

# SPHEREx Data Pipeline: Data Products and Tools

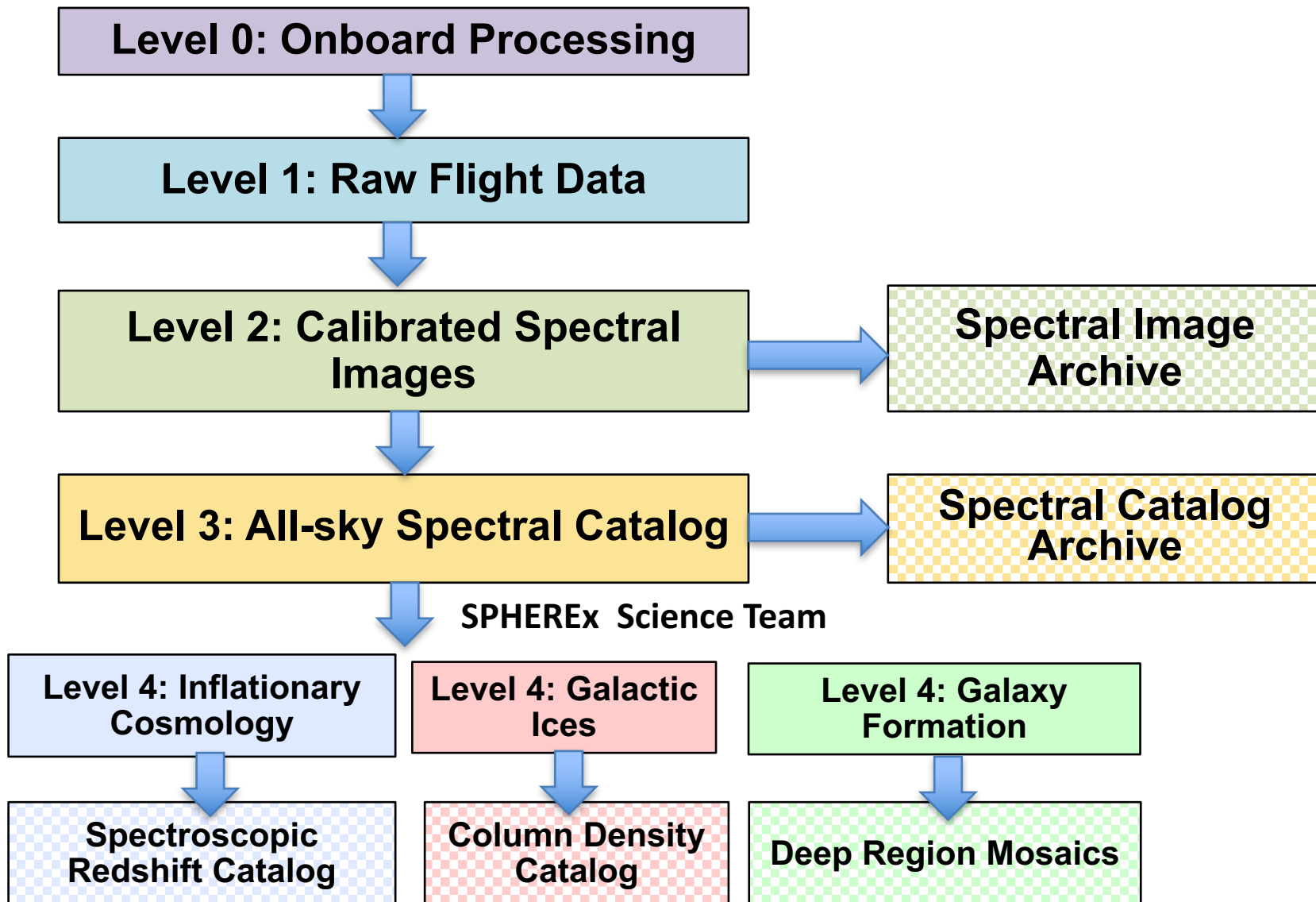
Brendan Crill

JPL

# Outline

- Data processing pipeline for the primary science themes of SPHEREx
- Data products and software tools for public distribution
- What additional products and tools could be funded with a Science Enhancement Option
- Please give us feedback throughout the workshop!
  - Friday 10:30am discussion

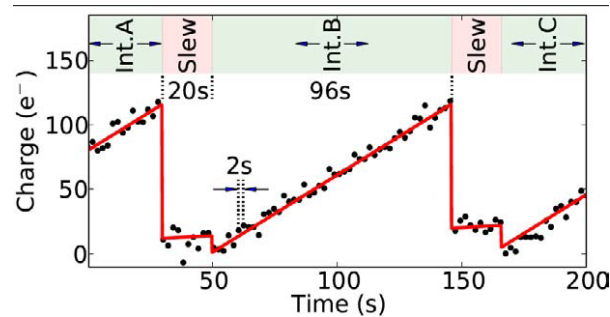
# Data Flow



# Organization

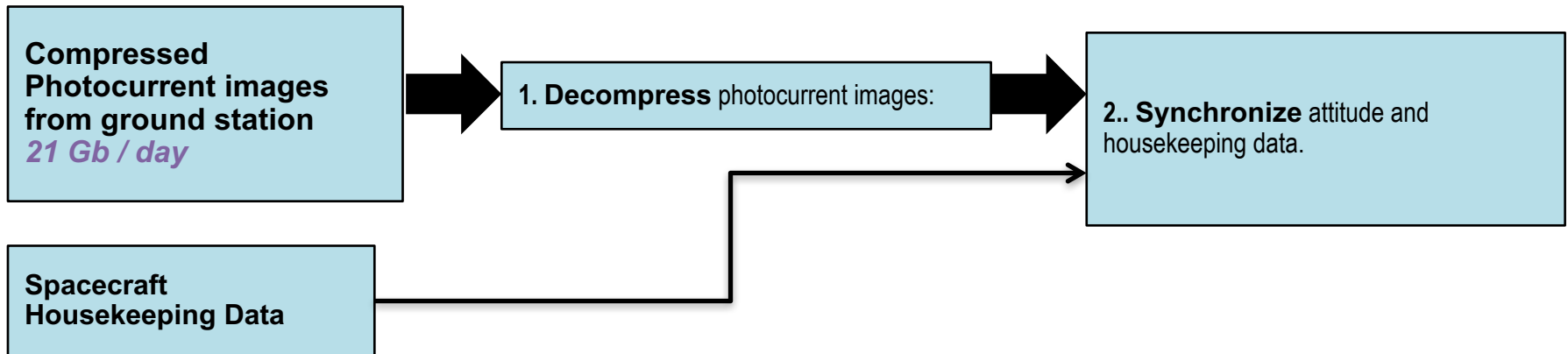
- JPL ground station
  - Retrieves spacecraft data packets from near earth network
- SPHEREx Science Center (Caltech)
  - processes daily flow of image data from spacecraft: cleaning, calibration, astrometry, PSF measurement (Level 1, Level 2)
  - Creates spectral catalogs from image data (Level 3)
  - Provides data quality checks
  - Delivers data products internally to SPHEREx science team and to the public archive
- SPHEREx Science team
  - Algorithms, software, science interpretation for the Inflationary Cosmology, Galactic Ice, and Extragalactic Background projects (Level 4)
  - See talks from Doré, Cooray, Melnick, Capak, Ashby, etc. for details..
- IRSA (IPAC)
  - Serves public data archive

# Level 0: On-board data processing



- H2RG detectors sampled at 0.5 Hz
- On-board processing:
  - linear fit to the samples-up-the-ramp for  $\sim 100$ s exposures
  - Flag cosmic ray hits
  - Flag saturation
  - Data compression
- Photocurrent images and flags telemetered to ground:  
 $\sim 21$  Gb/day

# Level 1: Raw image data



- Data pushed from ground station to SPHEREx science center
- Decompress images
- Synchronize with housekeeping and spacecraft attitude data

# Level 2: Calibrated Spectral Images

**Uncompressed Photocurrent Images**  
*36 Gb/day*

**1. Persistence Correction and Masking**

**Ground Calibration Products**

- Calibration matrix.
- Ancillary point source catalog.
- Spectral standard stars.

**2. Calibration**

- Absolute and channel to channel calibration.
- Flight Flat Field estimate.
- Bright pixel history.
- Dark current estimate.

**3. Astrometric Solution**  
to  $<1''$  per  $3.5^\circ \times 7^\circ$  field.

**5. Data Quality Assessment + Observation Planning**

**4. Measure Point Spread Function**  
(PSF) per field

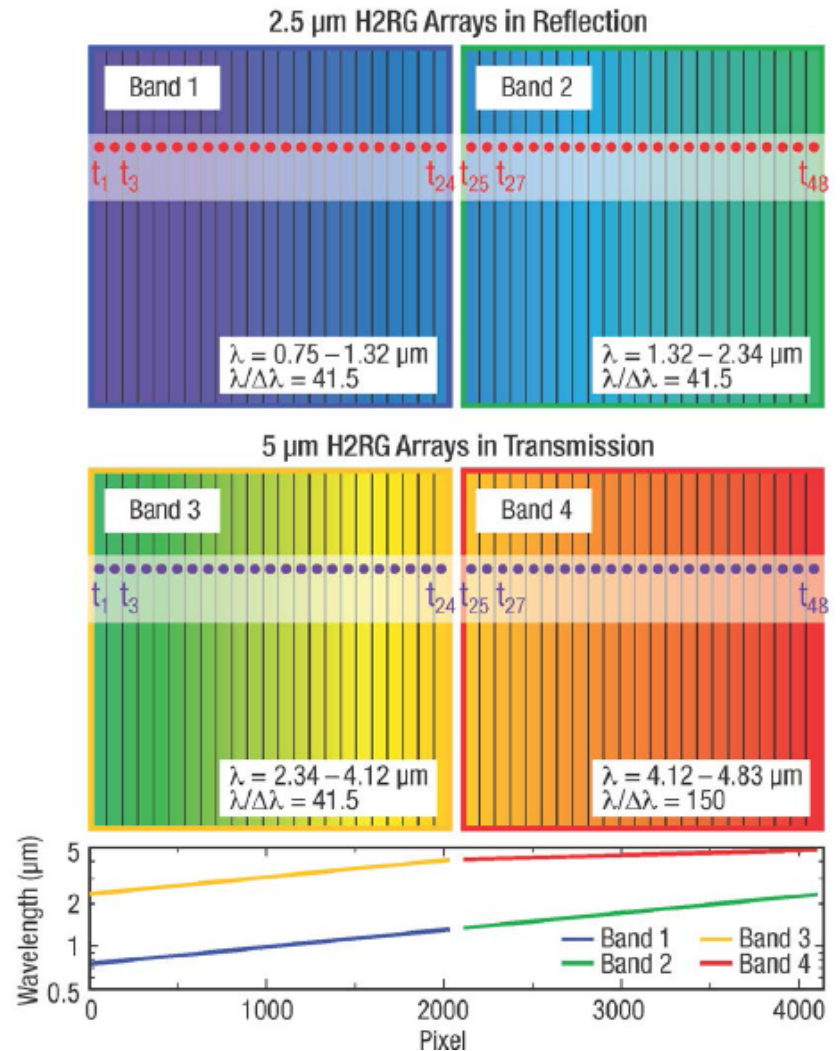
**Spectral Image Data Archive**

*3.2 Tb/3 month*

- Spectral Image database
- Flag images
- Wavelength images
- PSF data
- Documentation
- Code Examples
- Interface with IRSA

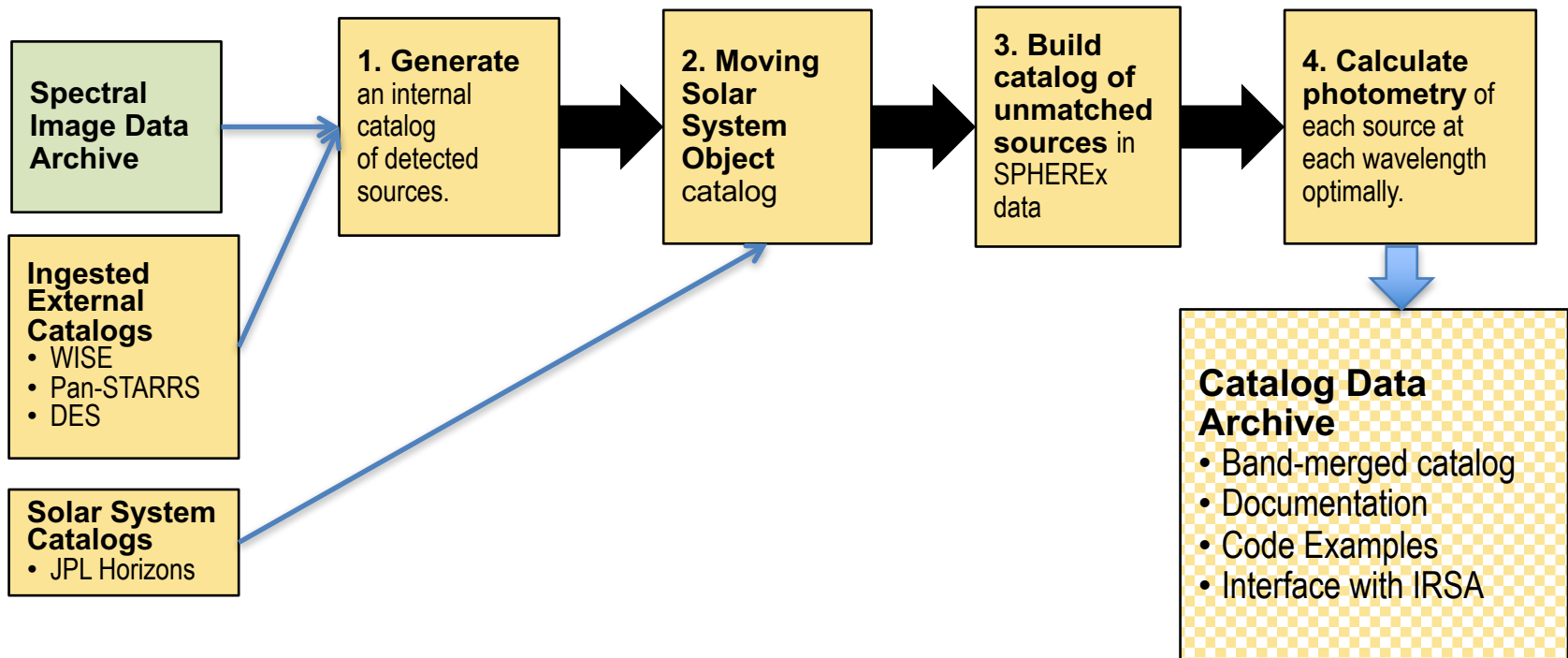
# Spectral Image Archive

- Four .fits image files per spacecraft pointing
- Cleaned, calibrated, astrometrically registered
- Bandpass varies across image
- Ancillary products:
  - PSF database
  - Image database to allow fast lookup of sky locations
  - Instrument info: spectral response, etc.

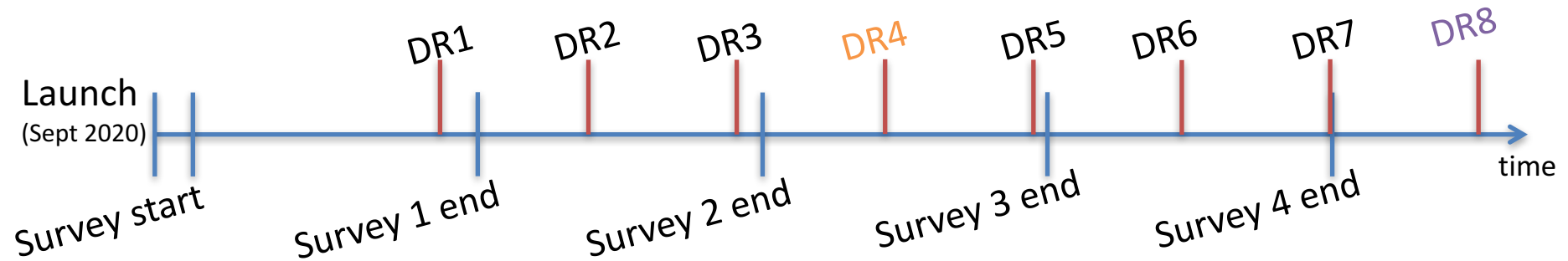




# Level 3: Spectral Source Catalog



# Baseline Public Data Release Plan



- Every 3 months starting 6 months after science survey begins: **Spectral Image Archive** delivered to IRSA
  - Includes SQL database for quick lookup of sky location
- DR4: 1-year **Spectral Catalog** delivery
  - Full spectral sampling
- DR8: End of mission: **Level 4 science products**
  - Deep field mosaic data cube
  - Column Density Catalog
  - Spectroscopic Redshift Catalog
- Includes documentation + software examples

# Science Enhancement Option: Legacy Archive Tools

- SPHEREx proposal includes a Science Enhancement Option:
  - Additional funding for a more user-friendly interface to the data archive, including additional tools and data products
- For the most part: SPHEREx science team internal tools and products will be better documented, made robust, and served via IRSA

# Science Enhancement Option Wish List

## Products/Tools

- Full-sky data cube viewer and extractor:
  - Full-sky mosaics
  - Full-sky Line intensity maps
  - On-the-fly mosaics (custom data cuts in time, etc)
- PSF-fit tool
  - For a given sky position and epoch, return PSF-fit photometry
- Variable source extractor
  - Flux density vs. time for variable sources
- Moving source flux density estimator
  - Determine flux density vs time for moving objects
- Unmatched source catalog
  - A catalog of sources detected by SPHEREx, with no counterpart in our external catalogs
- ...?
- Discussion on Friday at 10:30 am