

Developing Data Management Project: Planning

IST400/600
Jian Qin

A case: Mars Express project data management plan

- Questions need to be addressed:
 - What are the roles and responsibilities of data management and the data system?
 - What are the data products?
 - What are the processing levels of data products?
 - How should information about data products and processing be documented?
 - How will data quality be controlled?
 - How will data be archived?

http://www.aspera-3.org/ids/APAF_PDMP_V1.3.pdf

The Mars Express project

- An European Space Agency (ESA) project
- The mission helps to answer fundamental questions about the geology, atmosphere, surface environment, history of water and potential for life on Mars.
- <http://sci.esa.int/science-e/www/area/index.cfm?fareaid=9>



IST400/600 Scientific Data Management

3

Agencies involved in the project

- European Space Agency (ESA)
- NASA, USA
- Data products to be delivered to:
 - Analyzer of Space Plasma and Energetic Atoms (3rd Version) (ASPERA-3) team
 - Co-Investigators of Mars Express project
 - ESA Planetary Science Archive
 - NASA Planetary Data System

IST400/600 Scientific Data Management

4

Data collection instruments

- An orbiter spacecraft with 7 instruments on board (ASPERA-3 is one of the 7 instruments)
 - Performs remote sensing experiments
- A deployable lander
 - Performs exobiology and geochemistry research
- A network of ground and data processing stations

IST400/600 Scientific Data Management 5

Data flows

IDFS = Instrument Data File Set or Instrument Description File Set

Telemetry Flow

```

            graph LR
            Satellite[Satellite] -- Telemetry --> ESOC[ESOC]
            ESOC --> IRF[IRF]
            IRF --> SwRI_APAF[SwRI APAF]
            subgraph NASA
            SwRI_APAF
            end
            subgraph ESA
            ESOC
            IRF
            end
            
```

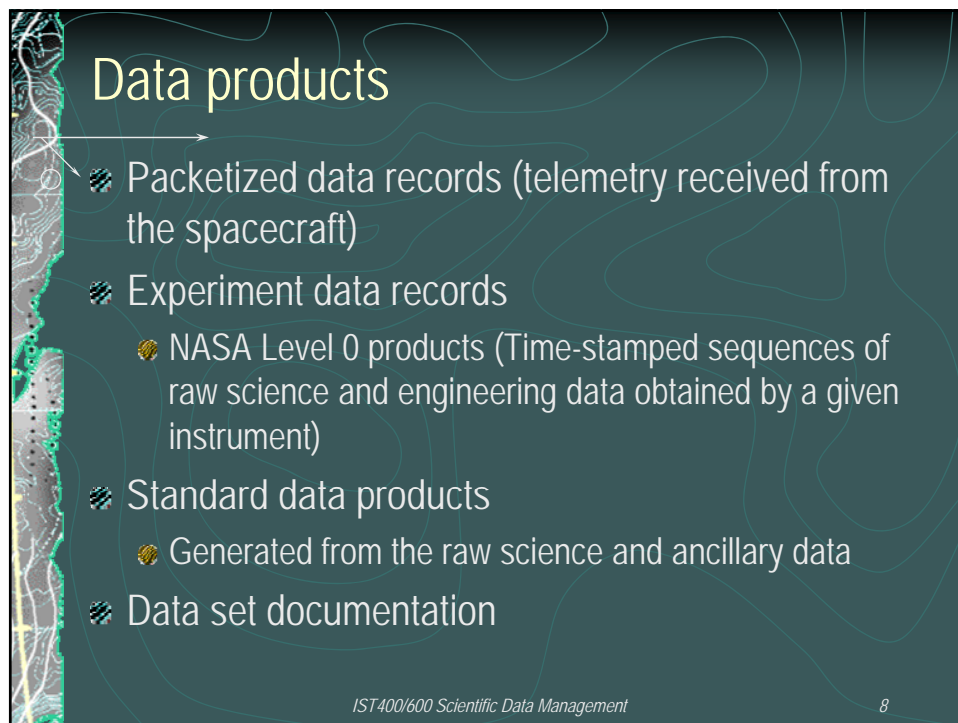
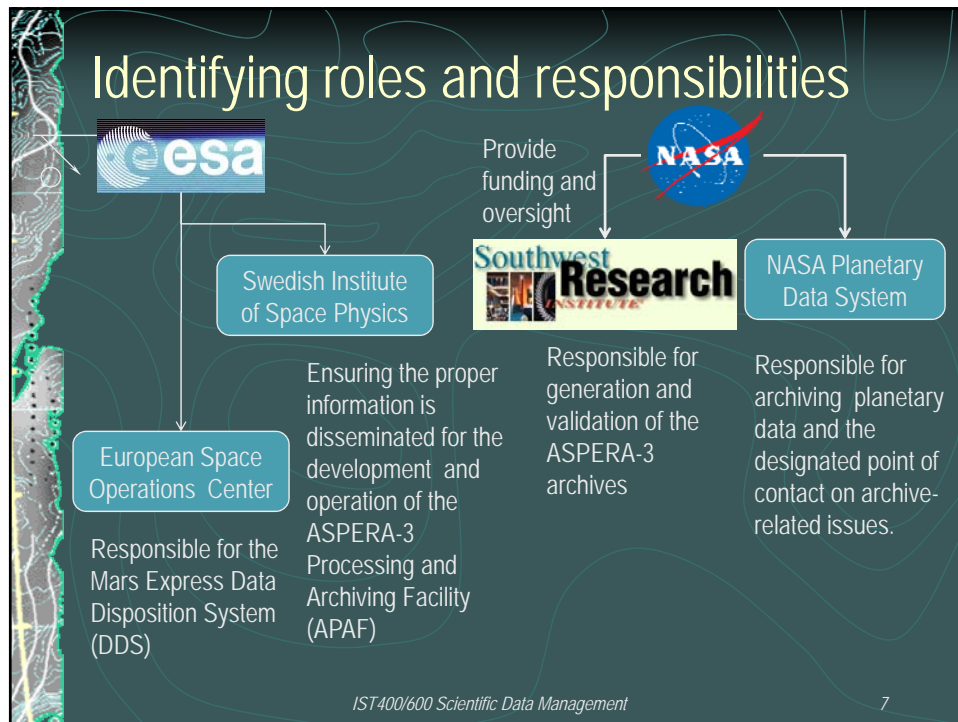
Processed Data Flow

```

            graph LR
            SwRI_APAF[SwRI APAF] --> IRF[IRF]
            IRF --> PSA[PSA]
            IRF --> CO_I_ESA((CO-I))
            SwRI_APAF --> CO_I_NASA((CO-I))
            SwRI_APAF --> PDS[PDS]
            subgraph NASA
            SwRI_APAF
            CO_I_NASA
            PDS
            end
            subgraph ESA
            IRF
            PSA
            CO_I_ESA
            end
            
```

SwRI = Southwest Research Institute
 APAF = ASPERA-3 Processing and Archiving Facility
 IRF = Swedish Institute of Space Physics
 ESOC = European Science Operations Center
 PSA = ESA Planetary Science Archive
 PDS = NASA Planetary Data System

IST400/600 Scientific Data Management 6



NASA definitions of processing levels

- **Packet Data:** Telemetry data stream as received at the ground station
- **Level 0:** Instrument science data (e.g., raw voltages, counts) at full resolution, time ordered, with duplicates and transmission errors removed.
- **Level 1A:** Level 0 data that have been located in space and may have been transformed (e.g., calibrated, rearranged) in a reversible manner and packaged with needed ancillary and auxiliary data (e.g., radiances with the calibration equations applied).
- **Level 1B:** Irreversibly transformed (e.g., resampled, remapped, calibrated) values of the instrument measurements (e.g., radiances, magnetic field strength).
- **Level 1C:** Level 1A or 1B data that have been resampled and mapped onto uniform spacetime grids.
- **Level 2:** Geophysical parameters, generally derived from Level 1 data, and located in space and time commensurate with instrument location, pointing, and sampling.
- **Level 3:** Geophysical parameters mapped onto uniform space-time grids.

IST400/600 Scientific Data Management

9

Standard data products

- ASPERA-3 data are archived in the Instrument Data File Set (IDFS) format
 - Data
 - Timing information
 - Metadata
- Archived at SwRI and accessed by the ASPERA-3 team for science data analysis

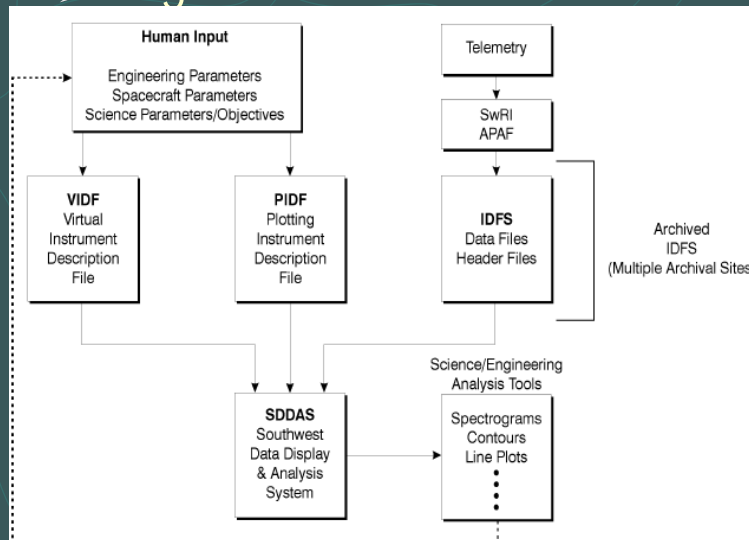
IST400/600 Scientific Data Management

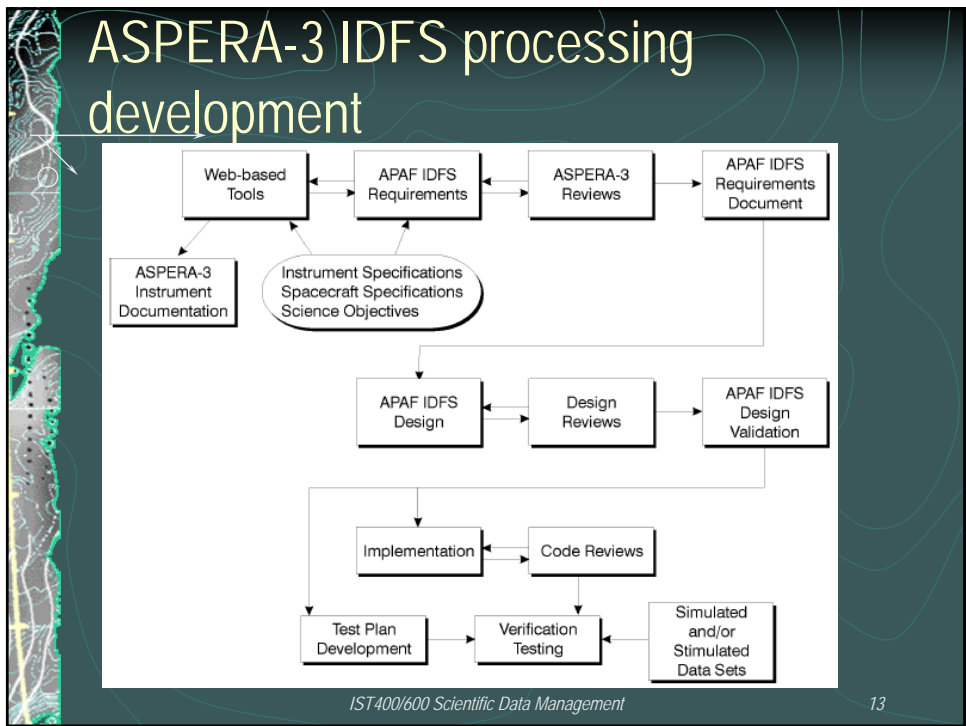
10

Data set documentation

- Data acquisition and processing history
 - Details of the data products to be archived
 - How they are generated
 - Validation process
 - Transfer procedure

Information flow for ASPERA-3 IDFS processing and use





Mars Express data system: data

<http://mexdata.space.swri.edu/data2.php>

Mars Express Data System

[HOME](#)
[DATA](#)
[PLOT](#)
[DOWNLOAD](#)
[LOGIN](#)

You are not logged in

Select one element from each list to create a hierarchy and then click one of the buttons below.

	Project	Mission	Experiment	Instrument
IDFS Hierarchy	MARS	MGS	ASPERA-3	ELS
	VENUS	Mars_Express	AUXILIARY MARSIS MODELS	IMA MU NPD NPI

Virtual Instrument

- ELSDLTL -- ELS Science Low Energy Range Data for Fixed Ener...
- ELSENG8 -- Electron Spectrometer 8-Bit Engineering Monitor D...
- ELSENGS -- Electron Spectrometer Status Data
- ELSSCIH -- Electron Spectrometer Science High Data
- ELSSCIL -- Electron Spectrometer Science Low Data
- ELSSWPH -- Electron Spectrometer Sweep High Data

Time Range

Beginning Year: Ending Year:

Day of Year: Day of Year:

[General Information](#)
[Available Data](#)

IST400/600 Scientific Data Management 14

Mars Express data system: plot

The screenshot shows the Mars Express Data System web interface. At the top, there are navigation tabs for HOME, DATA, PLOT, DOWNLOAD, and LOGIN. A message indicates "You are not logged in." Below this, there is a "Select a layout" dropdown menu currently set to "ELS Energy Flux Spectrograms (Common Scale)". To the left of the main form is a small thumbnail image of a spectrogram plot. The "Start time and duration" section includes a "Start Time" field with the value "2003-190/19:00:00" and a "Duration" field with the value "00:10:00:00". The "Advanced Options" section includes a "Plot Size" dropdown set to "700x600", "Multiple Windows" with radio buttons for "Yes" and "No" (where "No" is selected), and "Use Java" with radio buttons for "Yes" and "No" (where "Yes" is selected). A "Generate Plot" button is located at the bottom right of the form.

http://mexdata.space.swri.edu/plot_menu2.php

IST400/600 Scientific Data Management

15

What did we learn from this case? (1)

- Planning for data management involves understanding the science and engineering research
 - Instruments used for data collection
 - Data flows
 - Data products (formats, documentation)
 - Use of data

IST400/600 Scientific Data Management

16

What did we learn from this case? (2)

- Planning for data processing
 - Understand levels of processing
 - NASA definitions
 - CODMAC definitions
 - PSA definitions
 - USGS definitions
(http://landsat.usgs.gov/products_IP_levelsofproc.php)
 - Creating documentation about data
 - Metadata
 - Ancillary data

IST400/600 Scientific Data Management


17

What did we learn from this case? (3)

- Planning for long-term archiving and access
 - Storage media
 - Transfer
 - Format
 - Uniform structure of data files
 - Templates for dataset, instrument, reference, and personnel

IST400/600 Scientific Data Management

18



What did we learn from this case? (4)

- Querying and analyzing data
 - Information about datasets (metadata, or header file)
 - Tools:
 - Search for data files
 - Analysis of data (visualization)

IST400/600 Scientific Data Management 19