aligned and co-added)
- Maps/mosaics (multiple adjacent images combined into a larger image)

Spectrographs

- Raw images (science, calibration, flats, darks, etc)
- Basic image reduction identical to the imagers (for array spectrographs)
- Wavelength and flux calibrated data
- Extracted 1D spectra
- Combined spectra

Heterodyne Instruments

- Raw data
- Basic reduced data
  - Includes basic calibration due to the nature of the instruments

Keywords and Data in the FITS Headers

- There will be a standard set of SOFIA defined FITS keywords in all data products
- Additional keywords and data will be included as appropriate for each specific instrument and data type
- Will provide all necessary keywords so that data files work with standard community data analysis packages

Ancillary Engineering Data in the Archive

- Ancillary data will be available in the archive as well
  - Water vapor column density
  - Pointing information
  - Aircraft data (location, altitude, etc.)
- Housekeeping data
  - Some of this will be present in the science data FITS headers
  - All will be available in its raw format at the full data rate
  - Selected data will be indexed and easily searchable
- What else should we have?

Calibration Data in the Archive

- All calibration data used to process the images will also be available in the archive
  - Any specific calibration targets requested by the Guest Investigators
  - Calibration targets selected by Observatory staff
  - Science and calibration data will be indexed so it is known what calibration data was applied to what science data
- Calibration data will go through the same (appropriate) steps of the processing pipelines as the science data
- Will be available through the same user interface.

Archive Interface

- Access to the archive is web based
- Modeled on MAST and IRSA
  - Similar interface and parameters
  - Should be familiar to the community

Science Analysis Tools Provided by the SMO

- Currently no SOFIA specific science analysis tools are planned
  - Data formats designed to be usable in standard community tools
- Are there tools that we need to supply?

- We will have a process to make “contributed” software from the community available
  - The details for this are just starting to be discussed

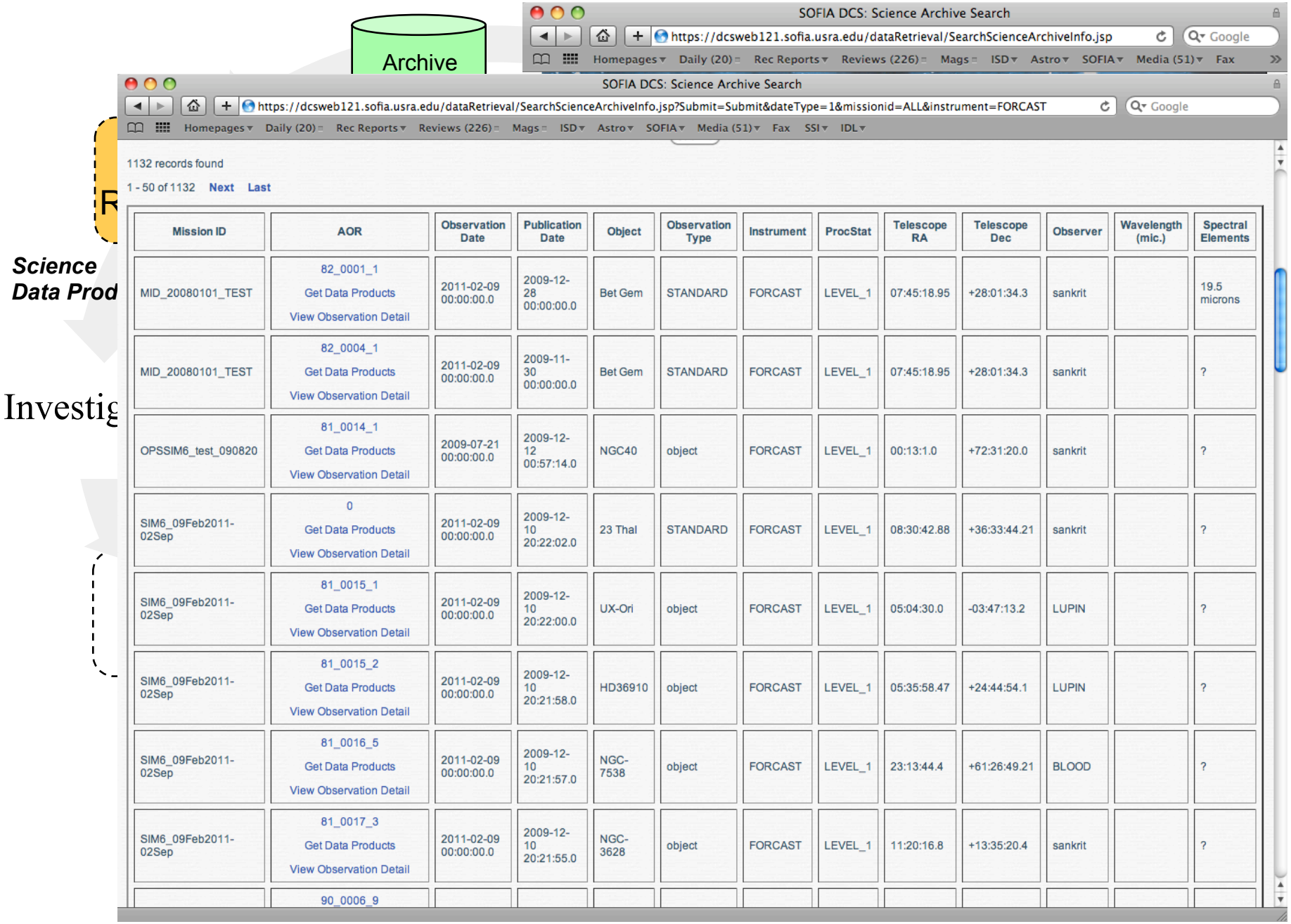
Help Desk

- SOFIA's help desk will provide assistance with
  - General questions about SOFIA and its instruments
  - Support for the proposal process
  - Support for Guest Observers relevant to observation preparation
  - Support for user of the data archive
- The help desk is also the place to provide suggestions and feedback on the SOFIA program.
  - Comments and suggestions from this workshop are welcomed and encouraged
- Help desk URL:
  - http://www.sofia.usra.edu/Contact/contact.html

Request for Input

- Science and calibration data products
  - Are there any data products missing that should be provided?
  - Are there any listed that are not needed?
  - What priority should the various products have relative to one another?
- What Ancillary/housekeeping data products are needed?
- Are there any SOFIA specific tools that the community believes should be provided?

Backup Slides



Mission ID	AOR	Observation Date	Publication Date	Object	Observation Type	Instrument	ProcStat	Telescope RA	Telescope Dec	Observer	Wavelength (mic.)	Spectral Elements
MID_20080101_TEST	82_0001_1 Get Data Products View Observation Detail	2011-02-09 00:00:00.0	2009-12-28 00:00:00.0	Bet Gem	STANDARD	FORCAST	LEVEL_1	07:45:18.95	+28:01:34.3	sankrit		19.5 microns
MID_20080101_TEST	82_0004_1 Get Data Products View Observation Detail	2011-02-09 00:00:00.0	2009-11-30 00:00:00.0	Bet Gem	STANDARD	FORCAST	LEVEL_1	07:45:18.95	+28:01:34.3	sankrit		?
OPSSIM6_test_090820	81_0014_1 Get Data Products View Observation Detail	2009-07-21 00:00:00.0	2009-12-12 00:57:14.0	NGC40	object	FORCAST	LEVEL_1	00:13:1.0	+72:31:20.0	sankrit		?
SIM6_09Feb2011-02Sep	0 Get Data Products View Observation Detail	2011-02-09 00:00:00.0	2009-12-10 20:22:02.0	23 Thal	STANDARD	FORCAST	LEVEL_1	08:30:42.88	+36:33:44.21	sankrit		?
SIM6_09Feb2011-02Sep	81_0015_1 Get Data Products View Observation Detail	2011-02-09 00:00:00.0	2009-12-10 20:22:00.0	UX-Ori	object	FORCAST	LEVEL_1	05:04:30.0	-03:47:13.2	LUPIN		?
SIM6_09Feb2011-02Sep	81_0015_2 Get Data Products View Observation Detail	2011-02-09 00:00:00.0	2009-12-10 20:21:58.0	HD36910	object	FORCAST	LEVEL_1	05:35:58.47	+24:44:54.1	LUPIN		?
SIM6_09Feb2011-02Sep	81_0016_5 Get Data Products View Observation Detail	2011-02-09 00:00:00.0	2009-12-10 20:21:57.0	NGC-7538	object	FORCAST	LEVEL_1	23:13:44.4	+61:26:49.21	BLOOD		?
SIM6_09Feb2011-02Sep	81_0017_3 Get Data Products View Observation Detail	2011-02-09 00:00:00.0	2009-12-10 20:21:55.0	NGC-3628	object	FORCAST	LEVEL_1	11:20:16.8	+13:35:20.4	sankrit		?
	90_0006_9											