Data Science and Computing on the Path to Autonomy

Richard J. Doyle

Program Manager, Information and Data Science, JPL

Jet Propulsion Laboratory
California Institute of Technology

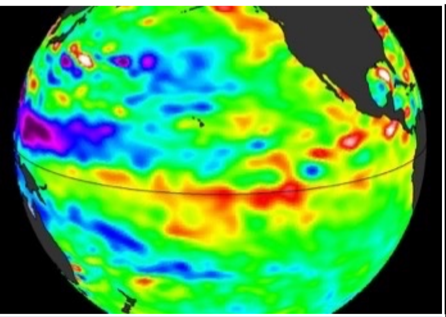
Astroinformatics 2019, Caltech, Pasadena

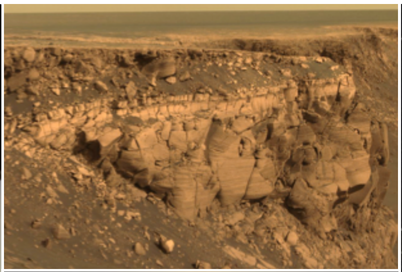
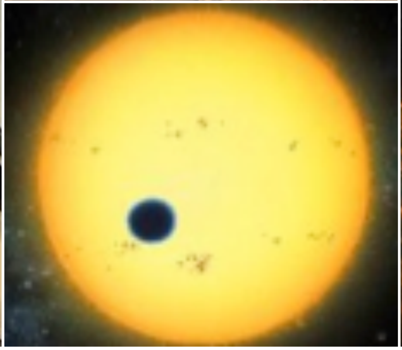
June 26, 2019

JPL's Mission for NASA

Robotic Space Exploration

Earth Science • Mars • Solar System • Astrophysics • Exoplanets • Interplanetary network

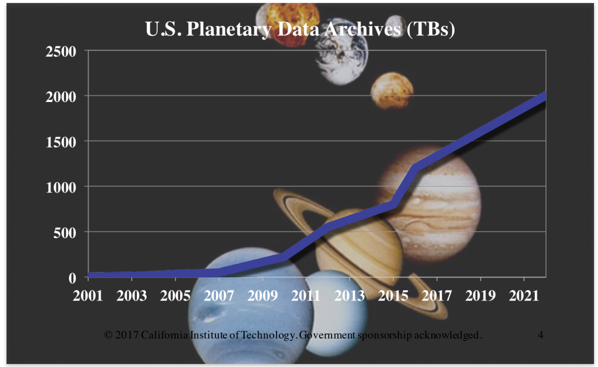



Our mission has introduced unique challenges for protecting space system assets and information

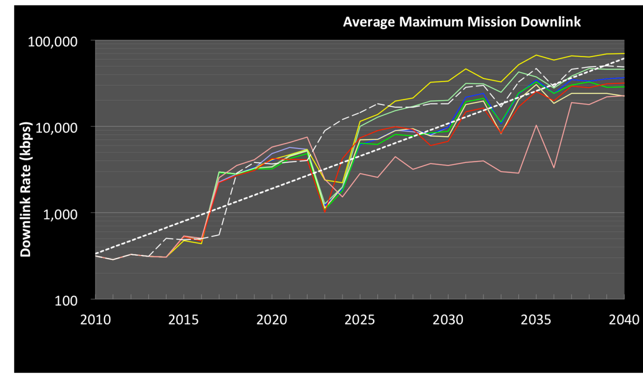
Data Challenges for Space Systems

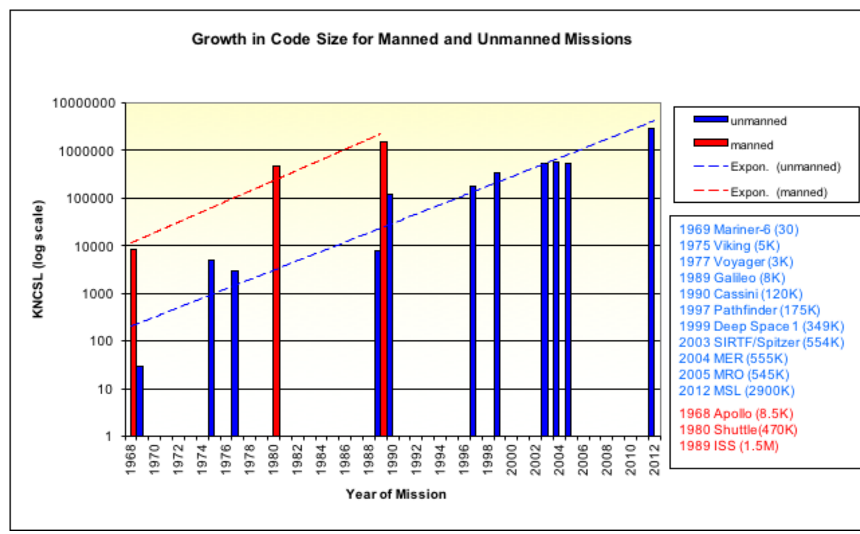
Scaling Pressures for NASA Missions



Increasing…

- ***data rates***
- ***data variety and complexity***
- ***data distributed across silos***
- ***challenges in data analysis***
- *data set sizes*
- *software complexity*



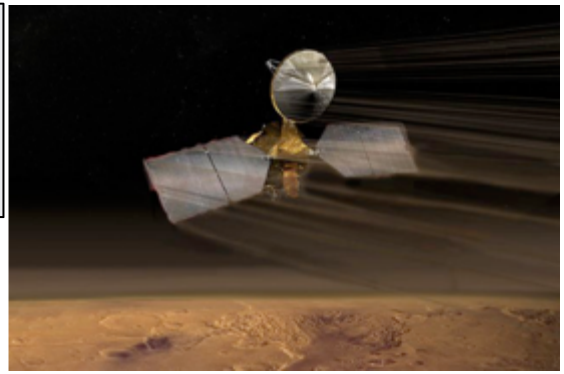


Emerging data science capabilities are critical to addressing these challenges

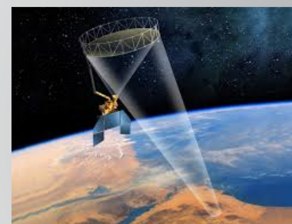
Data Lifecycle Model for NASA Missions

From Onboard Computing to Scalable Data Analytics

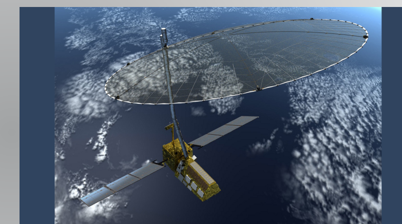
Emerging Solutions
- *Next-Generation Flight Computing*
- *Onboard Data Analytics*



Observational Platforms and Flight Computing



SMAP (Today): 485 GB/day



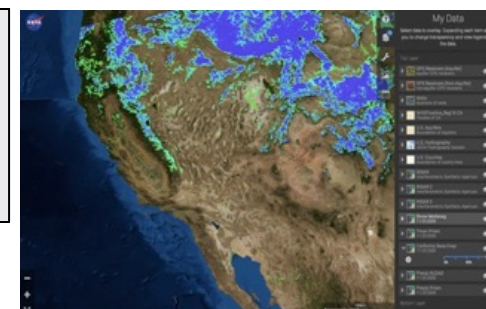
NISAR (2020): 86 TB/day

Scaling Pressures Expose the Need for an Integrated End-to-End Data and Computational Architecture

Emerging Solutions
- *Intelligent Ground Stations*
- *Agile Mission Operations*

Ground-based Mission Systems

Emerging Solutions
- *Data-Driven Discovery from Archives*
- *Scalable Computation and Storage*



Interactive Analytics and Visualization and Decision Support

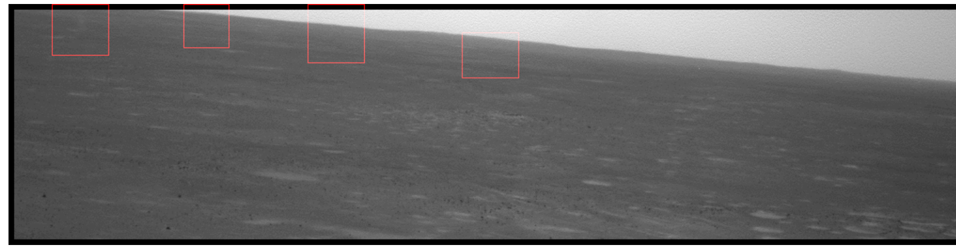
Opportunities Enabled by Data Science

1. Support <u>scalability</u> to capture and analyze NASA observational data

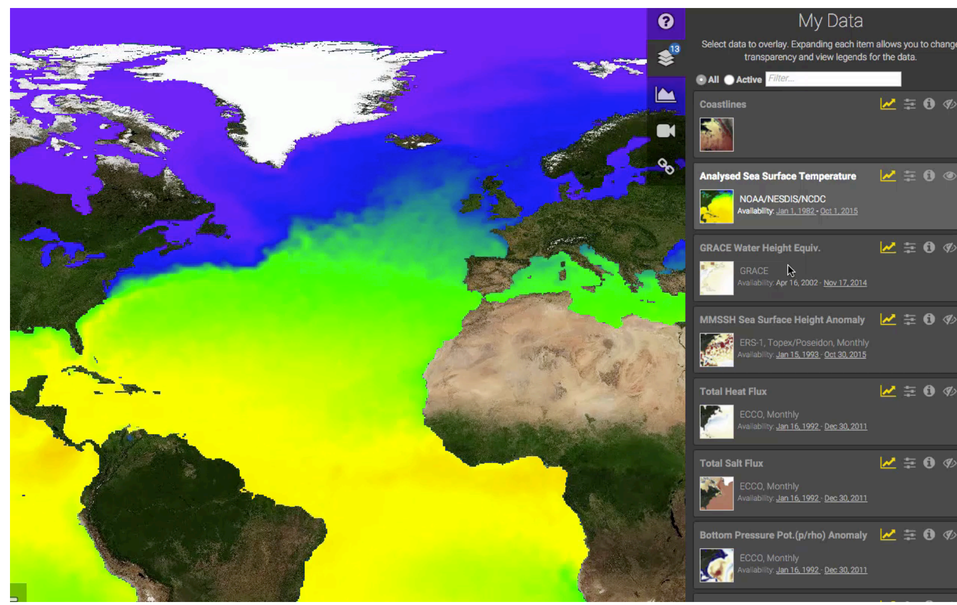
2. Apply <u>data-driven approaches</u> across the entire data lifecycle

3. Increase <u>access, integration and use</u> of highly distributed archival data

4. Increase <u>data science services</u> for on-demand, interactive visualization and analytics



Using ML to track events on Mars



Interactive Visualizations

Interoperable Ground Data Systems and Archives

An Analytics-Driven Ground Environment

Intelligent Ground Stations

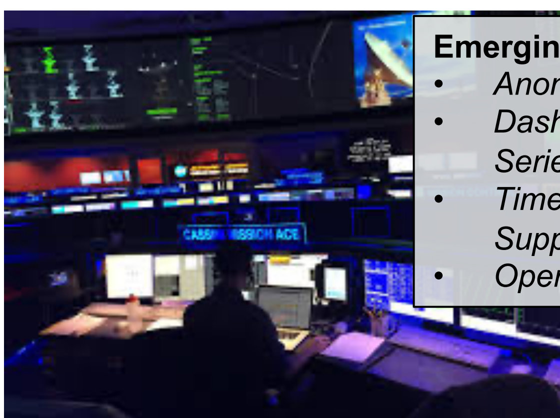
Emerging Solutions
- *Anomaly Detection*
- *Combining DSN & Mission Data*
- *Attention Focusing*
- *Controlling False Positives*

Data-Driven Discovery from Archives

Emerging Solutions
- *Automated Machine Learning - Feature Extraction*
- *Intelligent Search*
- *Integration of disparate data*

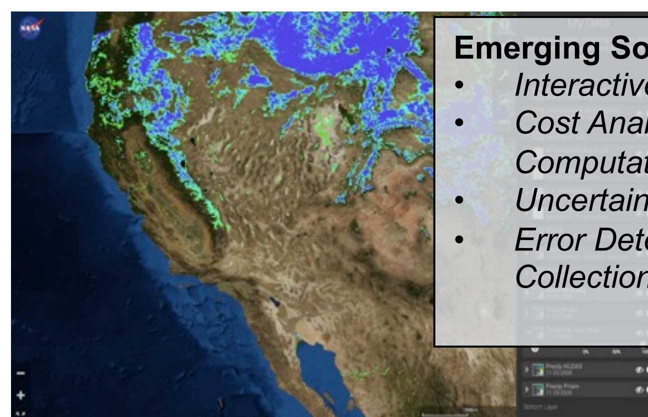
Technologies: Machine Learning, Deep Learning, Intelligent Search,
Data Fusion, Interactive Visualization and Analytics

Agile MOS-GDS



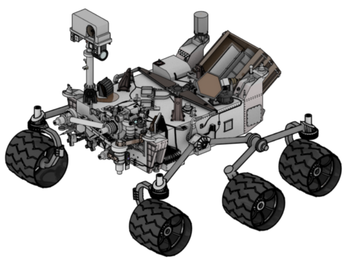
Emerging Solutions
- *Anomaly Interpretation*
- *Dashboard for Time Series Data*
- *Time-Scalable Decision Support*
- *Operator Training*

Data Analytics and Decision Support



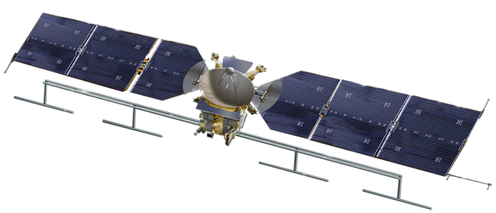
Emerging Solutions
- *Interactive Data Analytics*
- *Cost Analysis of Computation*
- *Uncertainty Quantification*
- *Error Detection in Data Collection*

ORCHIDS Infusion Pathways



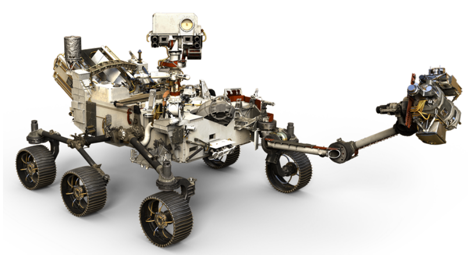
Mars Science Laboratory (MSL)

- Full engineering data system tracking
  - Commanding, telemetry, event monitoring, DSN monitoring, planning, scheduling, team staffing, vehicle parameters…
- Knowledge capture tool for human reporting
- Multi-subsystem in-flight data trending
- Anomaly investigations



Europa Clipper

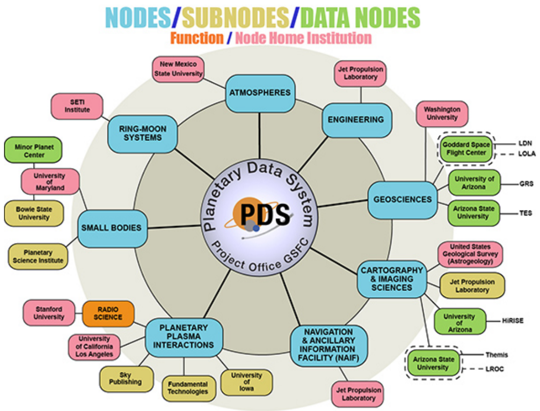
- Time correlation between science and engineering products
- Human knowledge capture
- Multi-instrument anomaly investigations



Mars 2020

- Active learning of anomalies - humans ‘on the loop’
- Increased data exploration capabilities in the mission operations environment
- Knowledge playback and training

NASA Archives: Access to Data*



International Planetary Data Alliance
IPDA

Planetary Science

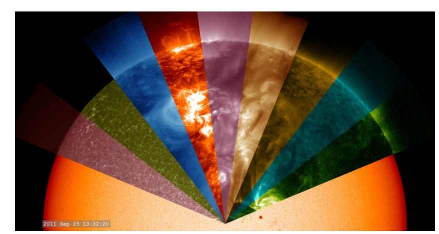
Highly
distributed/federated
Collaborative
Information-centric
Discipline-specific
Growing/evolving
Heterogeneous
International Standards &
Interoperability

Multiple Data Centers

IVOA

Astronomy

Earth Observation



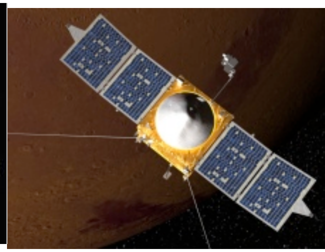
Multiple Data Centers

Heliophysics

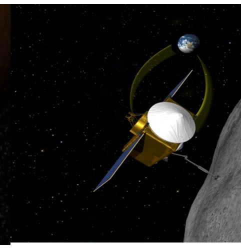
* Petascale environment that is moving to an exascale requirement

International Planetary Data Alliance: Building a World-wide Planetary Data Ecosystem

LADEE
(NASA)

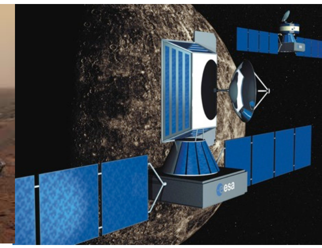


MAVEN
(NASA)

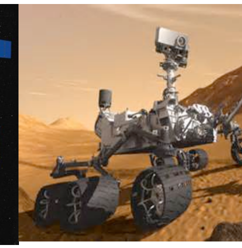


Osiris-REx
(NASA)

ExoMars
(ESA/Russia)

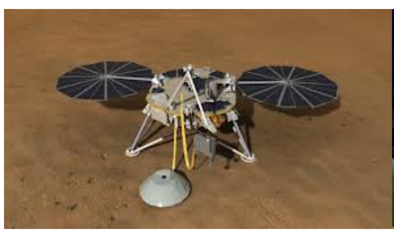


BepiColombo
(ESA/JAXA)



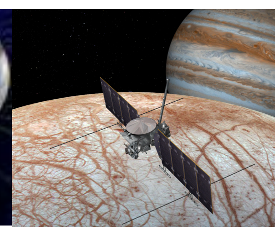
Mars 2020
(NASA)

Psyche
(NASA)



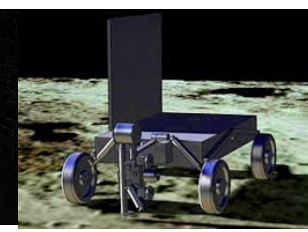
InSight
(NASA)

JUICE
(ESA)

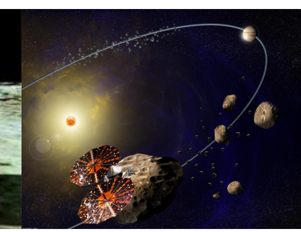


Europa
(NASA)

Hyabusa-2
(JAXA)



Chandrayaan-2
(ISRO)

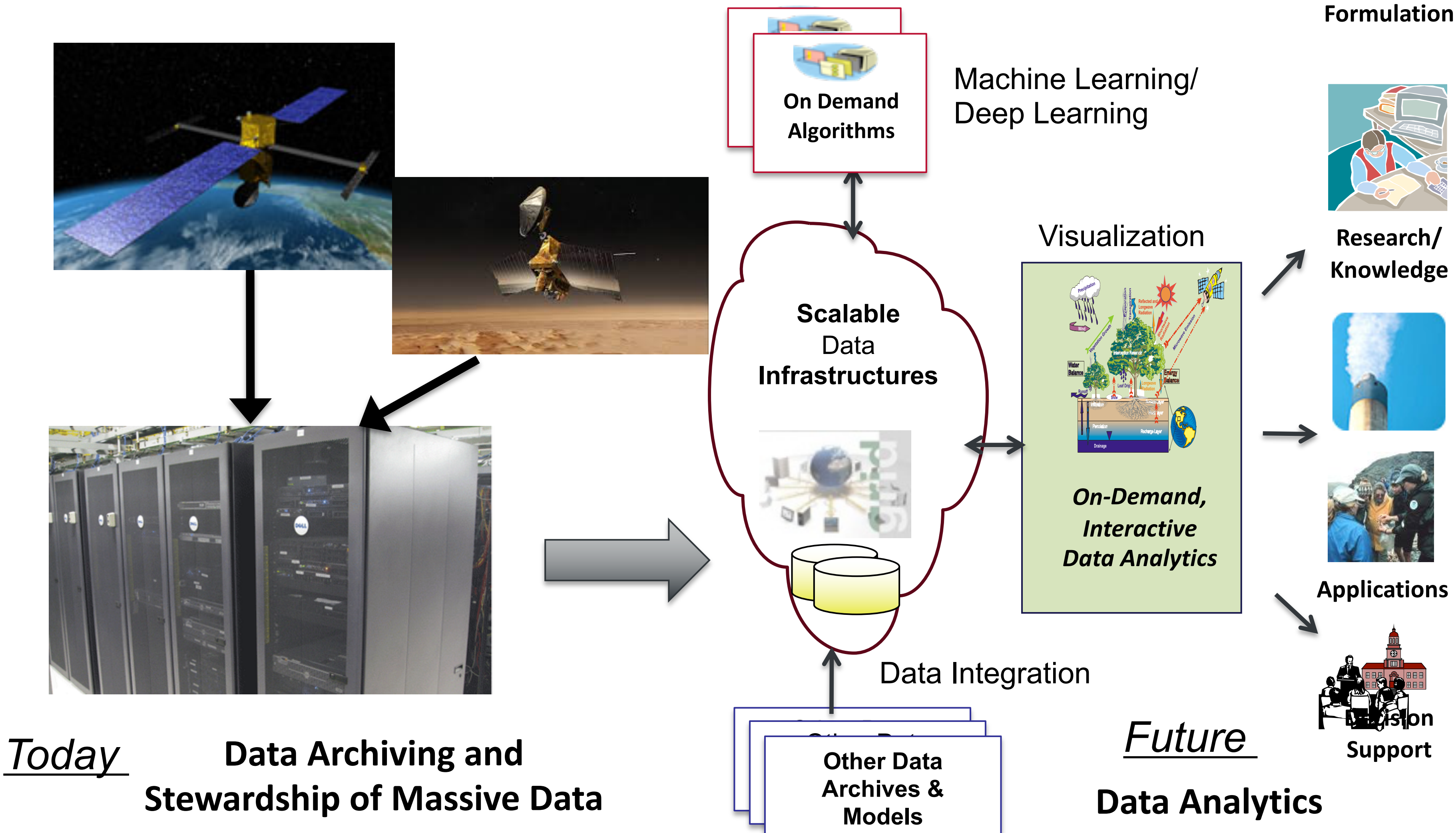


Lucy
NASA

Endorsed by the **International Planetary Data Alliance** in July 2012 –
https://planetarydata.org/documents/steering-committee/ipda-endorsements-recommendations-and-actions

Shifting toward Data Analytics

Expanding to Data-Driven Analytics



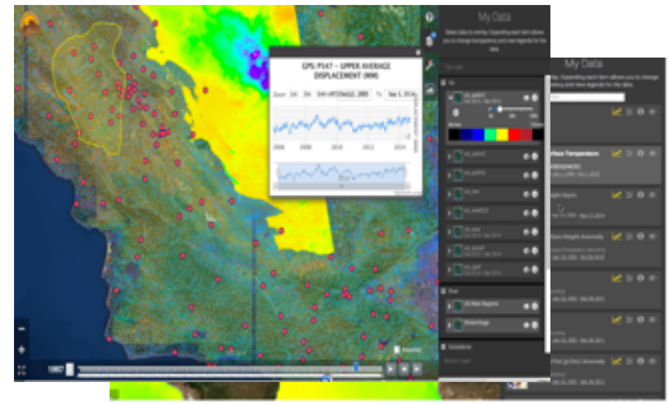
Reducing Data Wrangling: “There is a major need for the development of software components… that link high-level data analysis-specifications with low-level distributed systems architectures.”
Frontiers in the Analysis of Massive Data, National Research Council, 2013.

Interactive Analytics for Data Exploration

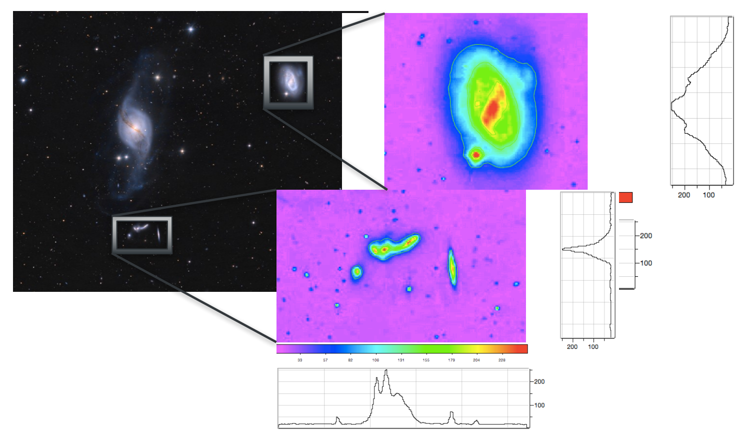
Exploration of Data from Planetary Missions

Examples: Planetary Image search, Mars and Moon surface navigation, feature extraction from Planetary images.

Examples: Hydrology and sea level rise

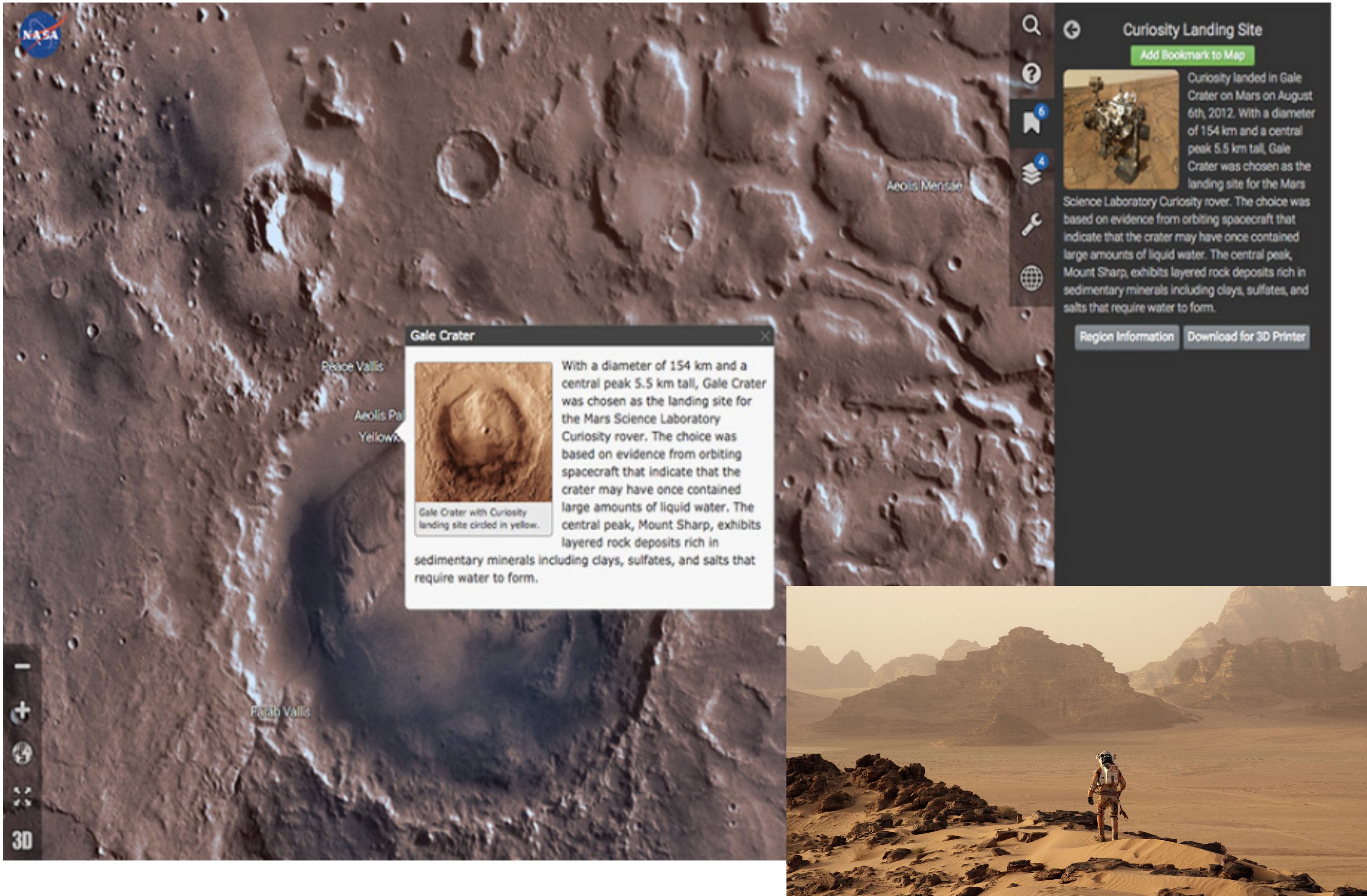


Analysis of Earth and Climate Science Observations and Models



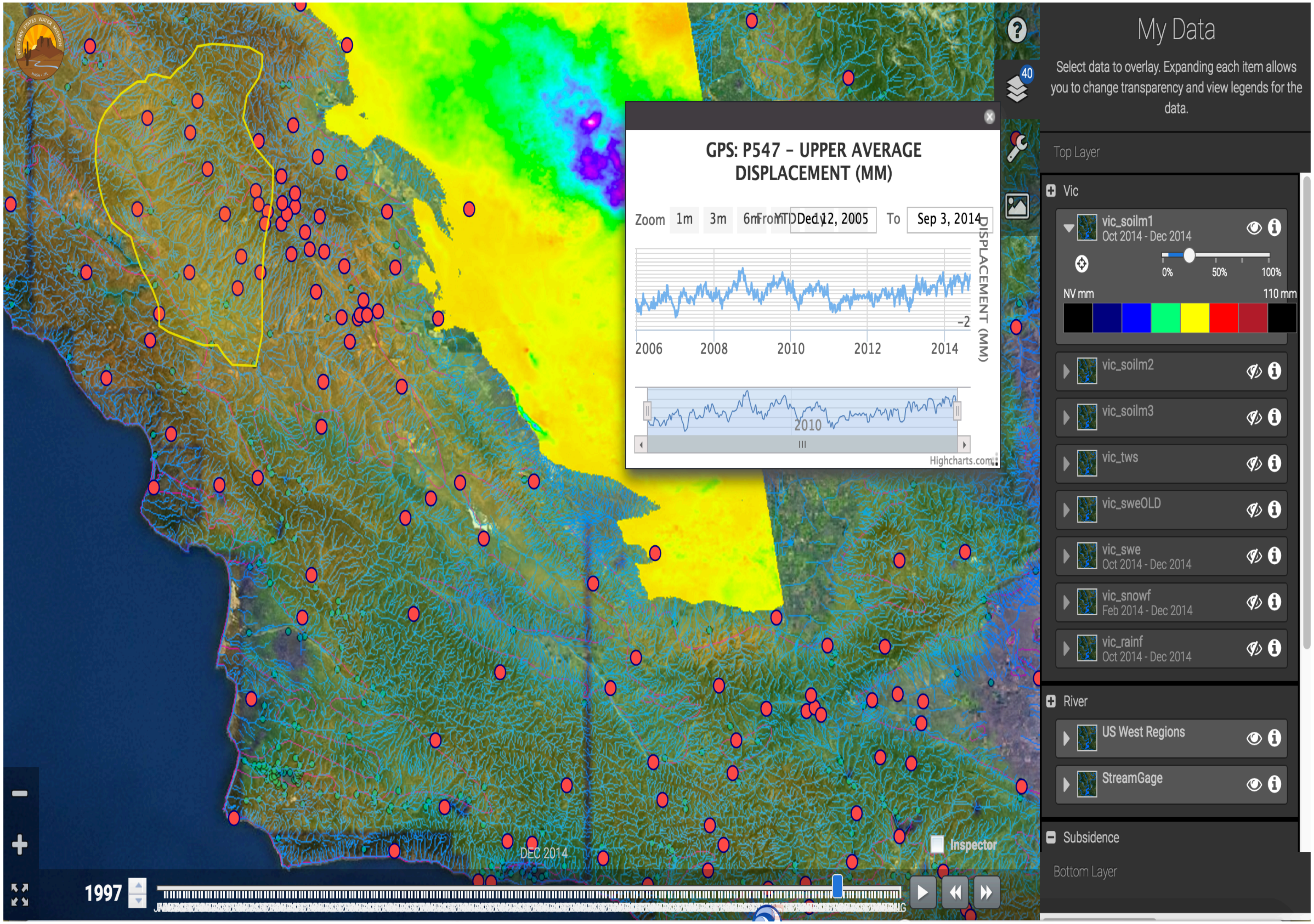
Classification and Analysis of Transient Events in Astronomy

Mars Trek: The Google Earth of Mars



Credit: Emily Law, Shan Malhotra

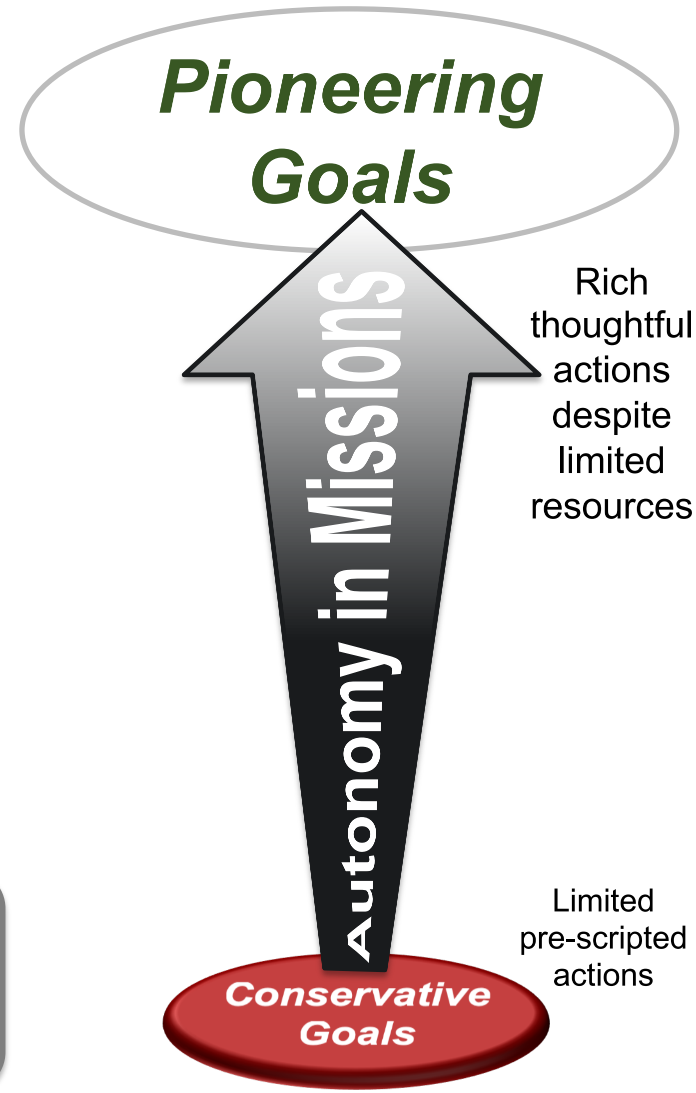
WaterTrek: Interactive Analytics for Western States Water Analysis



Credit: Jay Famiglietti, Cedric David, Shan Malhotra, D. Crichton

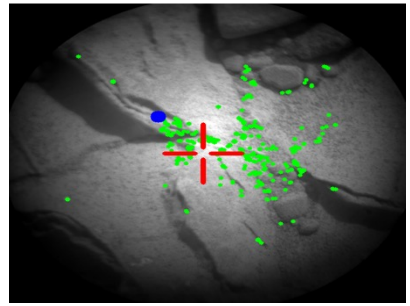
Computing, Autonomy and Analytics at the Far Edge

The Need for Autonomy



Autonomy enables pioneering missions to: Explore new destinations and increase science yield, robustness, operability

Enabling Onboard Autonomy Through Data Science and Computing

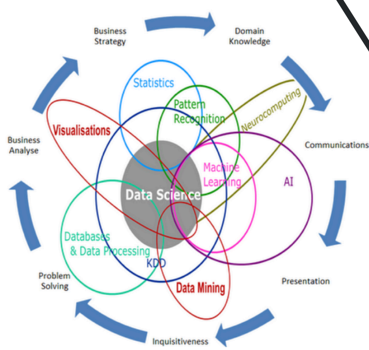


Onboard Autonomy

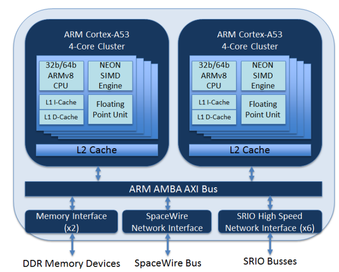
SPACE MISSION ARCHITECTURE

<u>Capabilities:</u>

- Data Analytics
- Scalable Data Services
- Data Processing
- Data Integration



Data Science



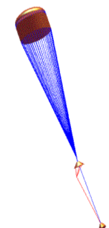
<u>Capabilities:</u>

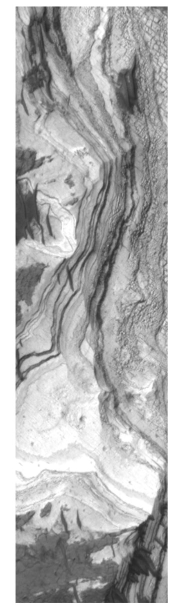
- Processors
- Memory
- Energy Mgmt
- Fault Tolerance

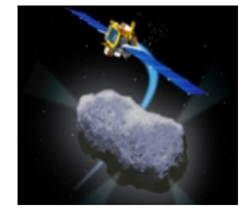
High Performance Spaceflight Computing

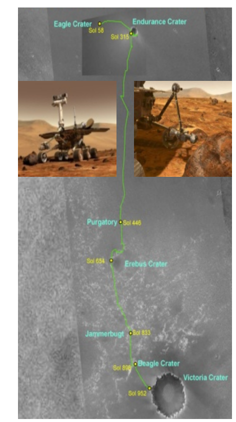
Future NASA Needs for Space-Based Computing

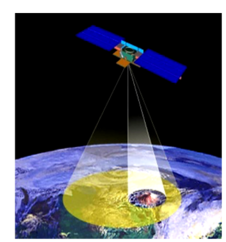
- Space-based computing has not kept up with the needs of current and future missions
- NASA has some unique requirements
  - Deep space, long duration, robotic and human missions
  - Higher performance, smaller spacecraft
    - Onboard science data processing
    - Autonomous operations
  - Extreme needs for low power and energy management, efficiency, fault tolerance and resilience

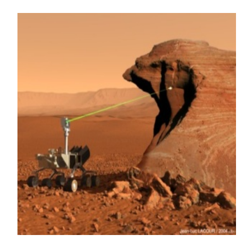








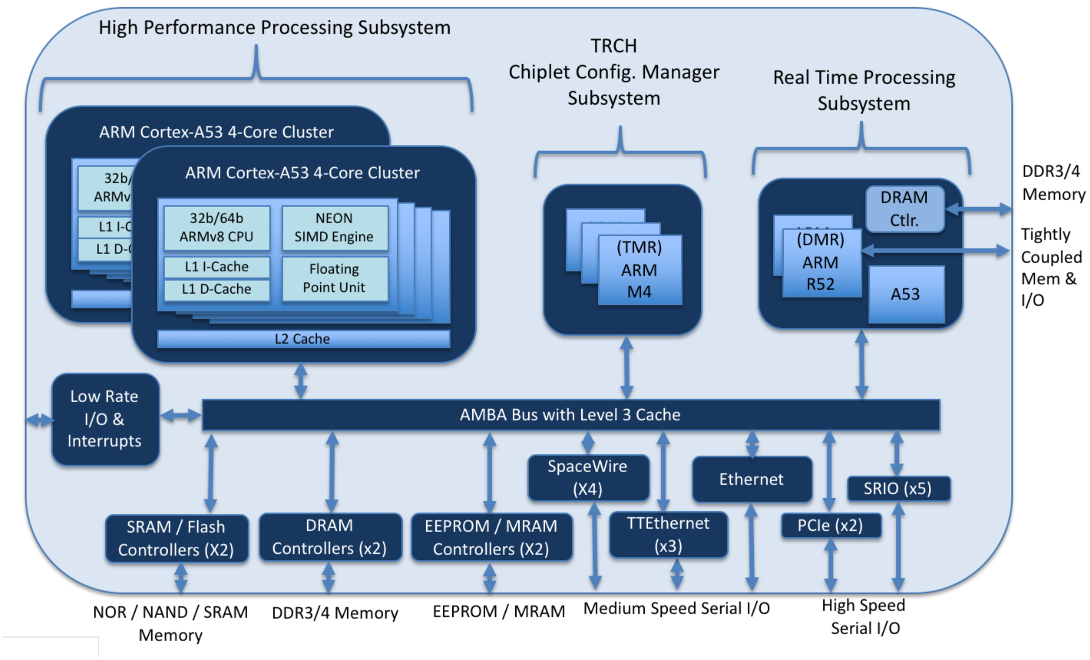


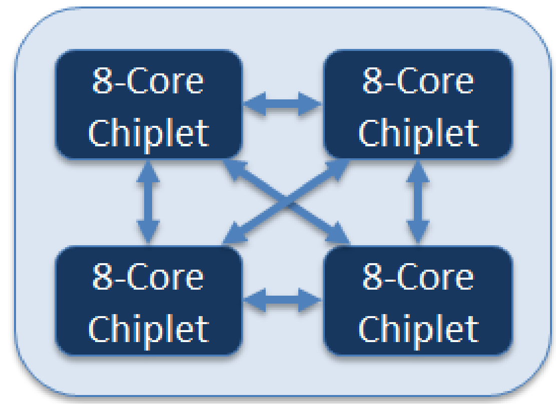


NASA Flight Computing Drivers

Computation Category	Mission Need	Objective of Computation	Flight Architecture Attribute	Processor Type and Requirements
Vision-based Algorithms with Real-Time Requirements	• Terrain Relative Navigation (TRN) • Hazard Avoidance • Entry, Descent & Landing (EDL) • Pinpoint Landing	• Conduct safe proximity operations around primitive bodies • Land safely and accurately • Achieve robust results within available timeframe as input to control decisions	• Severe fault tolerance and real-time requirements • Fail-operational • High peak power needs	• Hard real time / mission critical • Continuous digital signal processing (DSP) + sequential control processing (fault protection) • High I/O rate • Irregular memory use • General-purpose (GP) processor (10's – 100's GFLOPS) + high I/O rate, augmented by co-processor(s)
Model-Based Reasoning Techniques for Autonomy	• Mission planning, scheduling & resource management • Fault management in uncertain environments	• Contingency planning to mitigate execution failures • Detect, diagnose and recover from faults	• High computational complexity • Graceful degradation • Memory usage (data movement) impacts energy management	• Soft real time / critical • Heuristic search, data base operations, Bayesian inference • Extreme intensive & irregular memory use (multi-GB/s) • > 1GOPS GP processor arrays with low latency interconnect
High Rate Instrument Data Processing	High resolution sensors, e.g., SAR, Hyper-spectral	• Downlink images and products rather than raw data • Opportunistic science	• Distributed, dedicated processors at sensors • Less stringent fault tolerance	• Soft real time • DSP/Vector processing with 10-100's GOPS (high data flow) • GP array (10-100's GFLOPS) required for feature ID / triage

HPSC – Reinventing the Role of Computing in Space





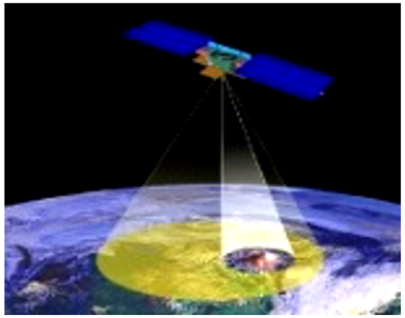
- **HPSC** offers **a new flight computing architecture** to meet the future needs of NASA missions.
- Providing on the order of **100X the computational capacity of current flight processors for the same amount of power**, the multicore architecture of the HPSC chiplet provides **unprecedented flexibility** in a flight computing system.
- By enabling the operating point to be set dynamically, **trading among needs for computational performance, energy management and fault tolerance**.
- **HPSC has been conceived to be highly extensible**. Multiple **chiplets can be cascaded together** for more capable computing, or HPSC **can be configured with specialized co-processors** to meet the needs of specific payloads and missions.
- **HPSC is a technology multiplier**, amplifying existing spacecraft capabilities and enabling new ones.
- The HPSC team anticipates that the chiplet will be **used by virtually every future space mission**, all benefiting from more capable flight computing.

HPSC – Mission Infusion Framework

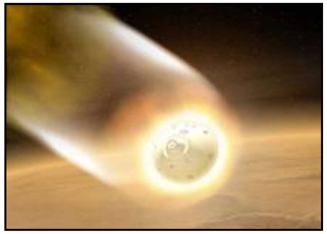
- NASA will develop HPSC-based, flight-qualified, single board computers (SBCs), ready for infusion into missions
  - Develop a NASA SBC reference design
  - Integrate the board with at least one set of flight-ready system software
  - Demonstrate flight readiness of the single board computer
  - Fund industry to develop standards-based HPSC SBCs

Deep Space

Surface Systems



High Data Rate Instruments



Landing Systems

Human Spaceflight

CubeSats, SmallSats

Entry, Descent and Landing

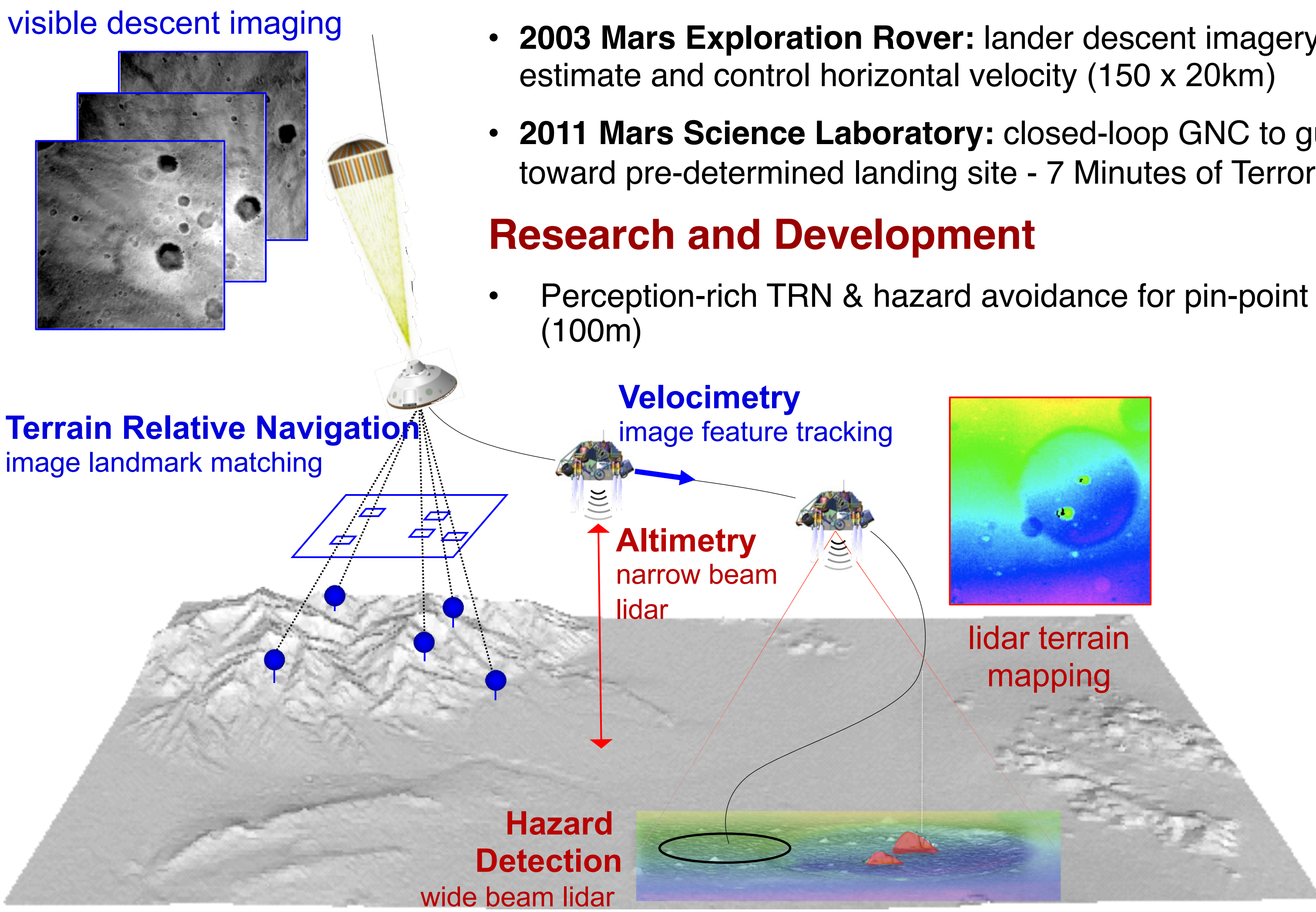
Terrain Relative Navigation and Hazard Avoidance

Flight Deployed

- **2003 Mars Exploration Rover:** lander descent imagery used to estimate and control horizontal velocity (150 x 20km)
- **2011 Mars Science Laboratory:** closed-loop GNC to guide EDL toward pre-determined landing site - 7 Minutes of Terror (20 x 7km)

Research and Development

- Perception-rich TRN & hazard avoidance for pin-point landing (100m)



Surface Mobility

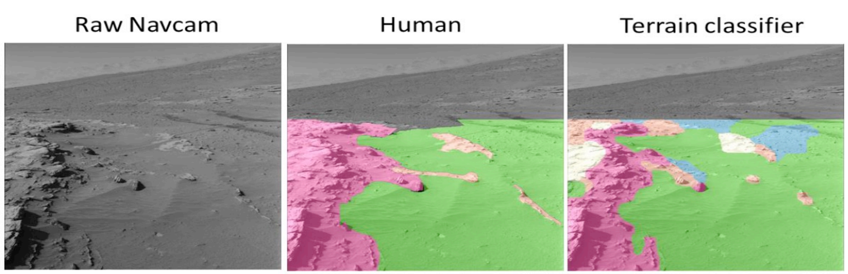
Mars Rover Navigation

Flight Deployed

- **1996 Mars Pathfinder:** obstacle avoidance with structured light
- **2003 Mars Exploration Rover:** obstacle avoidance with stereo vision; pose estimation and slip detection with visual odometry; goal tracking
- **2011 Mars Science Laboratory:** enhanced obstacle avoidance, visual odometry and goal tracking

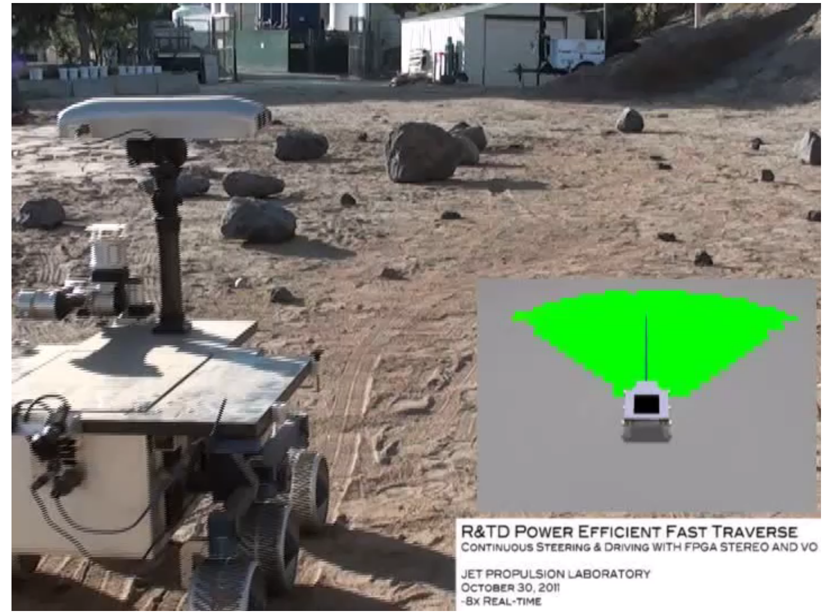
Research and Development

- Enhanced hazard detection, traversability analysis and motion planning for Mars 2020 and beyond



Terrain Classification

Athena



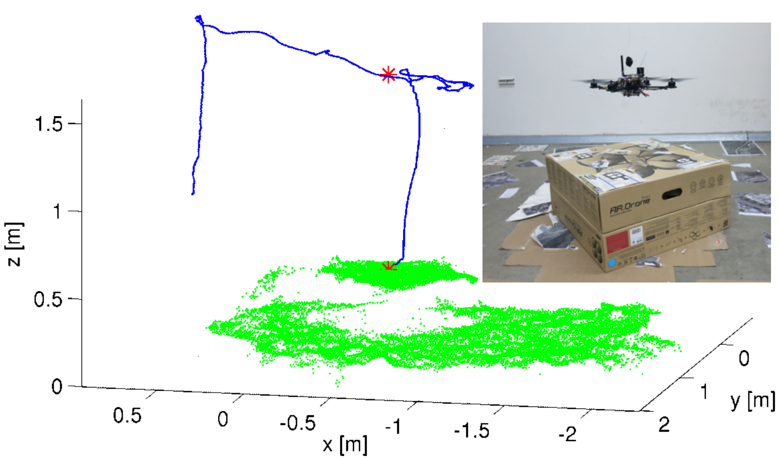
Fido

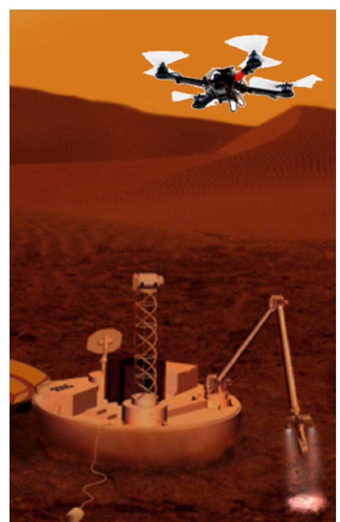
Above-Surface Mobility

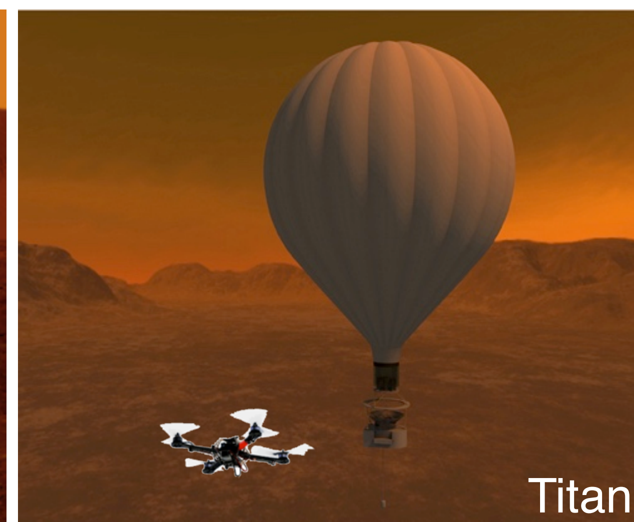
Rotorcraft and Balloon Mobility Research

Research and Development

- **Multiple applications:** (a) terrestrial (defense, intelligence, commercial, and science) and (b) planetary (Mars, Titan, and Venus)
- **Capabilities:** visual-inertial localization combines images with IMU for better estimate; autonomous landing with obstacle avoidance







Onboard Data Product Generation

Dust Devils on Mars

Dust devils are scientific phenomena of a transient nature that occur on Mars

- They occur year-round, with seasonally variable frequency
- They are challenging to reliably capture in images due to their dynamic nature
- Scientists accepted for decades that such phenomena could not be studied in real-time

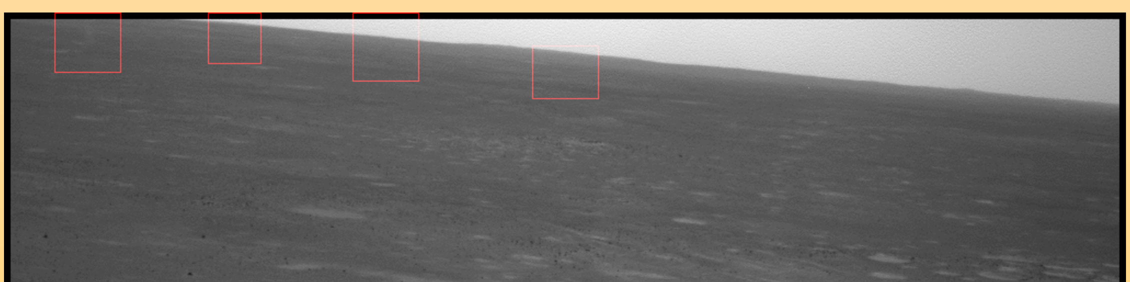
Spirit Sol 543 (July 13, 2005)

New onboard Mars rover capability (as of 2006)

- Collect images more frequently, analyze onboard to detect events, and only downlink images containing events of interest

Benefit

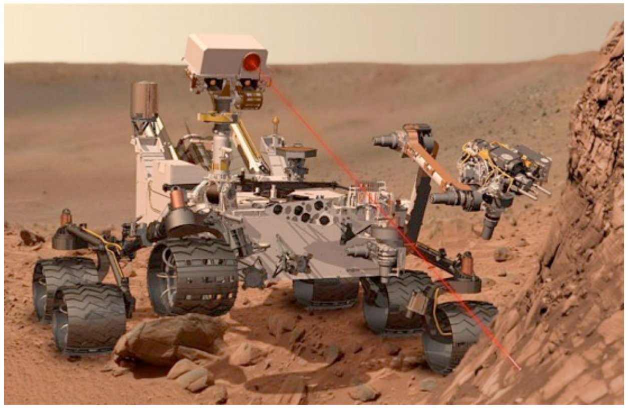
- < 100% accuracy can dramatically increase science event data returned to Earth
- *First notification includes a complete data product*



Credit: T. Estlin. B. Bornestein, A. Castano, J. Biesiadecki, L. Neakrase, P. Whelley, R. Greeley, M. Lemmon, R. Castano, S. Chien and MER project team

AEGIS Automated Targeting System
(Autonomous Exploration for Gathering Increased Science)

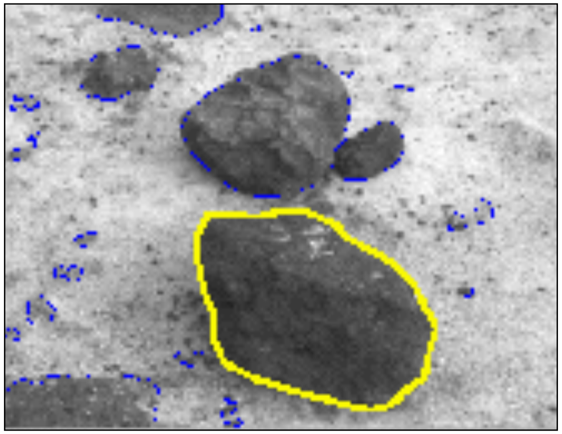
Description



- Intelligent targeting and data acquisition by
  - analyzing images of the rover scene
  - identifying high-priority science targets (e.g., rocks), *and*
  - taking high quality data of these targets autonomously with no ground communication required

Benefits

- New way of autonomously collecting rover science data



Status

Missions Deployed

- Mars Exploration Rover (MER) Mission
  - Found crater ejecta, outcrop, boulders, …
- Mars Science Laboratory (MSL) Curiosity Rover
  - Acquires data for ChemCam laser spectrometer

– Planned for M2020

- Winner of 2011 NASA Software of the Year Award

Partnering

Caltech-JPL Partnership in Data Science

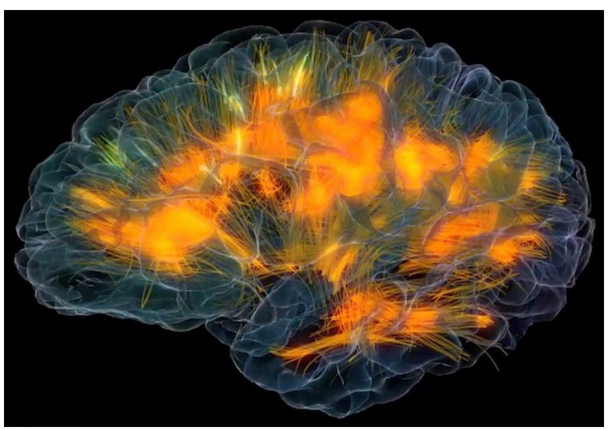
Center for Data-Driven Discovery on campus/Center for Data Science and Technology at JPL

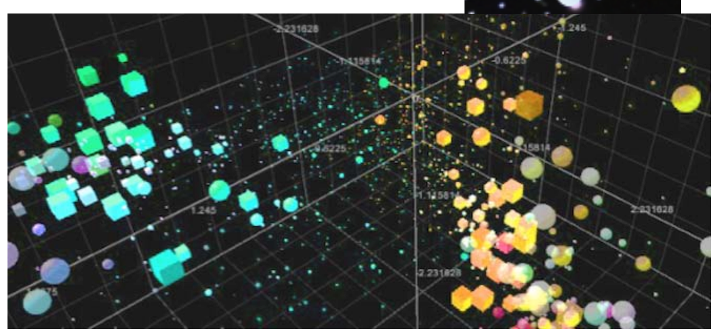
From basic research to deployed systems ~10 collaborations

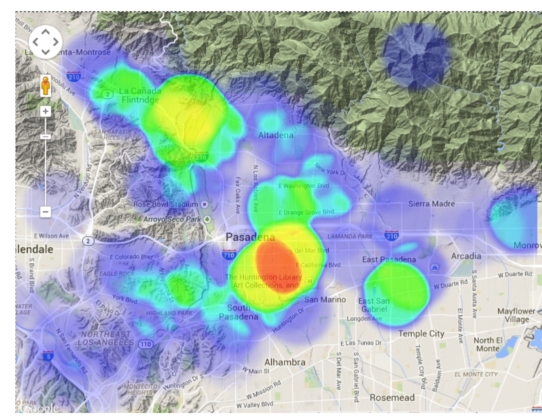
Leveraged funding from JPL to Caltech; from Caltech to JPL

CENTER FOR DATA-DRIVEN DISCOVERY

CDST







JPL Data Science Partnering Strategy

Universities

Caltech
UNIVERSITY OF CALIFORNIA
CSAIL

Non-NASA partnerships

NSF
U.S. DEPARTMENT OF ENERGY
NIH
National Institutes of Health

Open Source

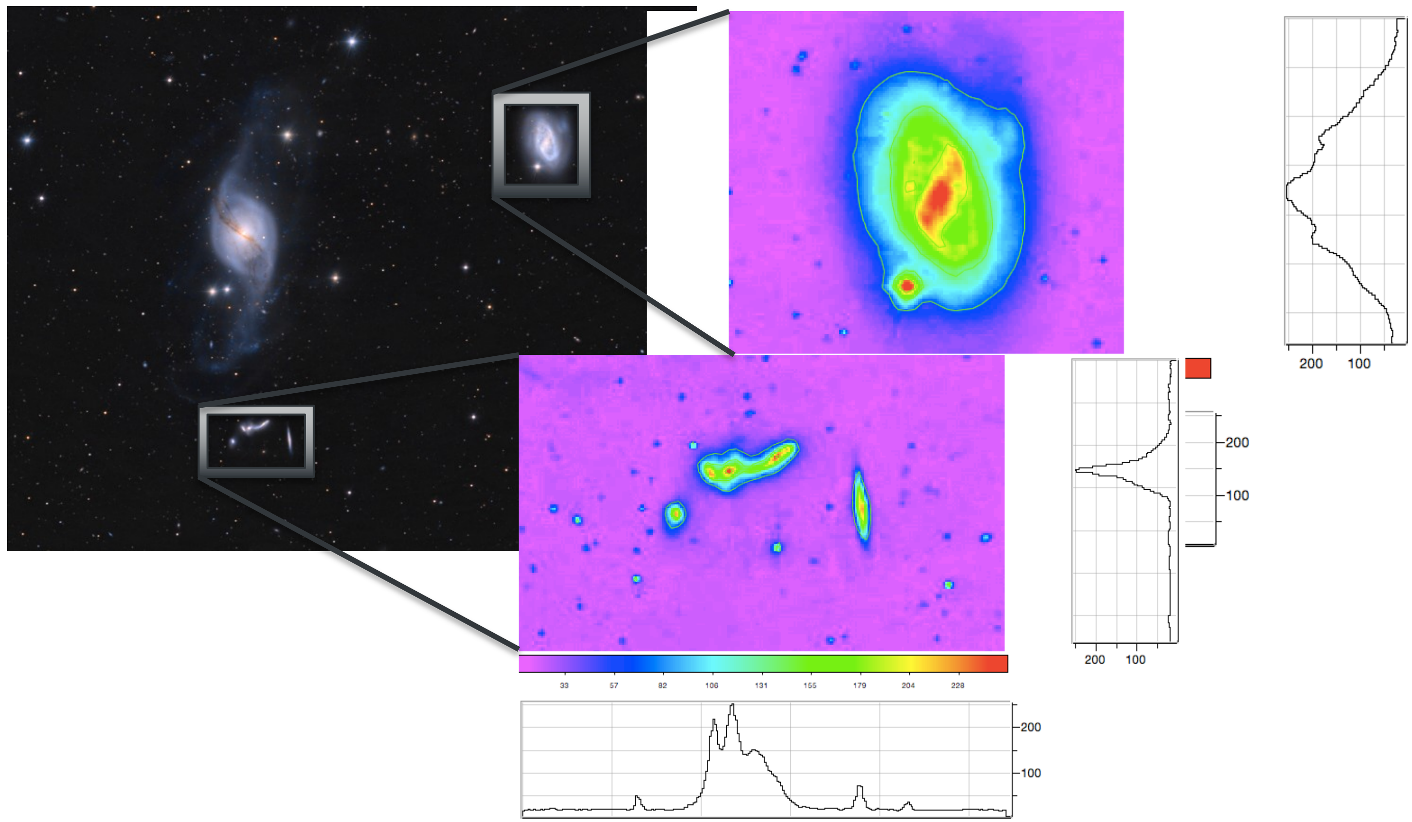
THE APACHE SOFTWARE FOUNDATION

Commercial

amazon web services
IBM Watson Analytics
NVIDIA

Methodology Transfer

From Astrophysics…

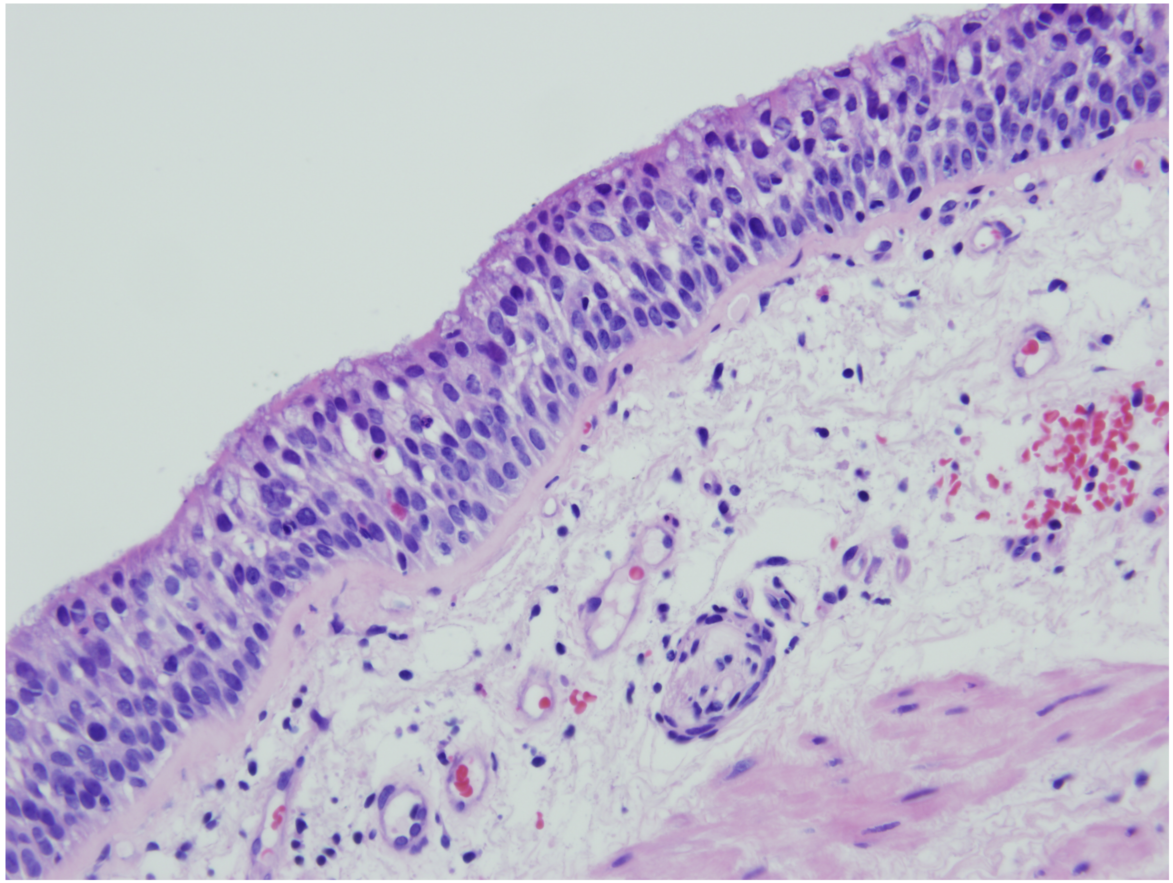


Description: Detecting objects from astronomical measurements by evaluating light measurements in pixels using intelligent software algorithms.

Image Credit: Catalina Sky Survey (CSS), of the Lunar and Planetary Laboratory, University of Arizona, and Catalina Realtime Transient Survey (CRTS), Center for Data-Driven Discovery, Caltech.

Methodology Transfer

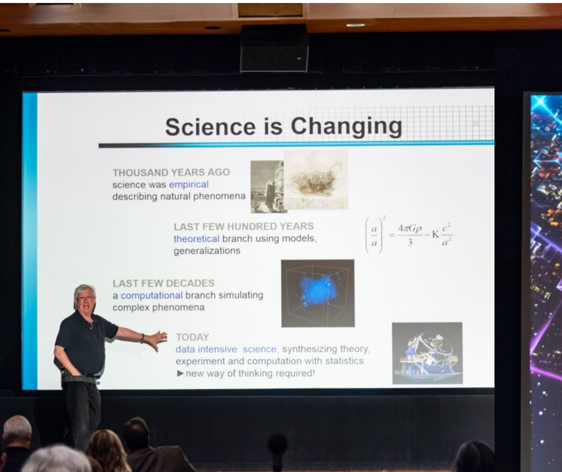
…to Biomedicine

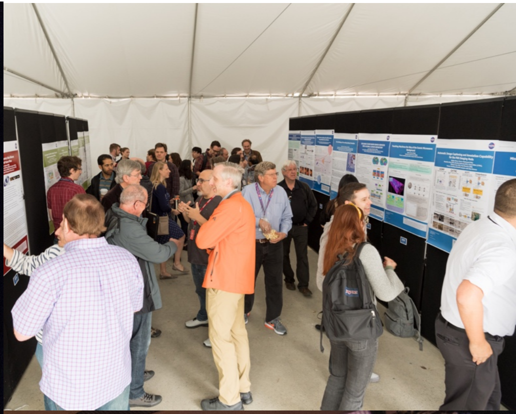


Description: Detecting objects from oncology images using intelligent software algorithms transferred to and from space science.

Image Credit: EDRN Lung Specimen Pathology image example, University of Colorado

An Emerging Community of Practice





2 Keynotes (Johns Hopkins, Caltech)
6 JPL Talks, 96 Posters, and live demos across JPL
Over 1000 attendees
…an emerging community of practice

NASA JPL Caltech