

# UVIT Data Pipeline Software

## A status Report



Rekshesh Mohan



# UVIT::Overview

- **Instrument**
  - Multiple filters and grism
  - FOV: 30' with ~2" resolution
  - Co-aligned with X-ray telescopes
- **Observing mode**
  - Pointed, ability to scan
  - 35 minute observations
- **Raw Data Volume**
  - 50 GB/day
  - 35 TB/year



# DESIGN::Considerations

- **Large datasets: TBs of data**
  - Efficient and Quick Search
  - Data+header in one file or separate metadata?
  - Web interface
- **Mission Lifetime: Years**
  - Different arch. & platforms
  - Data formats, database software, versions
  - Generations of developers & maintainers



# SOFTWARE::Philosophy

- **Open source (GPL)**
  - Java (First of its kind)
- **Platform neutral**
  - Support limited by manpower/availability.
- **Modern**
  - Modular
  - Self-documenting
  - Easily maintained
  - Reusable tools
- **Baselined for Sun Java 1.4.x on GNU/Linux**
  - Tested on both Intel and AMD arch.



# SOFTWARE::Components

- **Simulations**
  - UVIT responsibility
- **Pipeline**
  - UVIT responsibility
- **Archival/Dissemination**
  - ISSDC responsibility
- **High level products**
  - ISSDC, with inputs from Science team
- **Misc. tools – Exposure time calculator, Obs. planner**
  - Software Team (IIA)



# SOFTWARE::Personnel

- **1 Boss**
  - J. Murthy
- **3 PDFs**
  - R. Mohan
  - M. Safonova (TAUVEX)
  - P. Gopakumar (TAUVEX)
- **3 software trainees**
  - M. Fayaz
  - V. Sharan
  - L. Geetha



# UVIT::Pipeline

- **Tasks**

- Ingest
- Data validation
- Data correction
- Data registration
- Production of Level 2 data

- **Data Definitions**

- Level 0: Out of the Spacecraft, into Ground Station
- Level 1: Out of the Ground Station, into pipeline
- Level 2: Images from individual observations
- Level 3: Data from multiple Observations
- Level 4: Derived products



# PIPELINE::Operation

- Run from a shell script
- Operation governed by parameter files
- Data file produced at each logical step
- May be run interactively
  - **Steps can be skipped or substituted**
- Data file is self-documenting
  - **History of file contained in header.**





# DATA::Ingest

- **Read the data from ground station (level 1)**
  - Reformat data in software
  - All the work on Level 1 data dealt with by a single module
- **Expect a single data file (For one observation)**
  - Mission data can come separately
- **Writes out Level 1a data (will be archived)**



# DATA::Validation

- **Validation checks for data quality**
  - Mission parameters
  - Data parameters
- **Issues? => Keep it aside for manual processing**
- **Errors and info. Messages written to log file**
  - Blinking screen?
  - Message window pop-up?
  - Play music?
  - Speak out loud?
  - SMS concerned people?



# DATA::Processing

- **Data correction**
  - Distortion correction
  - Flat fielding
  - Calibration
- **Data registration**
  - *Work in progress*
- **Parameter files govern corrections**
  - Ancillary software required to produce and update these parameter files
- **Writes out Level 2 Data**



# DATA::Products

- **Level 2 data**
  - FITS images of individual observations
  - Point source lists
- **Level 3 data**
  - Data from multiple observations
- **Level 4 data**
  - Derived data products
- **Only Level 2 data is our (UVIT) responsibility**



# PIPELINE::Status

- **Stable version (1.0) is ready for TAUVEX**
  - UVIT specific modules need to be worked on!
  - Detailed testing required
  - Data registration not yet done
- **Need to run extensive simulations**
  - No shortage of time (hopefully)
  - Open issues – a few, none critical



# Open Issues

- **Level 0 data format – Not available**
- **Mission parameters**
  - Requires coordinates of the look direction
  - Information on SSM movements
- **Do we store all raw data?**

