NASA Science Data Management

Scientific Data Management Workshop

Joe Bredekamp
NASA Science Mission Directorate

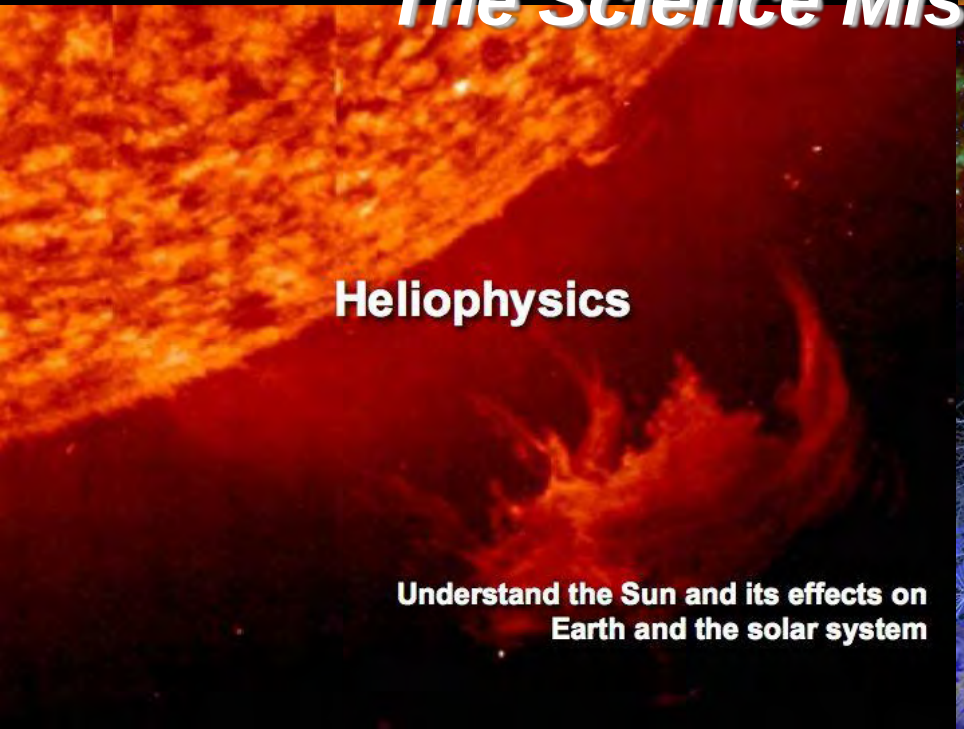
June 29, 2010

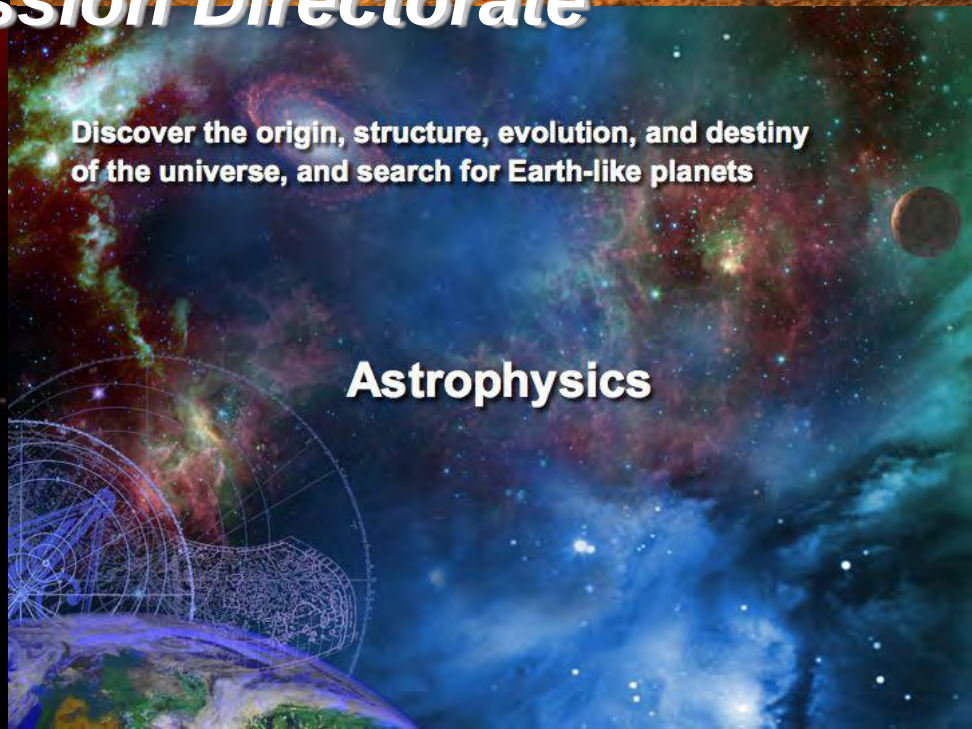
NASA Mission

To pioneer the future in space exploration, scientific discovery, and aeronautics research

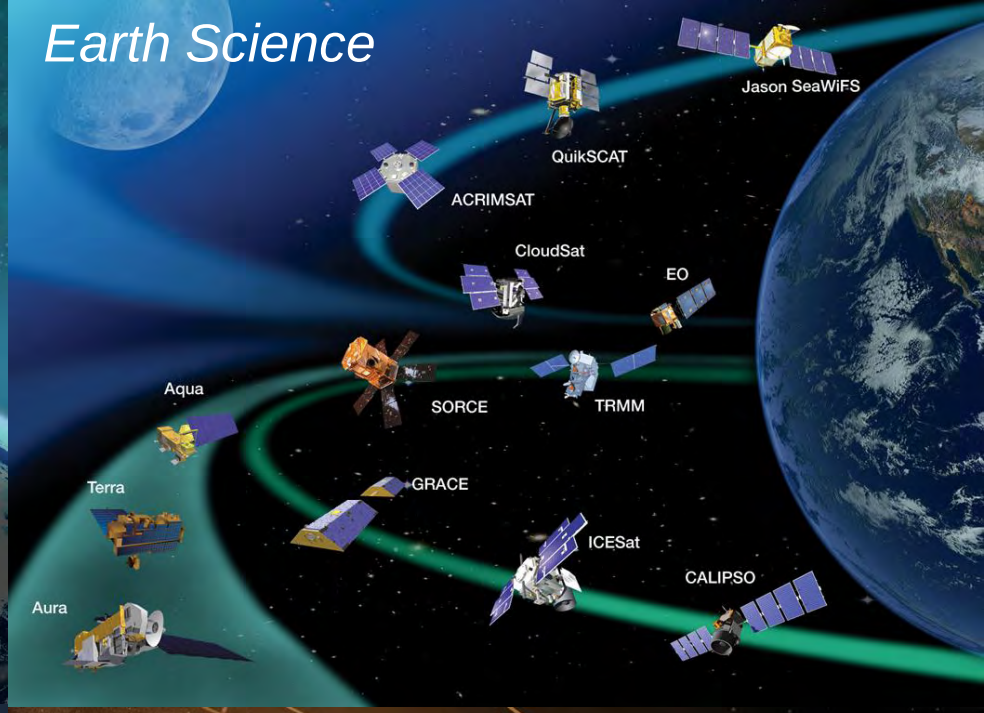
- ***Space Operations Mission Directorate*** *Extend the duration and boundaries of human space flight to create new opportunities for exploration and discovery*
- ***Exploration Systems Mission Directorate*** *Direct the identification, development, and validation of exploration systems and technologies*
- ***Aeronautics Mission Directorate*** *Enable a safer, more secure, efficient, and environmentally friendly air transportation system*
- ***Science Mission Directorate*** *Exploring the Earth-Sun system, our own solar system, and the universe beyond*

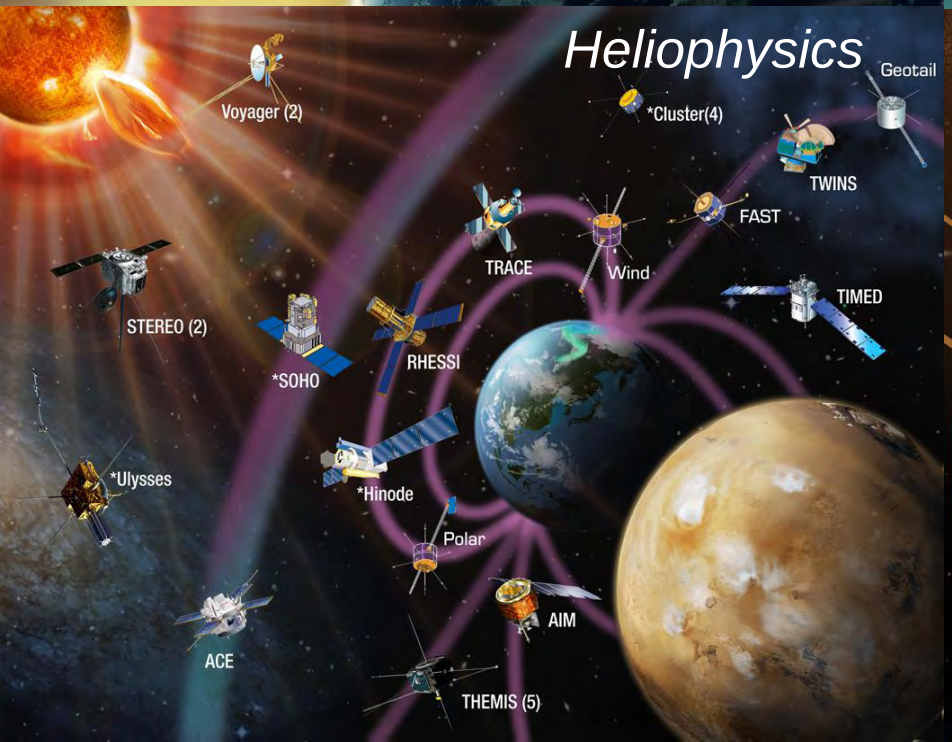
The Science Mission Directorate

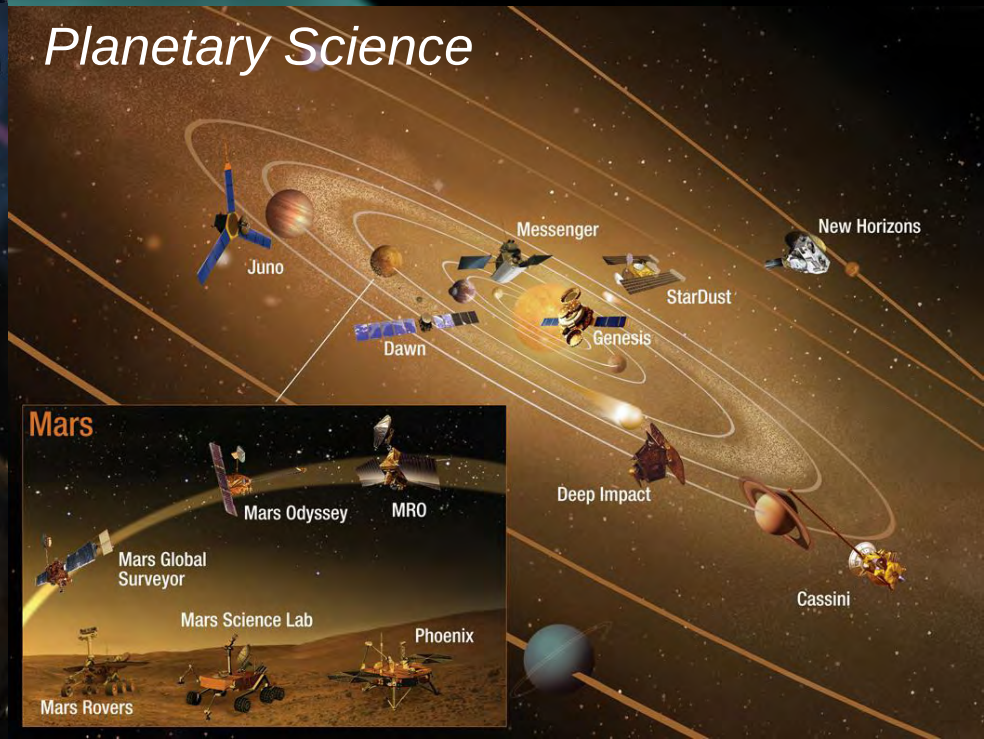




Astrophysics
RXTE
GALEX
JWST
Kepler
Chandra
WMAP
Swift
Spitzer
Hubble
XMM


Earth Science
Jason SeaWiFS
QuikSCAT
ACRIMSAT
CloudSat
EO
Aqua
SORCE
TRMM
Terra
GRACE
ICESat
CALIPSO
Aura


Heliophysics
Voyager (2)
Geotail
*Cluster(4)
TWINS
FAST
TRACE
Wind
TIMED
STEREO (2)
*SOHO
RHESSI
*Ulysses
*Hinode
Polar
AIM
ACE
THEMIS (5)


Planetary Science
Juno
Messenger
New Horizons
StarDust
Dawn
Genesis
Deep Impact
Cassini
Mars
Mars Odyssey
MRO
Mars Global Surveyor
Mars Science Lab
Phoenix
Mars Rovers

Science Computation and Information Management

Goal: *Provide comprehensive and robust infrastructure of data, computing, and modeling resources to maximize scientific productivity and knowledge enhancement*

Key Challenges:

- Multitude and diversity of missions
- Volume, richness, complexity, and breath of types of data
- Increasingly interdisciplinary nature of research
- Science cultural diversity
- Evolution: New technology and innovation
- Balancing competing constraints for open access vs. security/asset assurance

Science Data Management

Key Policy Objectives

- Stewardship responsibility for the preservation and utilization of science data assets as a national resource
- Open Data: Universal access and availability to science community, educators, and general public

Key Elements

- Project Data Management Plans: developed at onset of project describing data flows and delivery to public archive
- Timely delivery of data to public archives is a key performance measure throughout life of mission

Approach

- Each science division responsible for its data environment, often through partnerships with other federal agencies and international partners
- Coordination and federation as appropriate under auspices of SMD Chief Scientist

Science Mission Data Lifecycle

Mission Operations	***Science Data Operations***		***Research Operations***
Data Capture and Initial Processing	*Science Data Processing*	*Archival & Distribution*	*Scientific Inference and Discovery*
Generate time-ordered science data stream with transmission artifacts removed	*Calibration, other corrections, and generation of higher level data products*	*Ingest data products, metadata and supporting documentation*	*Explore vast data holdings to extract meaningful scientific insight and knowledge*
Mission Control and Data Handling Facilities	Mission Pipeline Processing Systems	Mission Data Archives	
	Science Investigator-led Processing Systems	Science Discipline-oriented Data Archives	

Key Instantiations:

EOSDIS Data Centers
Planetary Data System
Astrophysics Science Archive Research Centers
Heliophysics Data Environment: Solar Data Analysis Facility, Space Physics Data Facility

Lessons Learned/ Best Practices/ Future Vision

- One Size Doth Not Fit All:
  - Top-down mandates not generally effective
  - Organize by science discipline and allow for diversity and tailored implementations

- Clarify roles, responsibilities, requirements, “Rules of the Road” and other guidelines in management handbooks, Announcements of Opportunity, etc.

- Continuous User Involvement:
  - Science- driven around communities of practice
  - External working groups provide guidance and feedback

- Evolve the infrastructure to enhance quality, accessibility, and utilization of science data holdings
  - Provide advanced computational methods, algorithms, data management and analysis techniques, etc. through open calls that engage computer science and technology communities
  - Support archival data exploration for retrieval and analysis of mission datasets, as well as broader fusion and cross-correlation with other observations, datasets, theoretical models, etc.