

# EPDS ARCHIVE AND DATA MANAGEMENT

---

Rachel Akeson  
Deputy Director  
NASA Exoplanet Science Institute  
(NExSci)



# NExSci Role in EPDS

- NExSci is the NASA Exoplanet Science Institute and is the science operations and analysis center for NASA's Exoplanet Exploration Program
- As stated in the call, NExSci is expected to have responsibility for data processing and archiving after commissioning
- We have developed and are operating the following archives:
  - KOA – Keck Observatory Archive
  - LBTI Archive (no public data yet)
  - NASA Exoplanet Archive

# EPDS archive

- All science data will be in the archive and will go public after a 12 month proprietary period
  - Includes GTO time
  - Engineering and commissioning data will be evaluated for format compliance and usability before being placed in archive
- Selected team will produce raw (Level 0) data
- Selected team will provide a pipeline to generate internally calibrated (Level 1) data
  - Includes bad pixel removal, dark and background subtraction, wavelength calibration, and flat fielding
- Archive will also include ancillary data (e.g. weather)
  - Coordinated with NOAO

# Proposal requirements on data issues

- Step 1 and 2 proposals do not need to describe interface in detail (per call)
- NExScI will coordinate with selected teams during Instrument Concept Study to ensure submitted data pipeline plans meet requirements and work is properly scoped
- NExScI and NOAO will work with team selected for full implementation to develop data management plan (DMP) covering
  - Data transfer mechanisms and responsibilities
  - Level 0 format (FITS) and contents
  - Level 1 format (FITS) and contents
  - Pipeline requirements

# Why discuss data content now?

- Although Step 1/2 proposals do not require pipeline and data details, our experience is that the final product is better when the data issues are considered early
  - If keywords have to be added during commissioning, early data is less useful or even impossible to properly understand and archive

# Level 0 data header contents

- Header contents must include all information necessary to understand what data was taken and **to drive pipeline processing**
- Note that some of this information will have to be gathered from the telescope
- Telescope description
  - Location, date, telescope, image rotators
- Instrument state (anything set by observer should be recorded)
  - Filters, focus
- Observation parameters
  - Integration time, source name, sky coordinates, airmass

# Level 0 (continued)

- Calibration type
  - dark, flat
- Calibration source information
  - Lamp type, settings
- Detector information
  - Gain, bias, readout mode
- KOA Examples:
  - HIRES  
[http://www2.keck.hawaii.edu/koa/public/keywords/koa\\_keywords.php](http://www2.keck.hawaii.edu/koa/public/keywords/koa_keywords.php)
  - NIRSPEC <http://www2.keck.hawaii.edu/koa/public/keywords/NirspeckKeywords.php>

# Archive-added keywords

- The archive will also add keywords to allow for additional data searches
  - Program ID and PI information
  - Basic weather information (if available)
  - Added for both Level 0 and 1

# Basic pipeline requirements

- Must produce instrument calibrated data using only files from the Level 0 archive
  - No configuration/instrument information should be stored elsewhere
  - Relevant calibration data identified through keywords and heuristics
- Should require minimal human intervention to process a full night of data
- Must have a scriptable mode, i.e. not GUI only
- Must provide some data quality metrics
  - Discuss as part of DMP definition

# Level 1 example keywords

- Same telescope, instrument, source keywords as Level 0
- Processing keywords
  - Pipeline settings
  - Calibration files used
  - Derived wavelength calibration
  - Warnings or errors found during processing
- Data quality metrics (if available)

# Archive interface

- User-specified, web-based searches on metadata, including
  - Source name and position
  - Program name
  - Observation date
  - Wavelength range
  - Instrument mode, exposure time, calibration type
- Level 0 data results
  - Science and calibration file lists
    - Relevant calibration files identified with same heuristics as pipeline
  - Files can be downloaded individually or as complete set meeting search criteria
  - Links to CCD data visualizations
- See [koa.ipac.caltech.edu](http://koa.ipac.caltech.edu) to try out searches on your own

# KOA advanced search interface

Basic Search More Search Options Released Programs

## 1. Choose Instrument: ?

- |   |  |   |
|---|--|---|
| <input checked="" type="checkbox"/> HIRES           | <input checked="" type="checkbox"/> NIRC2          | <input checked="" type="checkbox"/> NIRSPEC         |
| <input checked="" type="checkbox"/> iodine cell in  | <input checked="" type="checkbox"/> narrow-imaging | <input checked="" type="checkbox"/> low-dispersion  |
| <input checked="" type="checkbox"/> iodine cell out | <input checked="" type="checkbox"/> medium-imaging | <input checked="" type="checkbox"/> high-dispersion |
|   | <input checked="" type="checkbox"/> wide-imaging   | <input checked="" type="checkbox"/> imaging         |
|   | <input checked="" type="checkbox"/> spectroscopy   |   |
| <input checked="" type="checkbox"/> DEIMOS          | <input checked="" type="checkbox"/> ESI            | <input checked="" type="checkbox"/> LRIS            |
| <input checked="" type="checkbox"/> MOSFIRE         | <input checked="" type="checkbox"/> OSIRIS         |   |
|   | <input checked="" type="checkbox"/> imaging        |   |
|   | <input checked="" type="checkbox"/> spectroscopy   |   |

Check All Clear All

(To retrieve public Keck Interferometer data, use the dedicated KI Search Form)

## 2. Choose File Types to Return: ?

- Return **any files** (science and/or calibration)
- Return **only science files**:
- Also show calibration files that are appropriate for data reduction (these files may not match all the search criteria)
- Return **only calibration files** that are of type:
- |                                  |                                      |                                |                                    |
|----------------------------------|--------------------------------------|--------------------------------|------------------------------------|
| <input type="checkbox"/> arclamp | <input type="checkbox"/> flatlamp    | <input type="checkbox"/> focus | <input type="checkbox"/> dark      |
| <input type="checkbox"/> bias    | <input type="checkbox"/> flatlampoff | <input type="checkbox"/> trace | <input type="checkbox"/> undefined |
| <input type="checkbox"/> polcal  | <input type="checkbox"/> fscal       |                                |                                    |

## 3. Enter Search Criteria: ?

- Spatial Search & Name Resolver:**  NED first  SIMBAD first

Radius:   (0.1 arcsec - 6 deg)

**Single Object Center:**

Coordinates or object name, to be resolved by NED or SIMBAD. Examples: "DQ Tau",

<b>UT Observation Date(s):</b>	<input type="text"/>
mm/dd/yyyy or yyyyymmdd	
<b>UT Observation Time(s):</b>	<input type="text"/>
See allowed formats	
<b>Exposure Time:</b>	<input type="text"/>
# of seconds, i.e. <500, >=900 or 300-600	
<b>Program ID:</b>	<input type="text"/>
Ex: Hu33H	
<b>Semester:</b>	<input type="text"/>
Ex: 2004A	
<b>Principal Investigator:</b>	<input type="text"/>
Last Name	
<b>Program Title:</b>	<input type="text"/>
Keywords list, i.e., "deep impact", "binary"	
<b>Institution:</b>	<input type="text" value="All"/>
<b>Frame Number(s):</b>	<input type="text"/>
List of values or range, i.e: "32, 35", "< 90", ">= 150", "4-10"	
<b>KOAIID:</b>	<input type="text"/>
Ex: HI.20031212.18900 or HI.2003 or NS.20031212.18900 or NS.2003	
<b>HIRES/NIRSPEC/NIRC2/OSIRIS Only:</b>	
<b>Wavelength(s) Covered:</b>	<input type="text"/>
Ex: 5000Å or 2000-11000Å or 4µm or 2.5-6µm	

Search

Reset Form

### Retrieve ancillary weather data

UT Observation Date(s):

See allowed formats

Go

# Example Level 0 visualization

- Example from KOA: HIRES data on 51 Peg

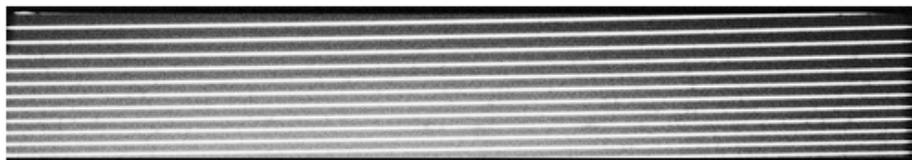
**Available Quicklook JPEG Previews  
for KOAID: HI.20040819.42748.fits**



**CCD: 1 (blue)**  
(HDU 01)



**CCD: 2 (green)**  
(HDU 02)



**CCD: 3 (red)**  
(HDU 03)

# Archive interface – Level 1 data

- Search returns file list and links to visualization of spectra
- Includes data metrics from Level 1 processing

CCD 2 (green)

Extracted Spectra Grades [\[Help\]](#)

Order #:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Grade:	Pass																	
	Plot																	

