The Challenges of Using Mars Mission Datasets for Science Analysis

Jeffrey R. Johnson
United States Geological Survey
Flagstaff, AZ
jrjohnson@usgs.gov

April 20, 2006 J.R. Johnson, USGS
The “customer perspective”

• **Who are the customers (users) of mission data?**
  - Funded R&A PIs, Co-Is, Grad students, post-docs
  - Mission planners
  - Scientists interested in *proposing to use* the data for R&A programs

• **What are their goals?**
  - **Detailed analyses**
    - begin with raw data, calibrate it themselves, accurately correlate with surface features, other data sets, incorporate custom algorithms, etc.
  - **Simple answers**
    - does appropriate data exist for the region of interest?
    - what are standard characteristics of region of interest (altimetry, radiometry, composition...?)
At first glance the data sets are overwhelming...

- Mariner, Viking (lander, orbiter), Hubble
- Mars Pathfinder: IMP, APXS, ASI/MET
- Mars Global Surveyor: MOC, TES, MOLA, MAG/ER
- Mars Odyssey: THEMIS, GRS, NS, MARIE, HEND
- Mars Express: HRSC, OMEGA, MARSIS/ASPERA/SPICAM
- Mars Exploration Rovers: Pancam, Mini-TES, MB, APXS, MI, RAT, Navcam, Hazcam
- Mars Reconnaissance Orbiter: HiRISE, CRISM, CTX, MARCI, MCS, SHARAD
- **Phoenix**: RAD, MARDI, SSI, TEGA, MECA, MET
- **Mars Science Laboratory**: MastCam, MARDI, MAHLI, ChemCam, APXS, CheMin, SAM, DAN, RAD, REMS
Where’s the data?
Get thee to a PDS website....

http://pds.jpl.nasa.gov/
Or the Imaging Node....

http://pds-imaging.jpl.nasa.gov/
Or the Geosciences Node....
http://pds-geosciences.wustl.edu/

Welcome to the Geosciences Node

The Geosciences Node of NASA's Planetary Data System (PDS) archives and distributes digital data related to the study of the surfaces and interiors of terrestrial planetary bodies. We work directly with NASA missions to help them generate well-documented, permanent data archives. We provide data to NASA-sponsored researchers upon request, along with expert assistance in using the data. Our focus is mainly on serving the planetary science community, but we also provide some support for the general user interested in geoscience data.

Where's the Data?

Click on DATA AND SERVICES in the black navigation bar above to browse our data holdings.

The Geosciences Node is part of the Earth and Planetary Remote Sensing Laboratory in the Department of Earth and Planetary Sciences at Washington University in St. Louis.

April 20, 2006
Typical end-user complaints about mission data products (L. Crumpler):

- **The released data cannot be “used” easily by non-mission scientists:**
  - Searches are stymied by data format challenges
  - Data are too “raw” to be useful….
    - data are organized/archived along mission timelines/orbits instead of spatial location
    - …Or data are too highly processed (gridded, resampled) for users who wish to analyze “actual observations”
  - “Quick look” overviews of available data for particular regions require substantial time commitments
    - MOLA: “There is not an easy way to find the right track, so a lot of time can be spent downloading a track that turns out to be not the one that is needed.”
But there are *lots* of tools out there to locate and/or process data….

- **MOC**
- **THEMIS**
- **MER: Analyst’s Notebook:** [http://anserver1.eprsl.wustl.edu/](http://anserver1.eprsl.wustl.edu/)
- **Combined orbital data sets**
  - ASU---JMARS: [http://jmars.asu.edu/](http://jmars.asu.edu/)
Where’s the software?

• The increasing complexity of Mars datasets results in custom software (s/w) development by instrument teams that is often the only means to visualize/analyze data

  • This s/w may be released at end of mission (but not always)
  • This s/w may be well documented/supported and available for a variety of platforms (but not always)
    • ISIS provides support for some (but not yet all) missions/data sets
  • Instrument/mission team s/w may require other proprietary tools to execute (e.g., IDL), but not always
    • Use of “virtual” IDL allows s/w use, but limited production ability (?)
  • Lack of this s/w results in redundant/complementary efforts to write tools necessary to visualize/analyze data
What’s the problem?

• **Argument Style 1**

  • Mars state-of-the-art instruments (remote sensing, in-situ) produce unique datasets that are complicated. But there are lots of smart folks around who can deal with it.

  • Tools exist (or are in continuous development by missions and/or R&A folks) to search, quality-check, calibrate, spatially locate, and analyze many data sets.

  • If not, instrument documentation is always available to inform users how to work with the data.

  • Users need do their homework and learn how to work with complicated data sets.
What’s the problem?

• **Argument Style 2 (more productive)**

  • Complicated datasets have significant “learning curves” that limit effective and efficient use of the data to only the original instrument PI’s team and/or mission teams.

  • This inhibits widespread community involvement in understanding/interpreting datasets vital to our understanding of Mars and the future of the Mars Exploration Program.

  • These “roadblocks” cause some potential researchers to forego proposals to R&A, Participating Scientist programs.

    • “I thought about writing a proposal, but it I didn’t have the time to become sufficiently fluent in all the datasets, so I bailed….”
So what are the solutions? (1)

- Continue efforts toward synergizing data from multiple missions and making the data, software, derived products available in standardized, easily-accessible formats

  - Expand/connect GIS-related servers
    - JMARS, PIGWAD, JPL, Google?
    - cf. Report of the Mars Environmental GIS Workshop, Oct. 5-6, 2005
  - Continue to advertise new tools via DPS/EGS/AGU/LPSC abstracts
  - Fund instrument teams to conduct workshops (on-line or in-person) for their instruments and datasets. It’s E/PO!
    - TES, THEMIS have done this, others?
    - JMARS, ISIS already do this, too.
So what’s the solution? (2)

- Encourage release of “lite” versions of mission-related software prior to end-of-mission
  - Mainly for “quick-look” use, not necessarily detailed analysis
  - Curtail necessity of proprietary (expensive) s/w packages → IDLV?

- Missions need to have a cartographic planning component, and they should follow a consistent cartographic coordinate system and data format

- Planetary image data should be systematically processed in a coordinated way and made available in an easily accessed form (e.g., a GIS).

- But relatively small budgets permit only partial solutions – need to encourage increased funding.
Promote Dataset Synergy (aka PDS)

“GIS-ready” data that are easily used by researchers for display and analysis via Planetary Data Systems (PDS) or other

Better defined processing steps for commonly used/requested data sets

Easier tools/capabilities for processing raw data from each instrument into standard format(s)

Provide validated/calibrated data in both raw and map projected formats

Need improved data catalog / discovery capability whether the datasets reside at a centralized entity or individual research facility
System Architecture

PDS

Standard Formats

Generic Processing Software (e.g., GIS)

Other sources

Need to identify / define / develop these components

Report of the Mars Environmental GIS Workshop, Oct. 5-6, 2005
Report from Break-out Group #2: IT Considerations

GIS Recommendations (cont)

Report of the Mars Environmental GIS Workshop, Oct. 5-6, 2005

- Provide “on-line services” to help process datasets that are not easily derived as a single final product (e.g., MOC narrow angle, THEMIS visible images).
- Work with future mission planners to “task” instruments and define processing steps to meet geodetic standards.
- Develop outreach activities to educate the planetary community about the benefits of:
  - GIS software* for spatial analyses
  - Community-supported data formats

• *i.e., geospatial applications in general, preferably those that can share datasets

• "One application will simply not work for everyone." The goal is to standardize the formats and/or data distribution mechanisms to allow many applications to utilize the same data in different ways. --- Trent Hare