

1000 Silos Planetary Data From An Engineer's Viewpoint

Eric Lyness

GSFC Planetary Environments Lab

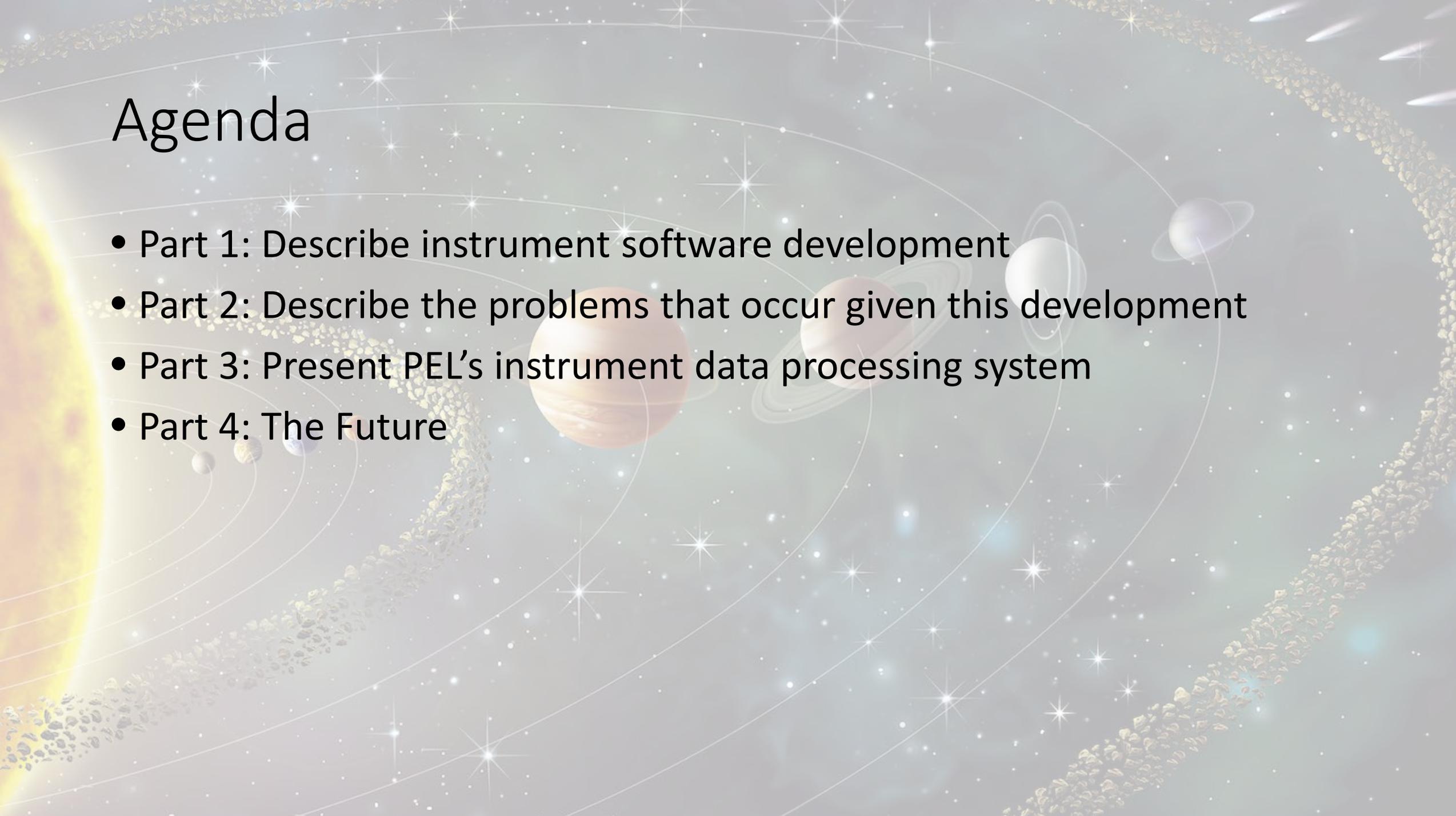
Code 699

About Me



- Software Engineer with Planetary Environments Lab at GSFC
- Develop software for Planetary Science Instruments
- Responsible for all aspects of ground software for Mars Organic Molecule Analyzer (MOMA)
- Previously Developed ground software for Sample Analysis at Mars, Maven Neutral Gas and Ion Mass Spectrometer and the LADEE Neutral Mass Spectrometer
 - Wrote the software and delivered the LADEE-NMS data to PDS in PDS4 format
 - I am a huge fan of PDS
- My customers are the scientists of 699. A big part of my job is making sure they have all the data and the best software we can make

Agenda



- Part 1: Describe instrument software development
- Part 2: Describe the problems that occur given this development
- Part 3: Present PEL's instrument data processing system
- Part 4: The Future

Part 1: Instrument Software Development



Instrument Software: Scope

The complete set of software required to support a complex science instrument contains many layers, written in many languages, developed over several years. I will present 7 layers.



Instrument Software: Flight Software

At the lowest level, the flight software defines the communications protocol and unique data formats for the mission.

Flight Software

```
prev_fil_b_state = None
curr_fil_b_state = None
fil_b_cycle_count = 0

seb_is_on = None
fil_emon_value = None

for pkt in tmfile:

    # Update our record of the SEB state:

    if (pkt.get_type() == tmread.MomaPktIds.digital_hk_pkt_id and
        not ignore_digital_status_pkt(pkt)
    ):
        seb_is_on = pkt.get_first_value(SEB_IS_ON_HKID)
        if seb_is_on is not None and not seb_is_on:
            curr_fil_a_state = "off"
            curr_fil_b_state = "off"

    if pkt.get_type() == tmread.MomaPktIds.seb_hk_pkt_id:
        fil_a_bit = pkt.get_first_value(IS_FIL_A_ON_HKID)
        fil_b_bit = pkt.get_first_value(IS_FIL_B_ON_HKID)
        try:
            EMON_HKID = FIL_A_EMON_HKID if fil_a_bit else FIL_B_EMON_HKID
            fil_emon_value = list(pkt.all_eng_data(EMON_HKID))[0].value
        except IndexError: # No value, channel is not being sampled
            fil_emon_value = None
        if fil_emon_value is None:
            curr_fil_a_state = "on" if fil_a_bit else "off"
            curr_fil_b_state = "on" if fil_b_bit else "off"
        elif fil_emon_value > FIL_EMON_THRESHOLD:
            if fil_a_bit:
                curr_fil_a_state = "on"
            if fil_b_bit:
                curr_fil_b_state = "on"
        else:
            curr_fil_a_state = "off"
            curr_fil_b_state = "off"

    # Compare to previous state and count any new cycles:

    if prev_fil_a_state is not None:
        if prev_fil_a_state == "off" and curr_fil_a_state == "on":
            fil_a_cycle_count += 1

    if prev_fil_b_state is not None:
        if prev_fil_b_state == "off" and curr_fil_b_state == "on":
            fil_b_cycle_count += 1

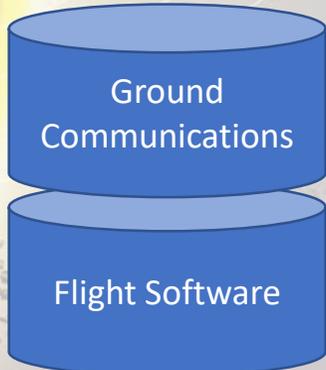
    # Store current state for next iteration:

    prev_fil_a_state = curr_fil_a_state
    prev_fil_b_state = curr_fil_b_state

return fil_a_cycle_count, fil_b_cycle_count
```

Instrument Software: Ground Support

Ground software communicates with flight software and provides commanding and telemetry monitoring.

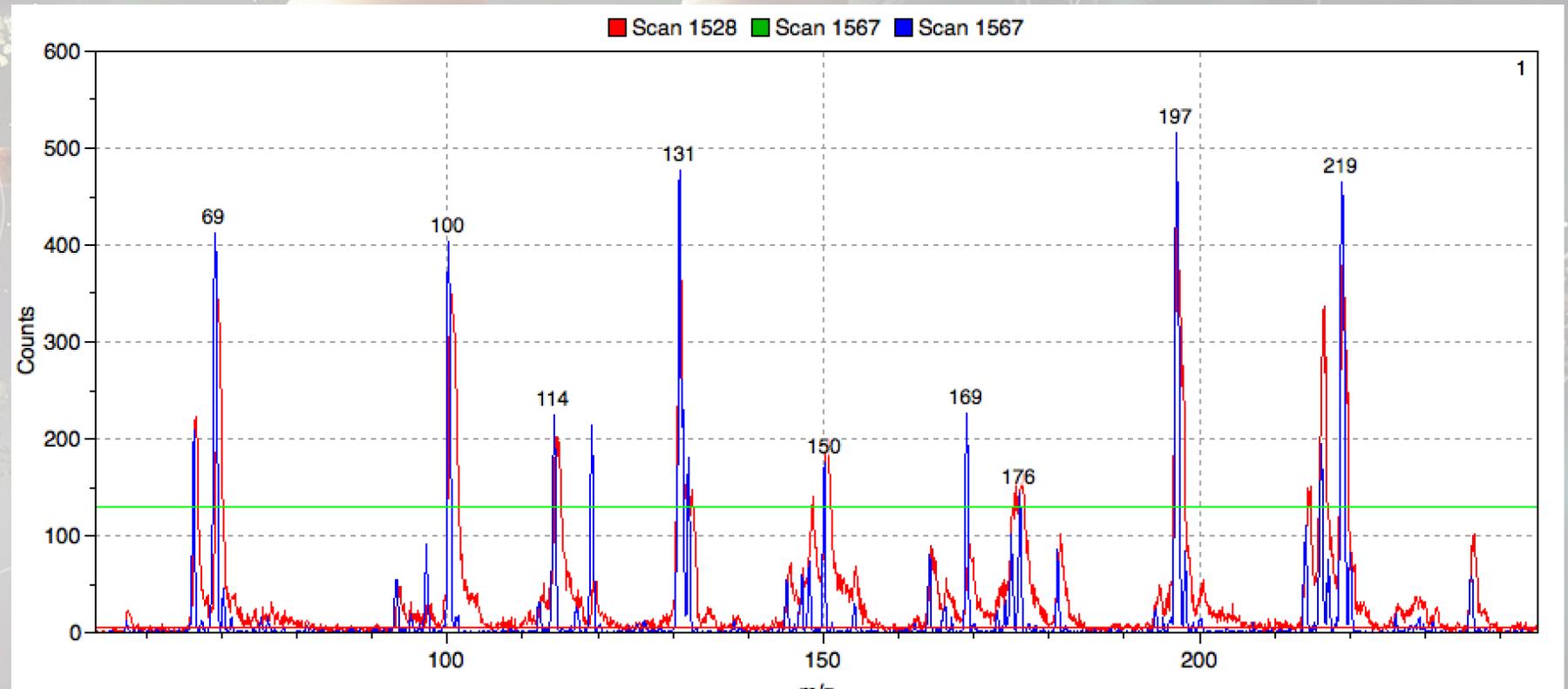
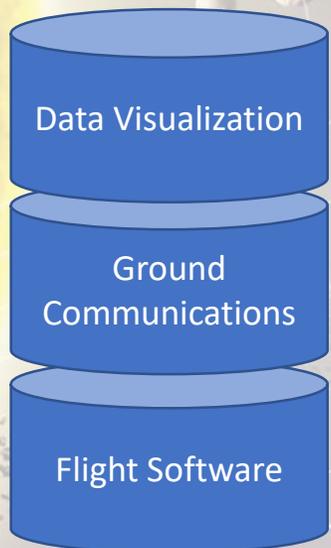


The screenshot displays a complex ground support software interface with several windows:

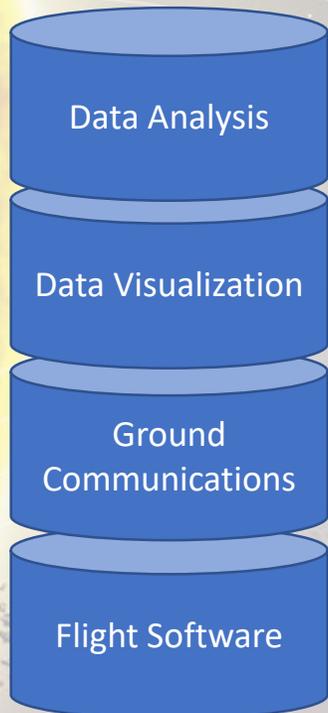
- NTSSE Event/Command Window:** Shows a log of system events and commands, including file operations, voltage checks, and housekeeping packets.
- BC Activity:** A graph showing current (A) over time (Seconds).
- System:** Displays system management information such as disk space (Total: 304451.00 GB, Used: 112838.00 GB, Free: 192413.00 GB) and network status (DISCONNECTED).
- radback:** A large window showing a grid of status indicators for various subsystems (SA-1 to SA-18) with columns for Rollout, NoResp, and Busy.
- Command:** Shows bus status (RUNNING), channel (CHAN_A), and retry status (ENABLE).
- TELEMETRY HMI (SAT-2):** Displays message details, RPM counter (64405), and 1553 status (Num Mess Recv = 0, Num Cmd Exec = 251).
- Temperature Monitor Points:** Two windows showing sensor data for DPU and APU boards, including board sensor values and voltage/current readings.
- RFM:** Shows RFM status (15.000) and time (4.800).
- RFM DIGITAL TELEMETRY:** Displays digital telemetry data for DIGTLM 1-4, including LUT, status of WD, and timing information.

Instrument Software: Visualization

Software to visualize science and engineering data in real-time



Instrument Software: Data Analysis



Model Flight Model

Chart Mode Marker-Relative Time

Chart Data Ion Count

Calibration Mode Mass

Ion Selection

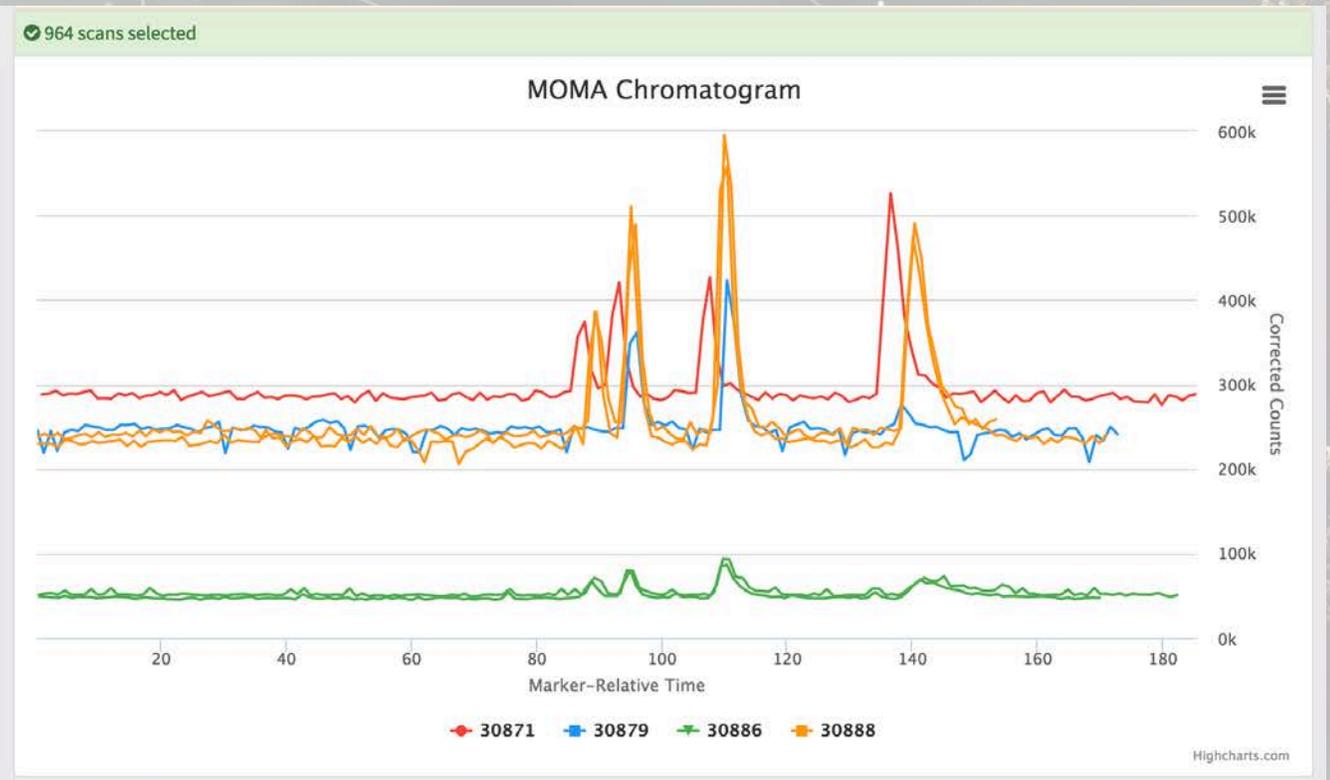
- Total
- [43.5, 44.5)
- [54.5, 55.5)
- [68.5, 69.5)

Normalize Correct Counts

Scans (0 records, 1 expressions)

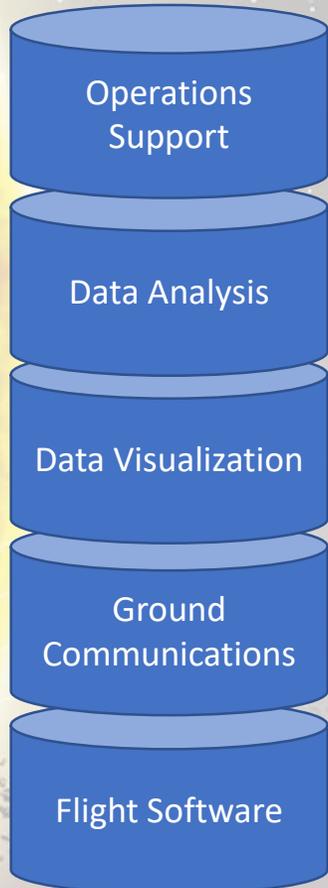
1 expression(s) selected

``TID` >= "30850" and `MK Name` starts with "GC 7:"`

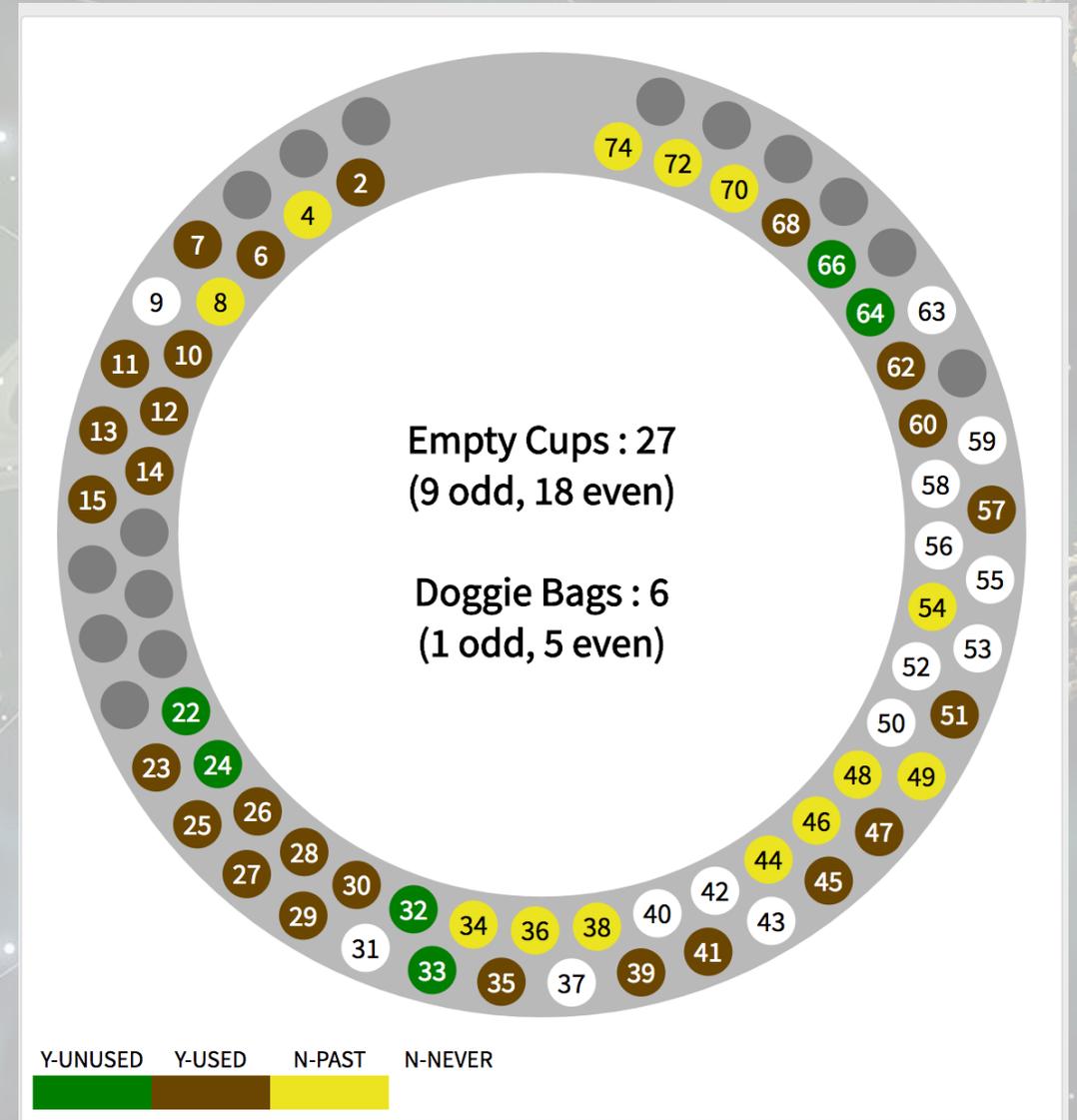


Post-Processing software to trend data and make instrument-specific analyses

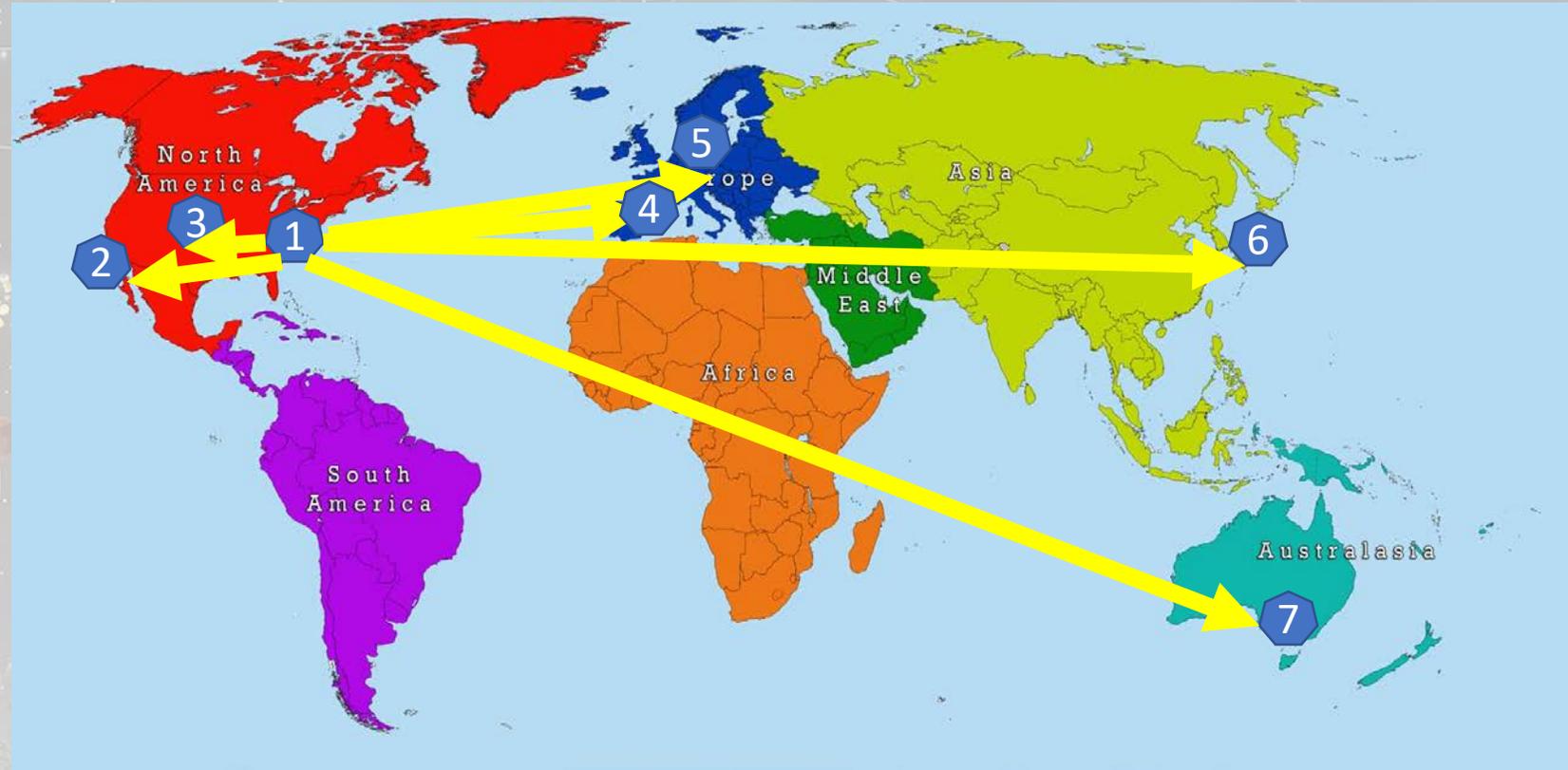
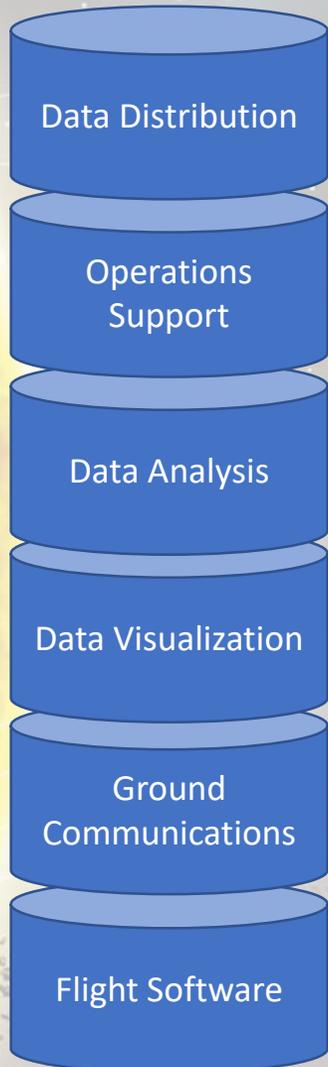
Instrument Software: Operations Support



Instrument operations require software to process data from the spacecraft and to support the operations team.

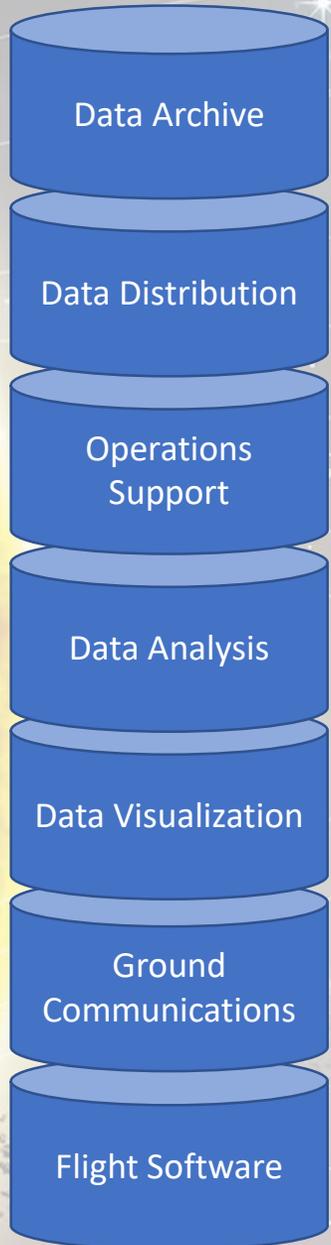


Instrument Software: Distribution

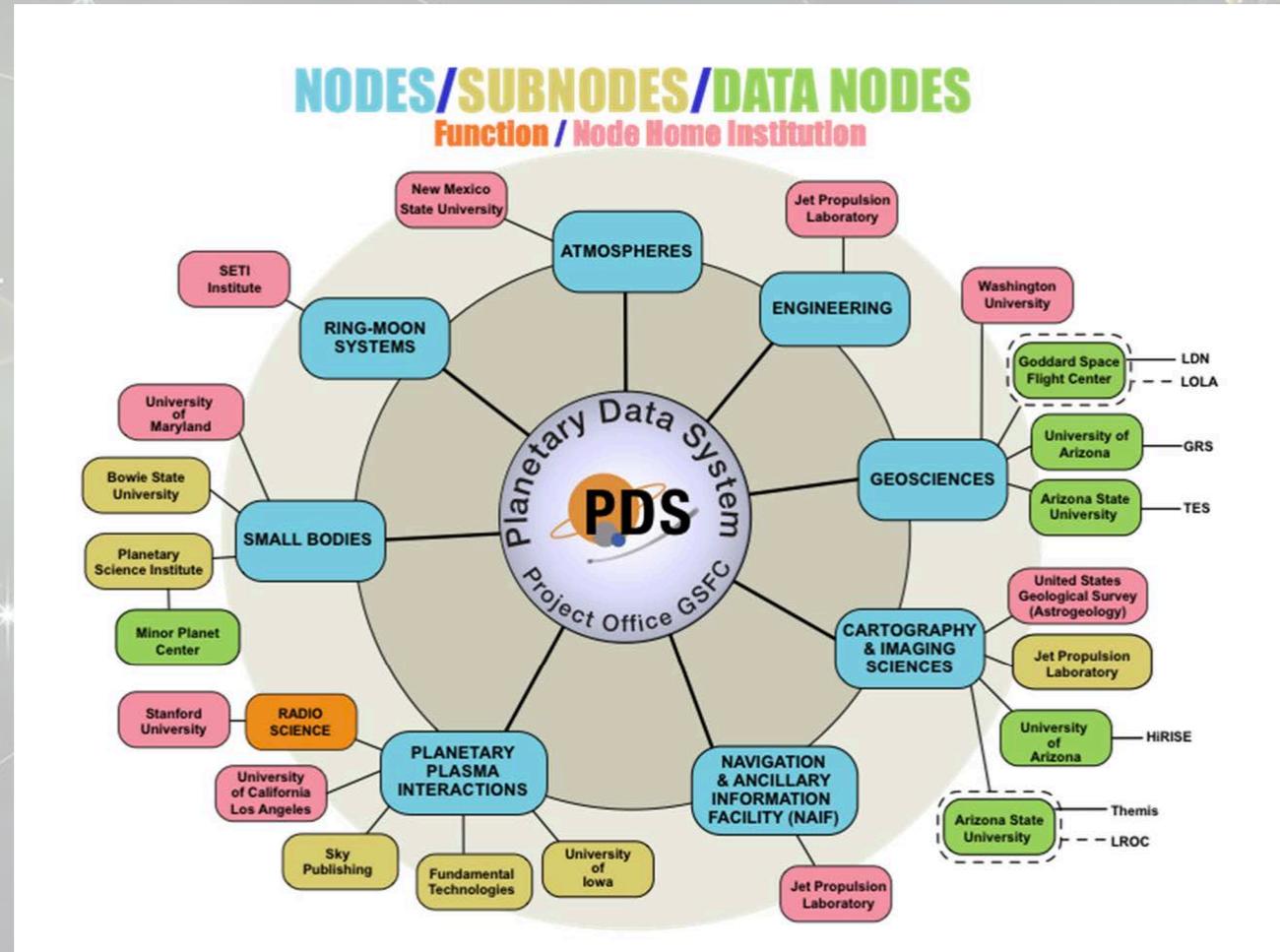


Operations requires distribution to a wide spread science team.

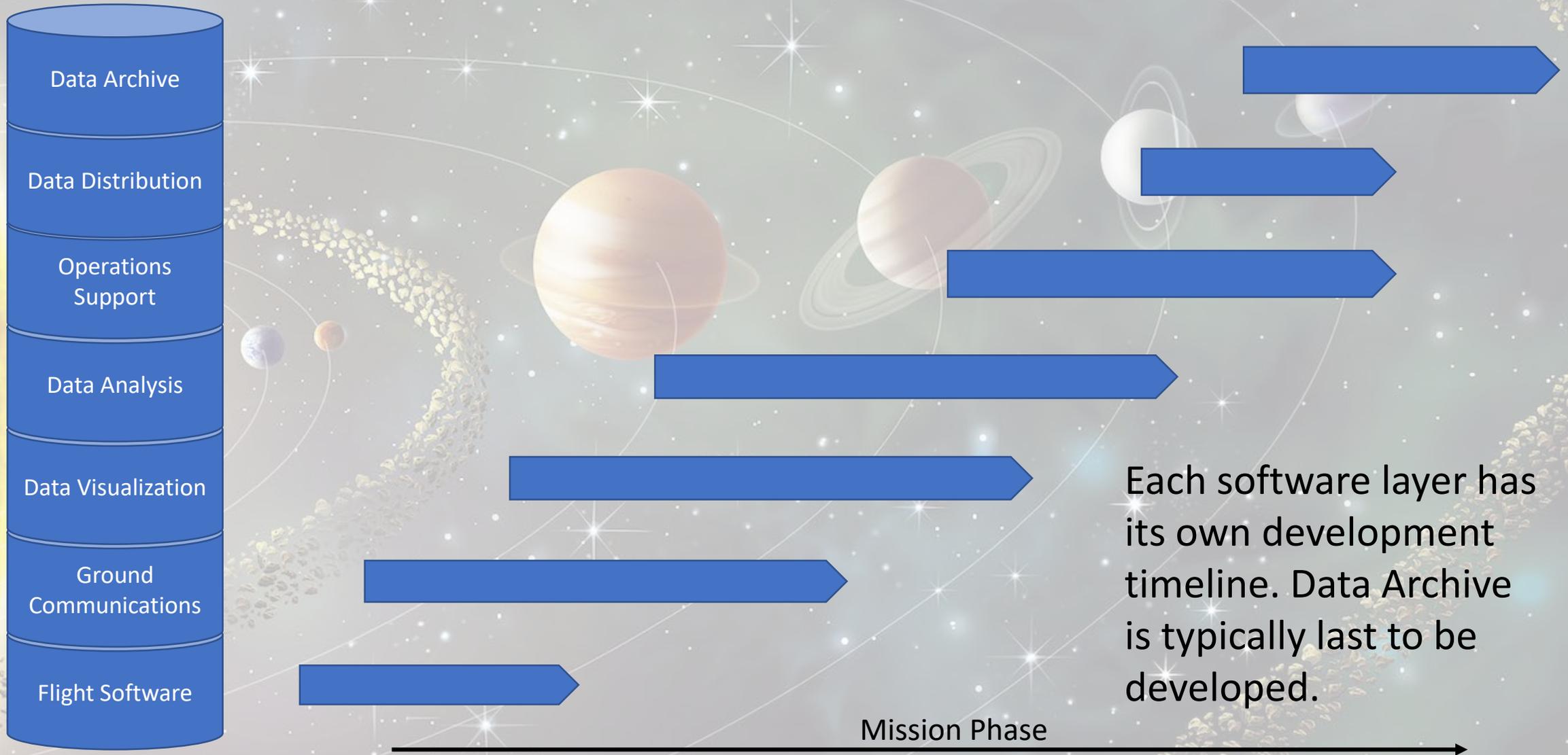
Instrument Software: Data Archive



The tip of the iceberg is software to generate the Planetary Data System archive

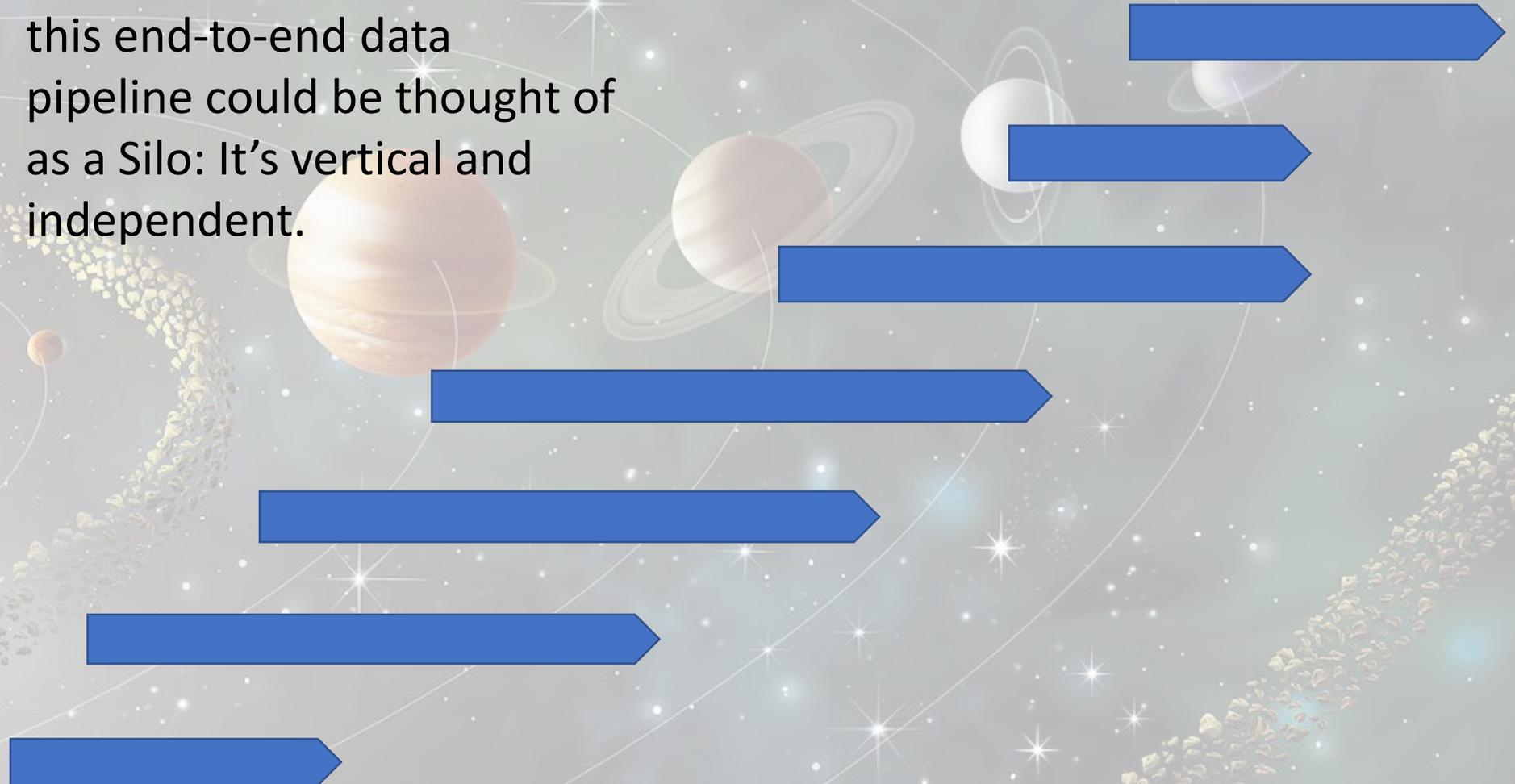
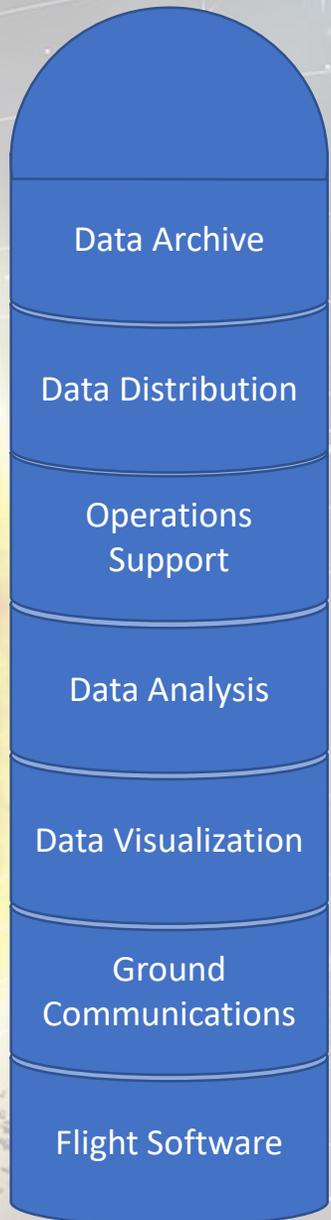


Instrument Software Development



Instrument Software Development

The set of tools to support this end-to-end data pipeline could be thought of as a Silo: It's vertical and independent.

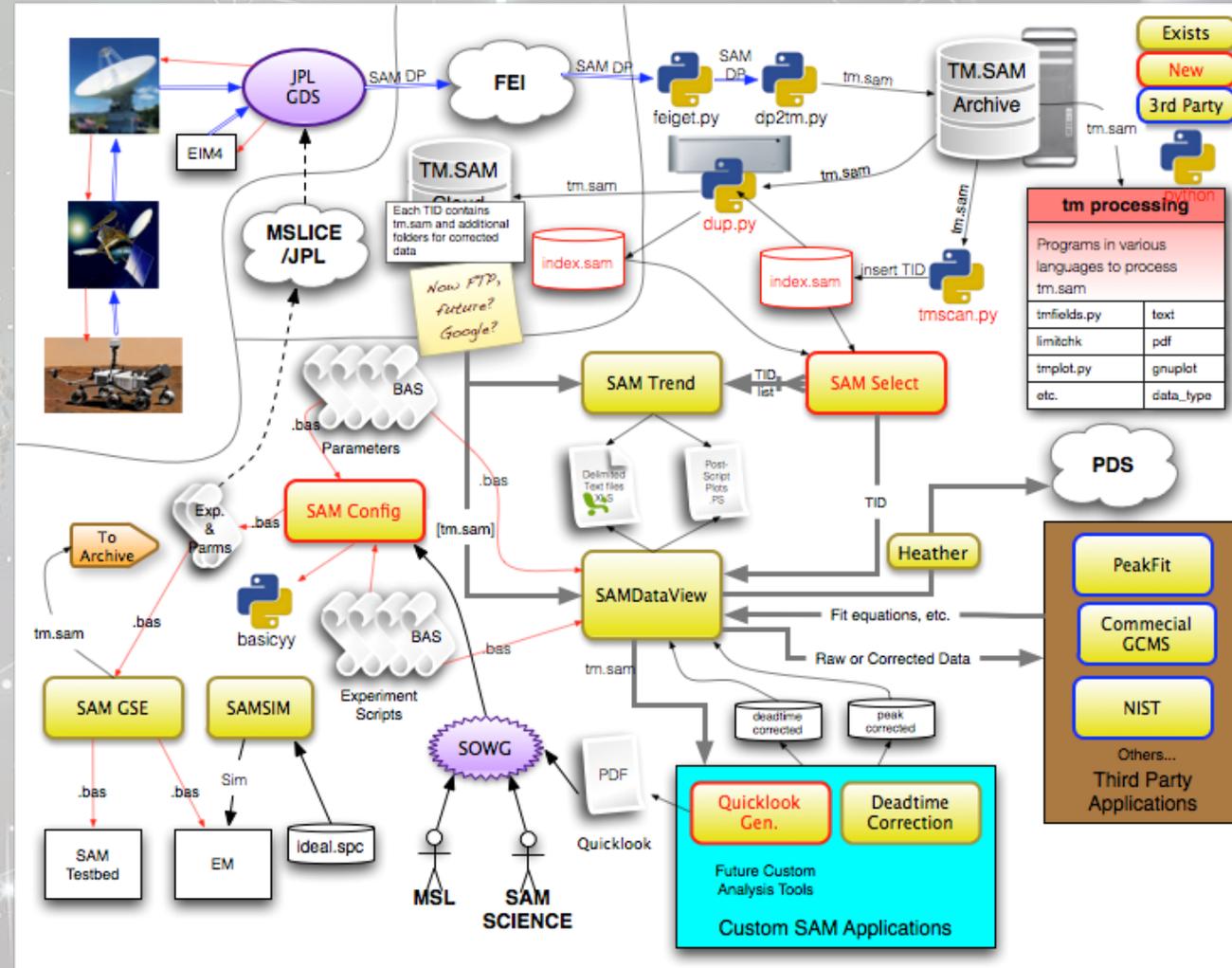


Mission Phase

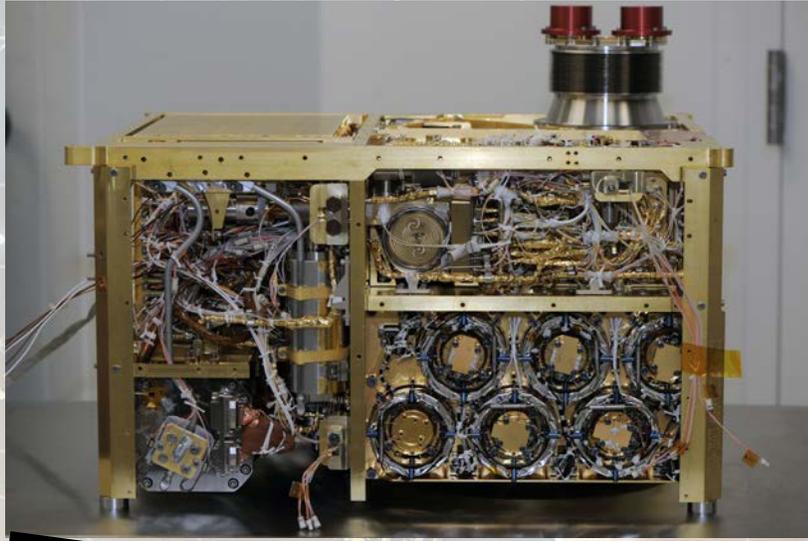


Instrument Software Silo Example

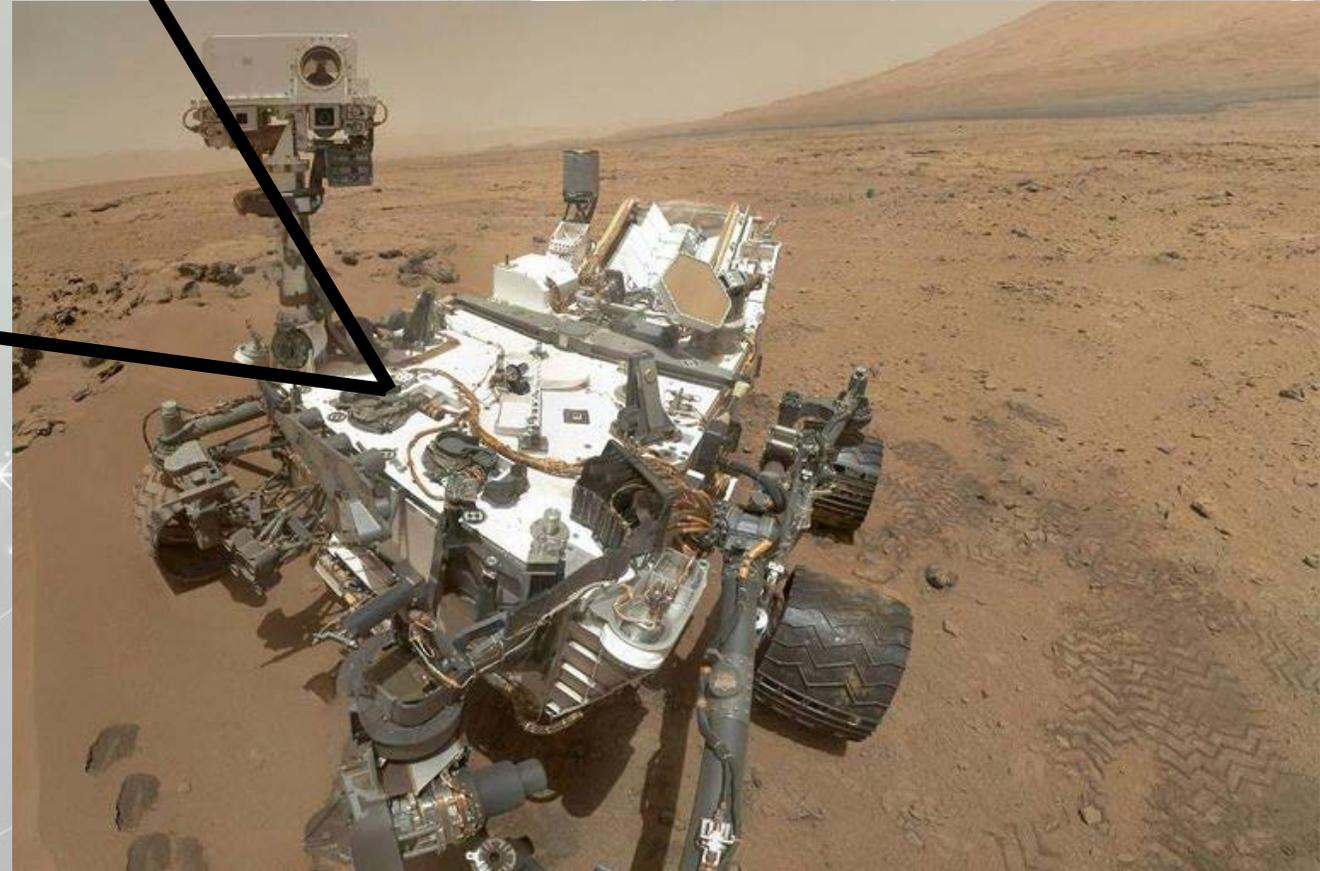
For the Sample Analysis at Mars (SAM) instrument on Curiosity, an entire software “silo” was created.



MSL

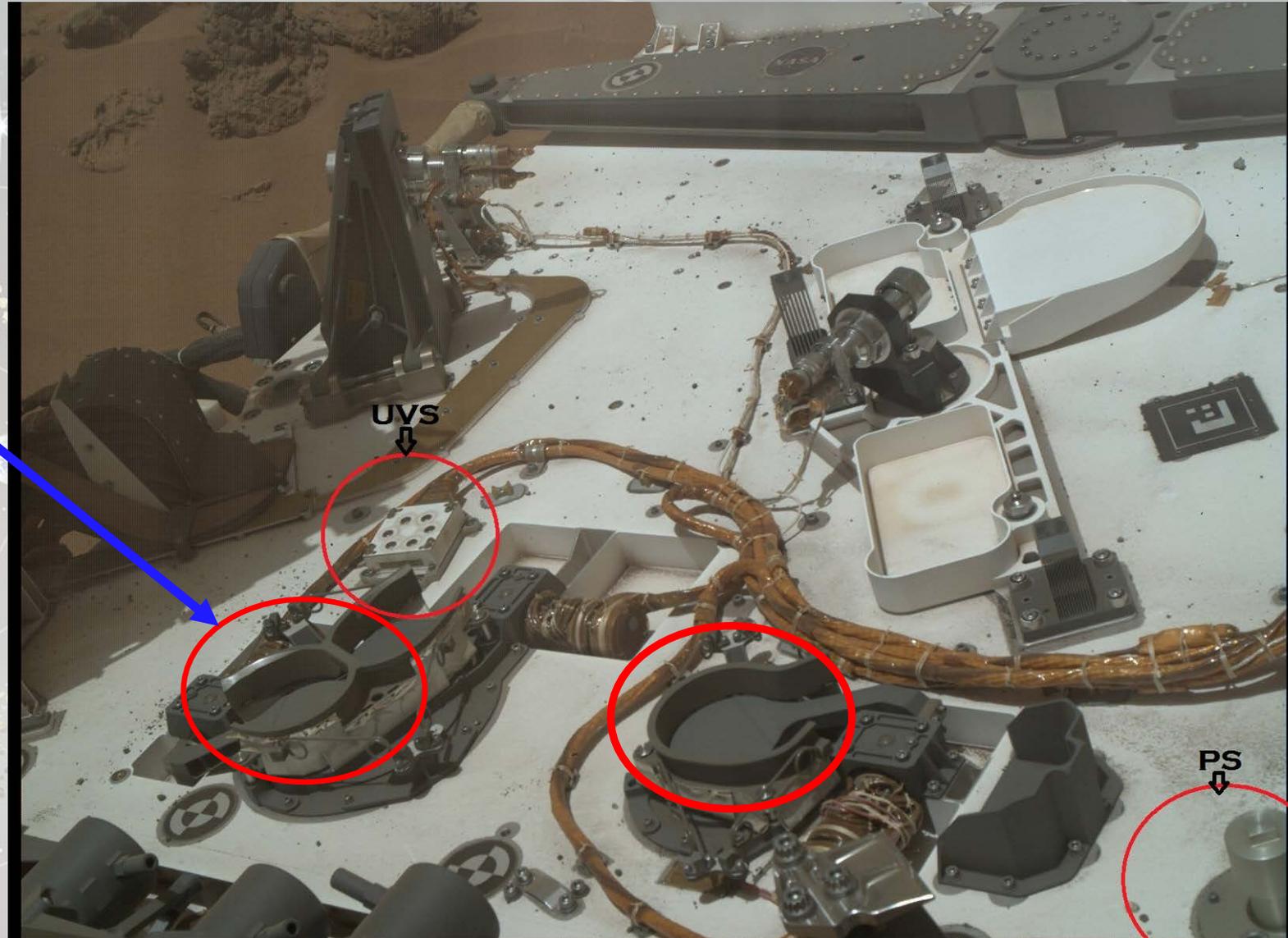


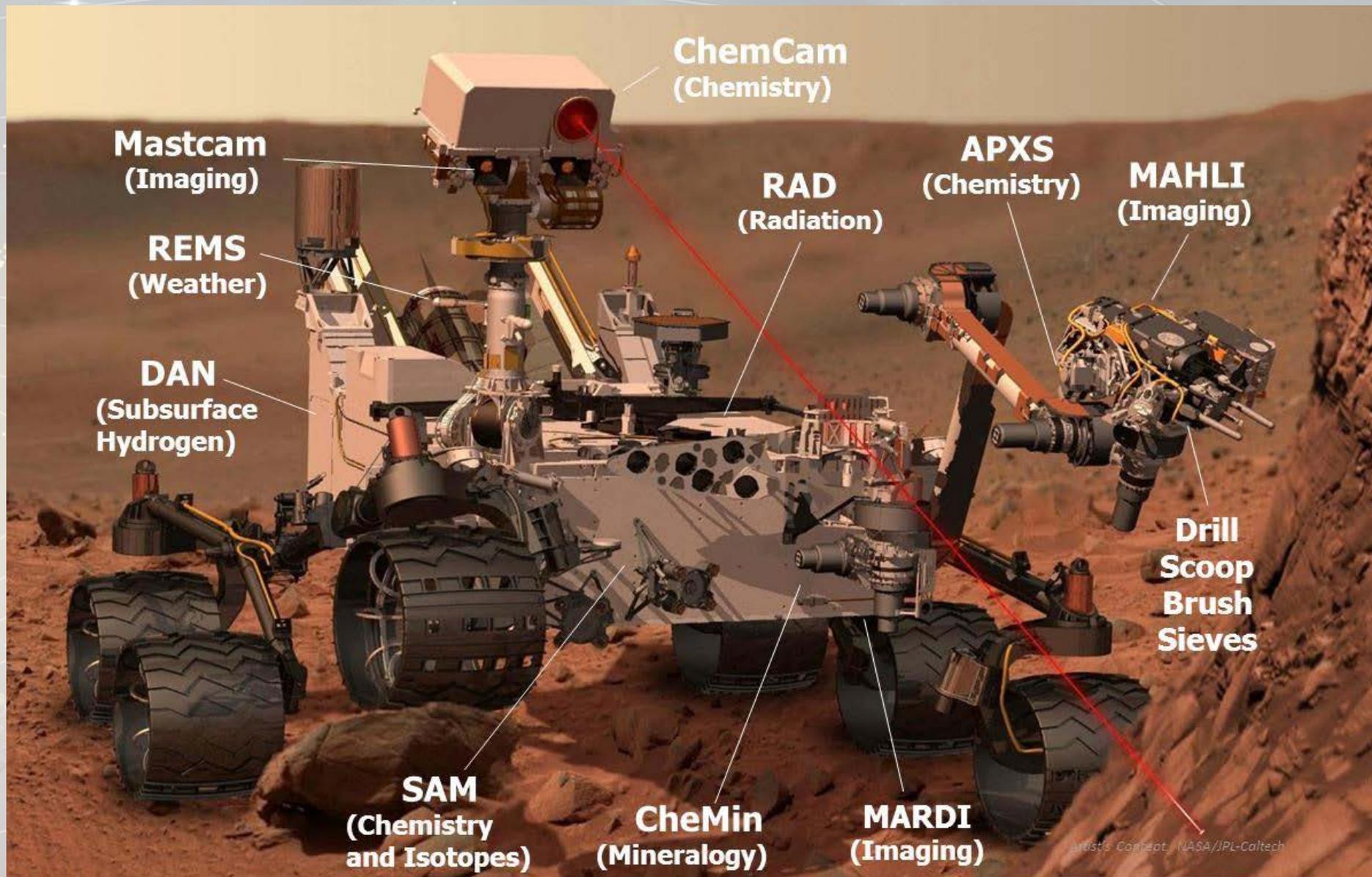
SAM sits in the belly of the rover and receives solid samples dropped by the arm into an inlet.

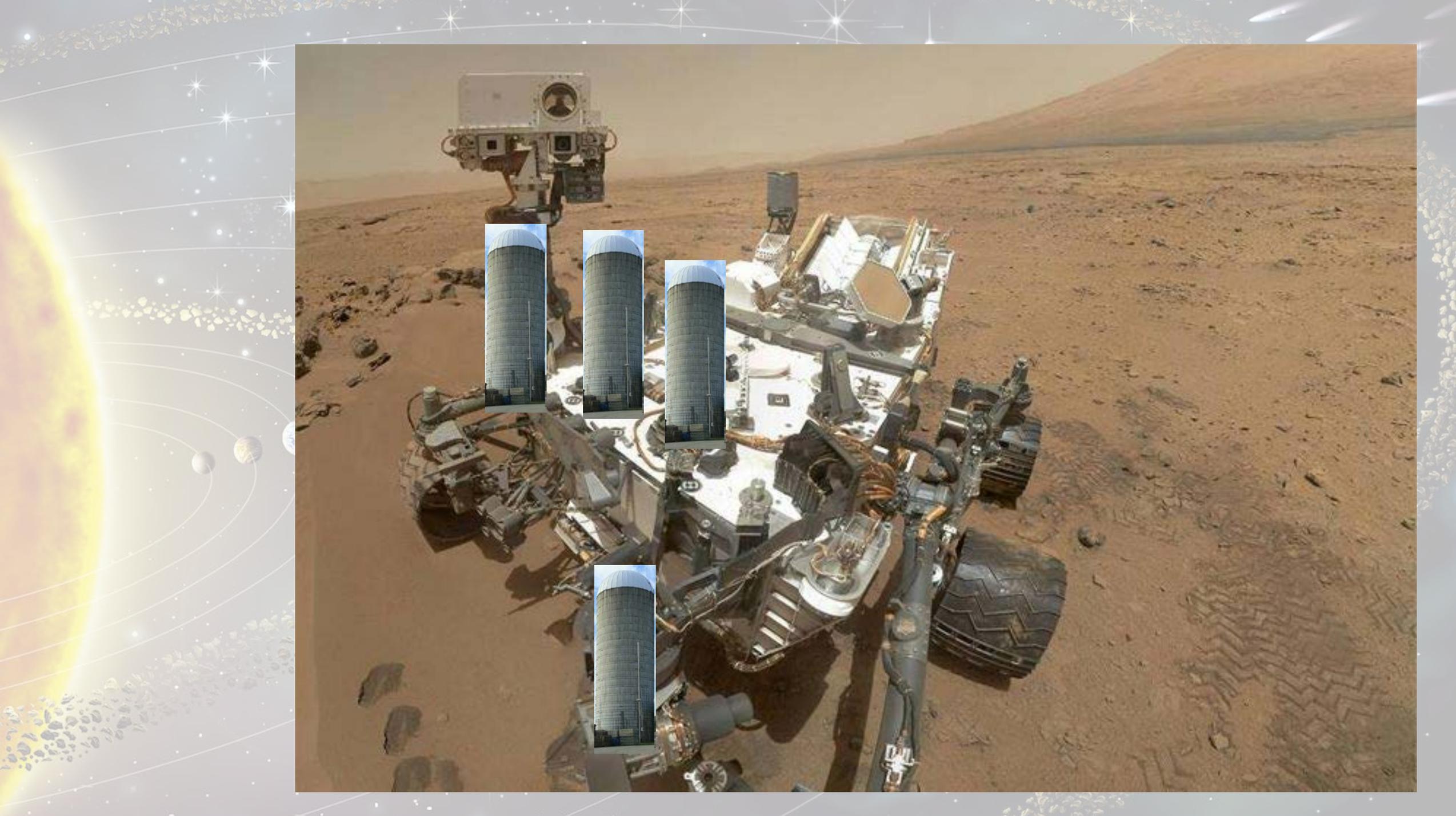


Curiosity's Deck

SAM's inlets







Part 2: The Trouble with Silos

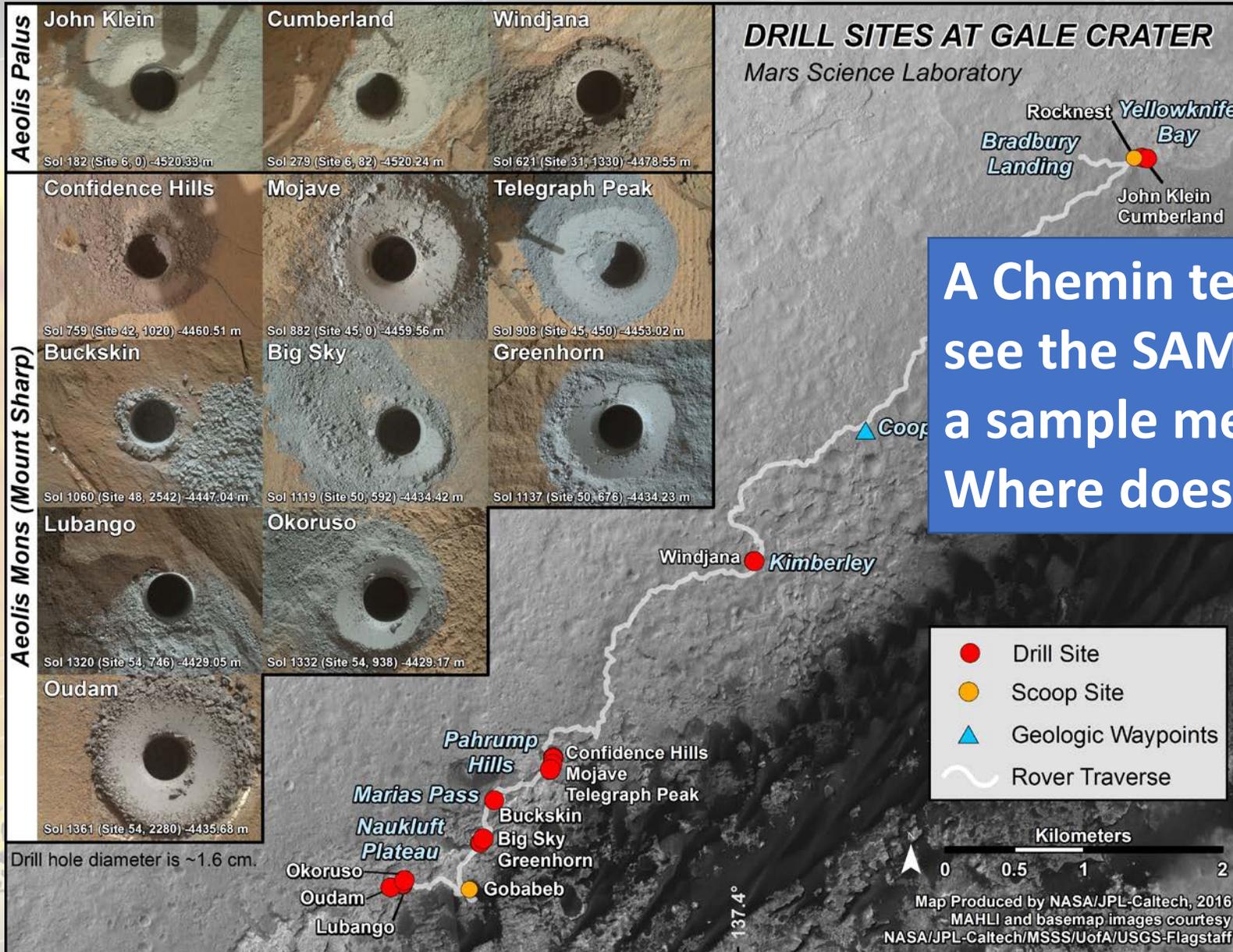
- Software silos yield data silos.
 - Scientific collaboration is difficult between instruments.
 - Lost meta-data because each team stores separately
- Instrument teams often do not use their own PDS products.
 - PDS software development is typically done at the end, after other tools are complete.
- Instrument teams redundantly develop pipelines.
 - Time consuming and costly process repeated by each team.
 - They ARE doing this.

Part 2: The Trouble with Silos



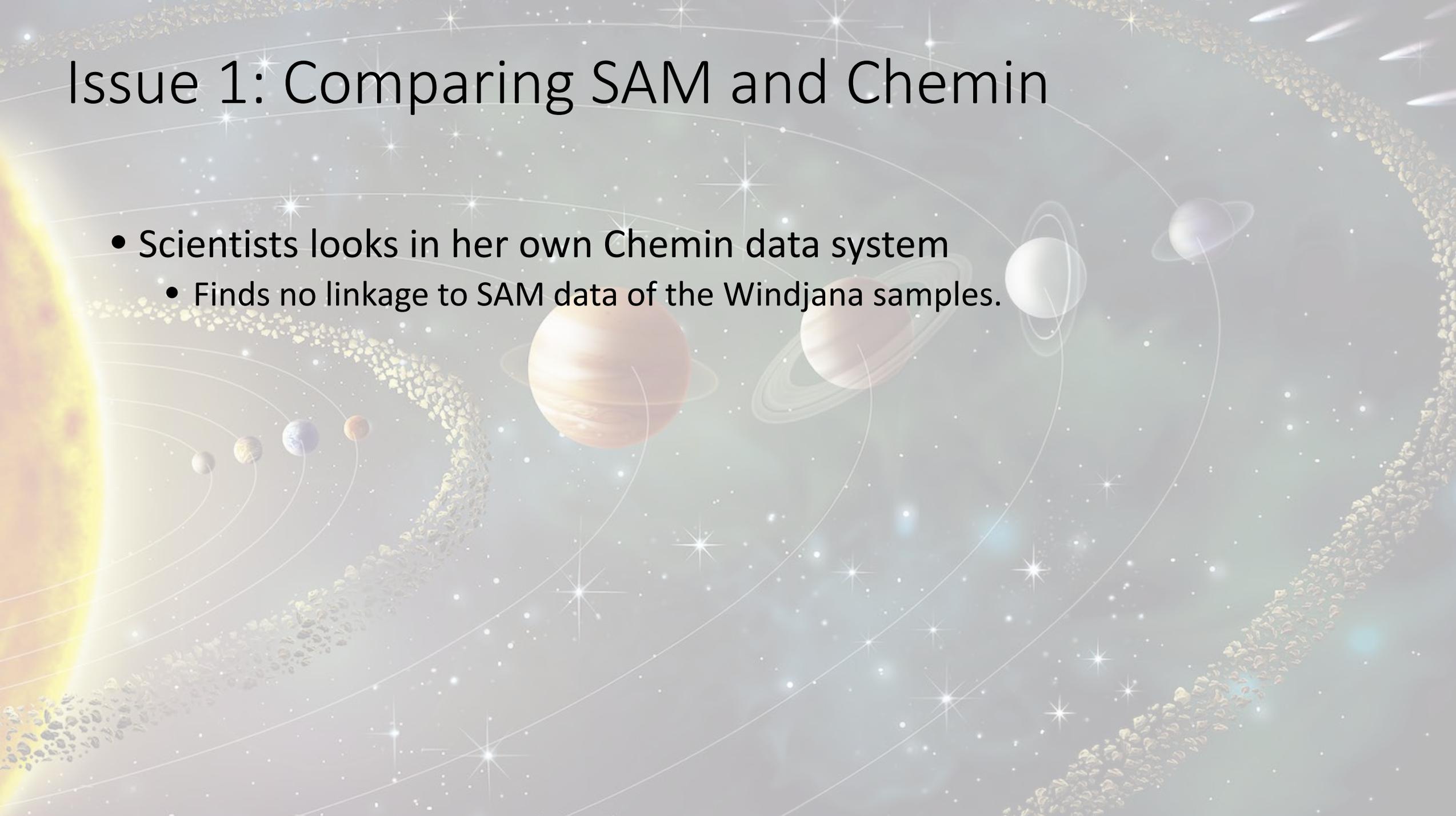
1. SAM and Chemin: A scientist's investigation
2. Redundant Development
3. PDS Products and the Data Pipeline

Issue 1: Comparing SAM and Chemin



A Chemin team scientist wants to see the SAM mass spectral analysis of a sample measured at Windjana. Where does she start?

Issue 1: Comparing SAM and Chemin



- Scientists look in her own Chemin data system
 - Finds no linkage to SAM data of the Windjana samples.

Issue 1: Comparing SAM and Chemin



APSX



SAM



Chemin



MAHLI

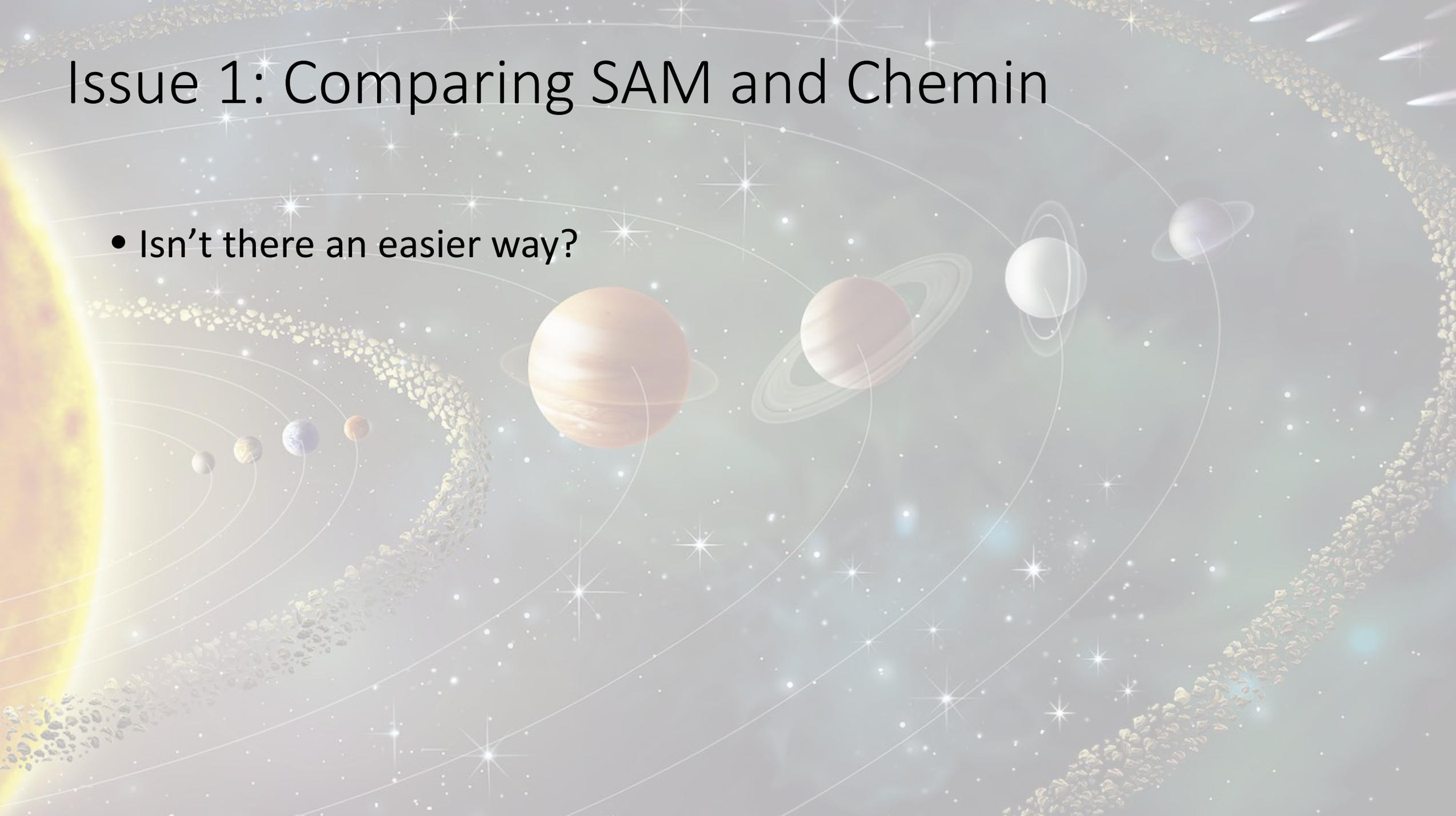
Scientist travels to the SAM silo...

Issue 1: Comparing SAM and Chemin

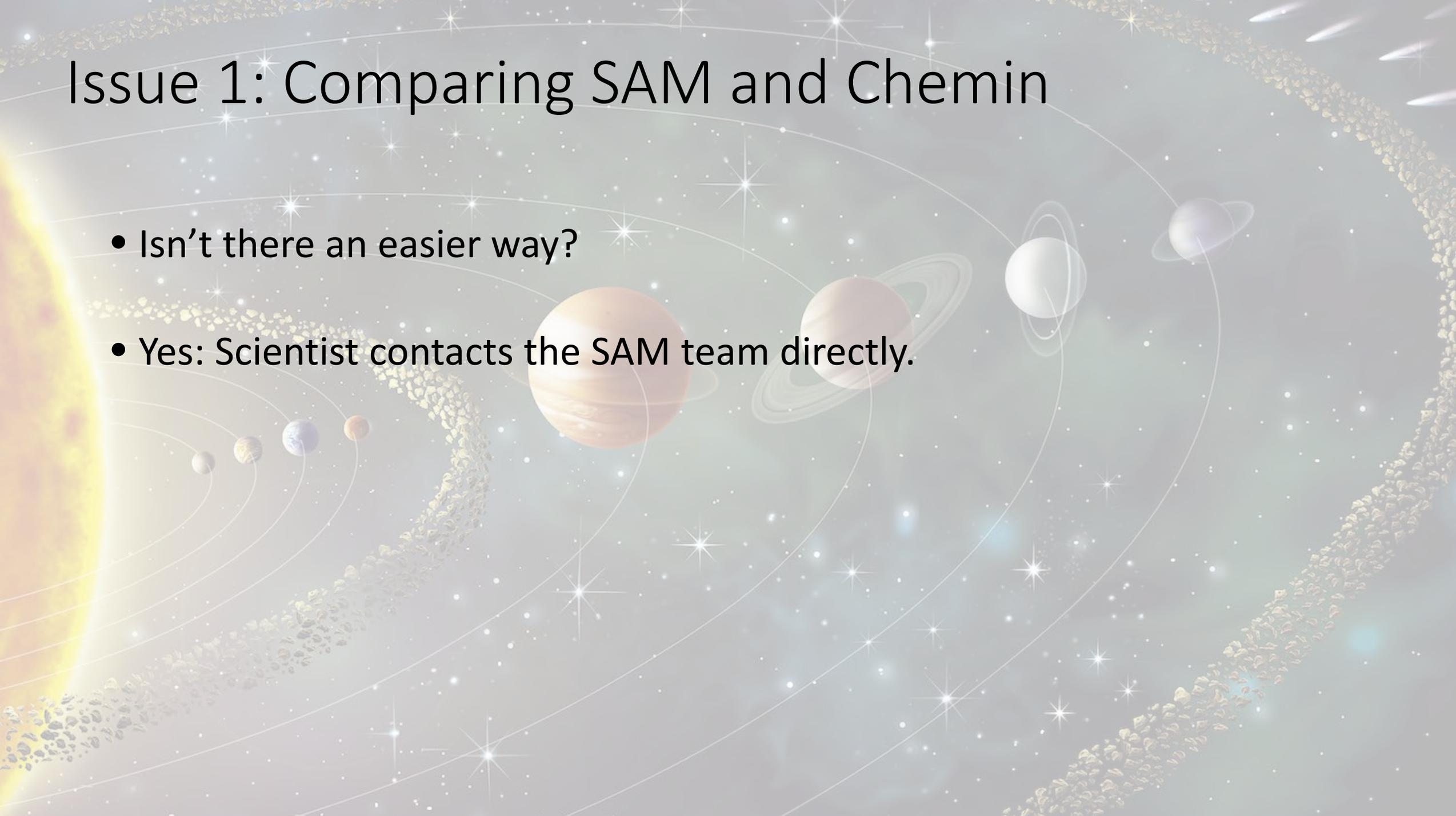
- What does she find in SAM PDS?
 - SAM Data is linked to the Sol of analysis, not sol of collection.
 - SAM often analyzes samples long after the sol it is drilled.
 - Samples are sometimes analyzed several times.
 - Samples may be cached and analyzed many months later.
 - This data is stored as comments in text files.

Issue 1: Comparing SAM and Chemin

- Isn't there an easier way?

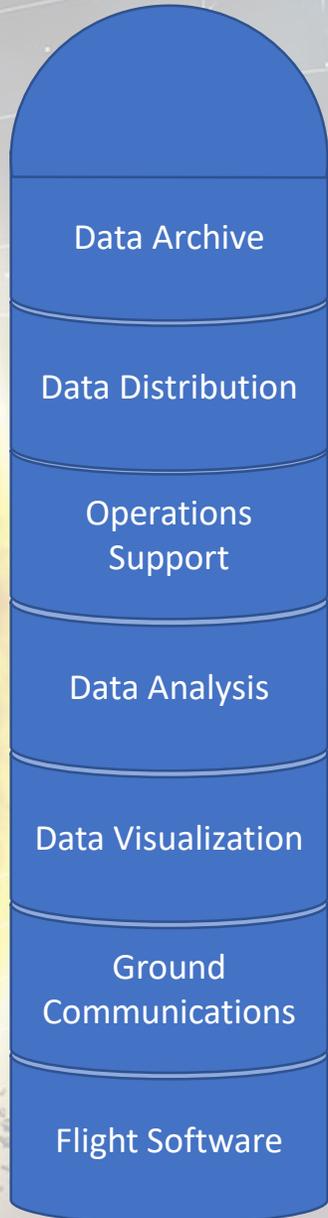


Issue 1: Comparing SAM and Chemin



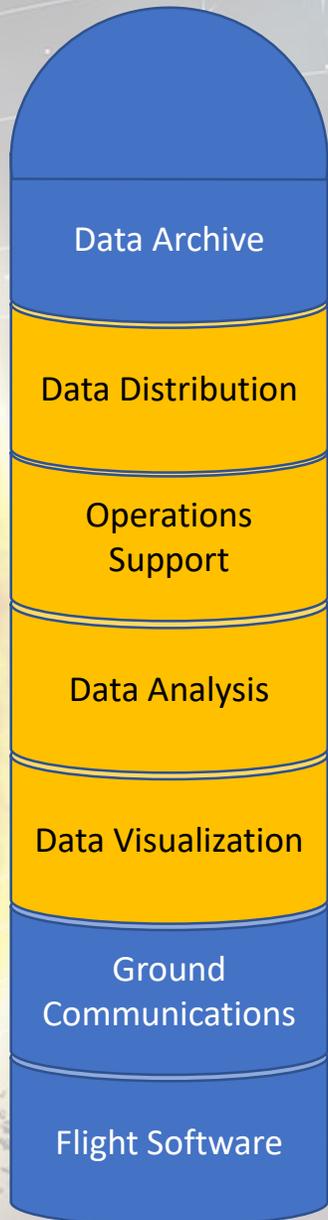
- Isn't there an easier way?
- Yes: Scientist contacts the SAM team directly.

Issue 2: Redundant Development



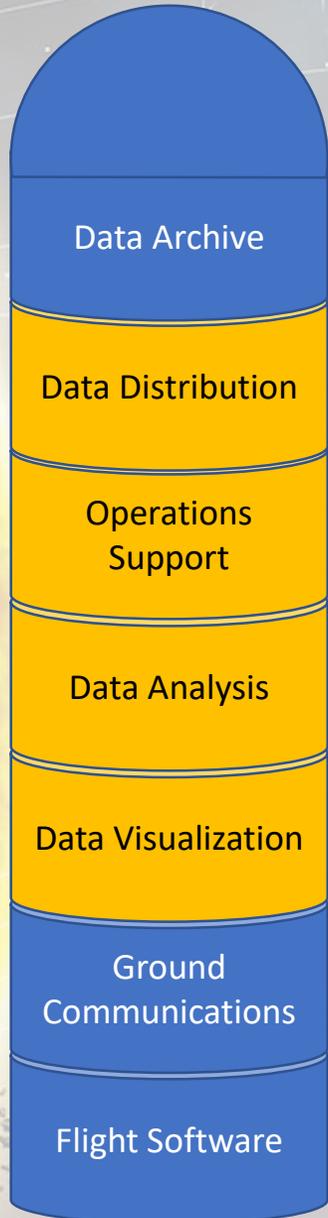
- Some of the Data Pipeline for each instrument is necessarily unique.
 - Science data formats
 - Communication protocols
 - Packet formats

Issue 2: Redundant Development



- Some of the Data Pipeline for each instrument is necessarily unique.
 - Science data formats
 - Communication protocols
 - Packet formats
- Many parts could be abstracted.
 - Housekeeping data analysis
 - Data Distribution
 - Visualizations

Issue 2: Redundant Development



- Some of the Data Pipeline for each instrument is necessarily unique.
 - Science data formats
 - Communication protocols
 - Packet formats
- Many parts could be abstracted.
 - Housekeeping data analysis
 - Data Distribution
 - Visualizations
- **THE FUNDAMENTAL PROBLEM OF SILOS**
- Why does this happen?

Issue 3: PDS Data Products and the Data Pipeline

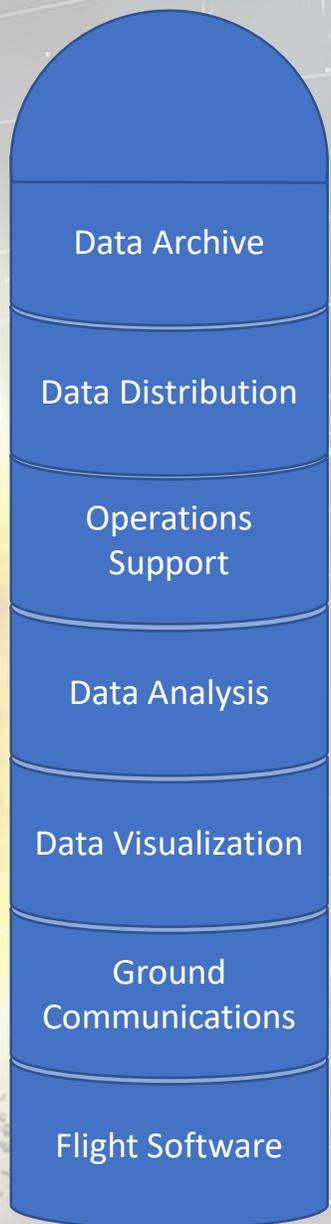


PDS encourages instrument teams to use PDS format data for their own internal analysis software.

Advantages are great:

- Data is constantly checked for accuracy by internal team.
- Necessary new metadata more likely to be identified and incorporated into PDS
- Software tools can potentially be distributed to other teams

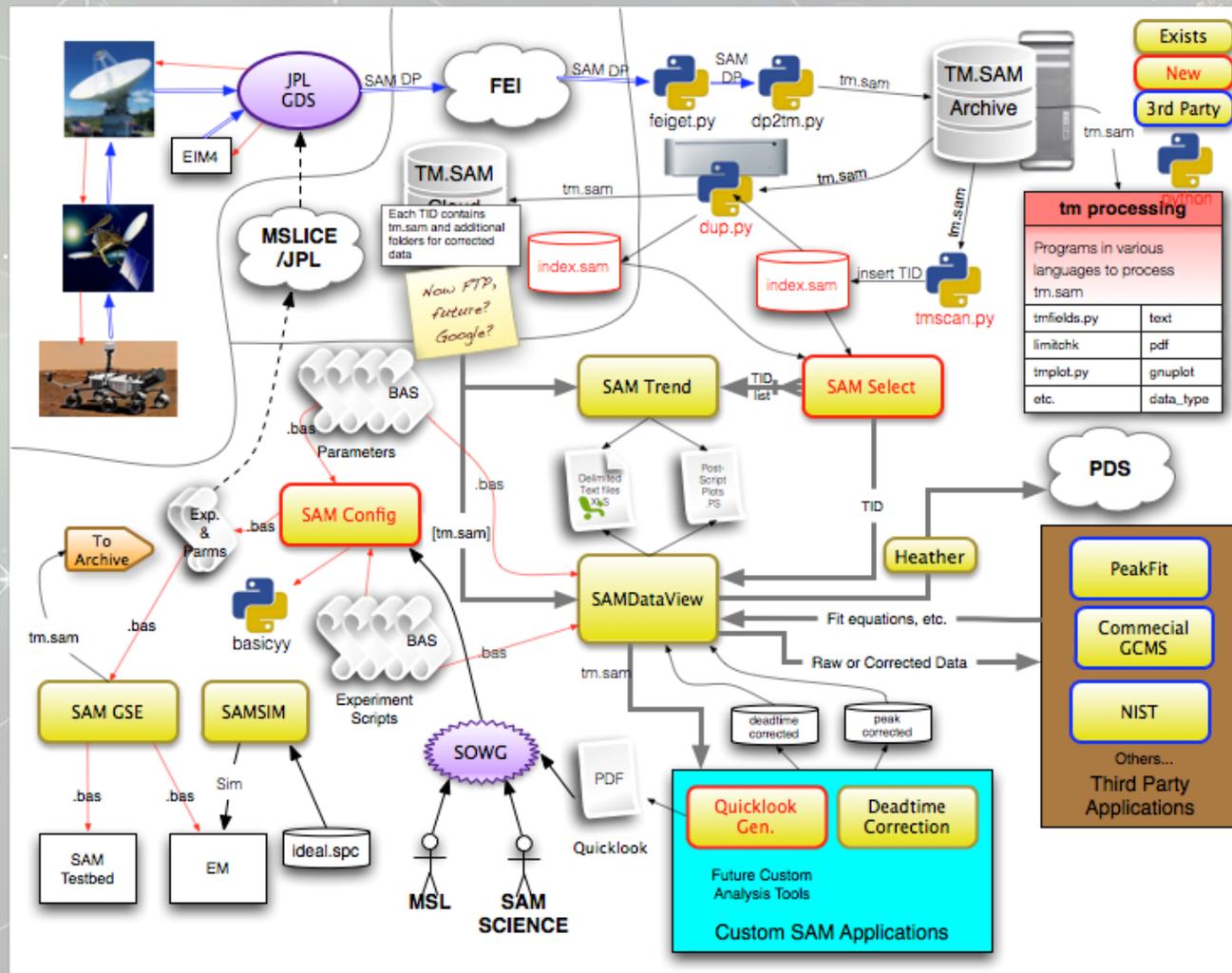
Issue 3: PDS Data Products and the Data Pipeline



The complex SAM data pipeline was *necessarily* developed long before the PDS delivery format was defined. Therefore internal tools use a different format.

During the mission, as operations were optimized, unofficial metadata gradually grew, often not immediately recognized as metadata.

Some metadata has been captured using ad hoc methods but not necessarily migrated to PDS.

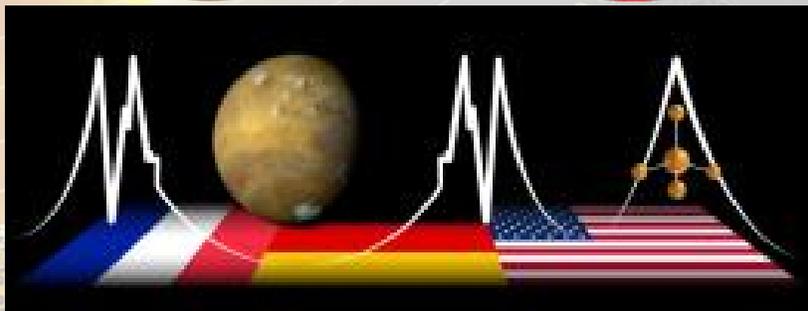


Part 3: The Planetary Environments Lab Solution

Our “Silo”



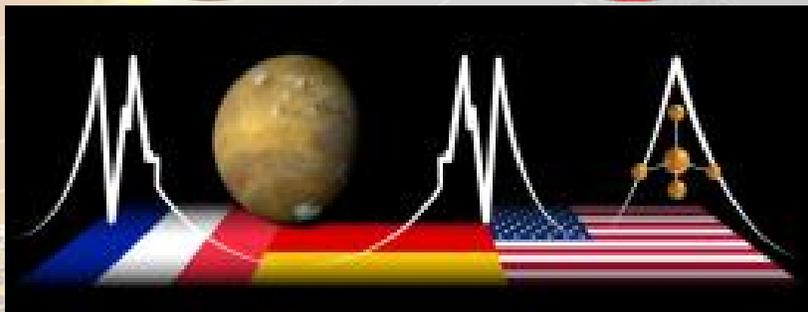
- PEL Software team supported 4 missions since 2006, all based on mass spectrometers.
- Software from one mission handed down and improved on the next mission.
- Gradually developed a skeletal framework from which to start support for new missions
- Complete end-to-end framework has been developed



Part 3: The Planetary Environments Lab Solution Our “Silo”

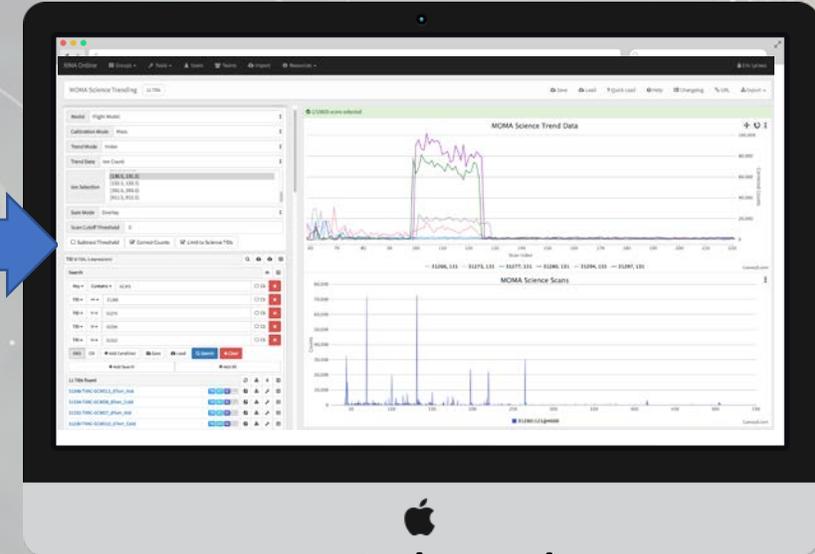


In the PEL, we have developed concepts and architectures about data, and processing data, that could be useful for the Planetary Data System.



Data Processing Line is like an industrial processing line

Instrument Telemetry



- **INPUT:** The processing line takes the telemetry as raw material and assembles it, adding value along the way, into a processed product for delivery to customers
- **OUTPUTS:**
 - PDS bundles can be one of the outputs
 - Web-based interface to the data



Raw Telemetry



De-commutated data



Calibrated data



Archived



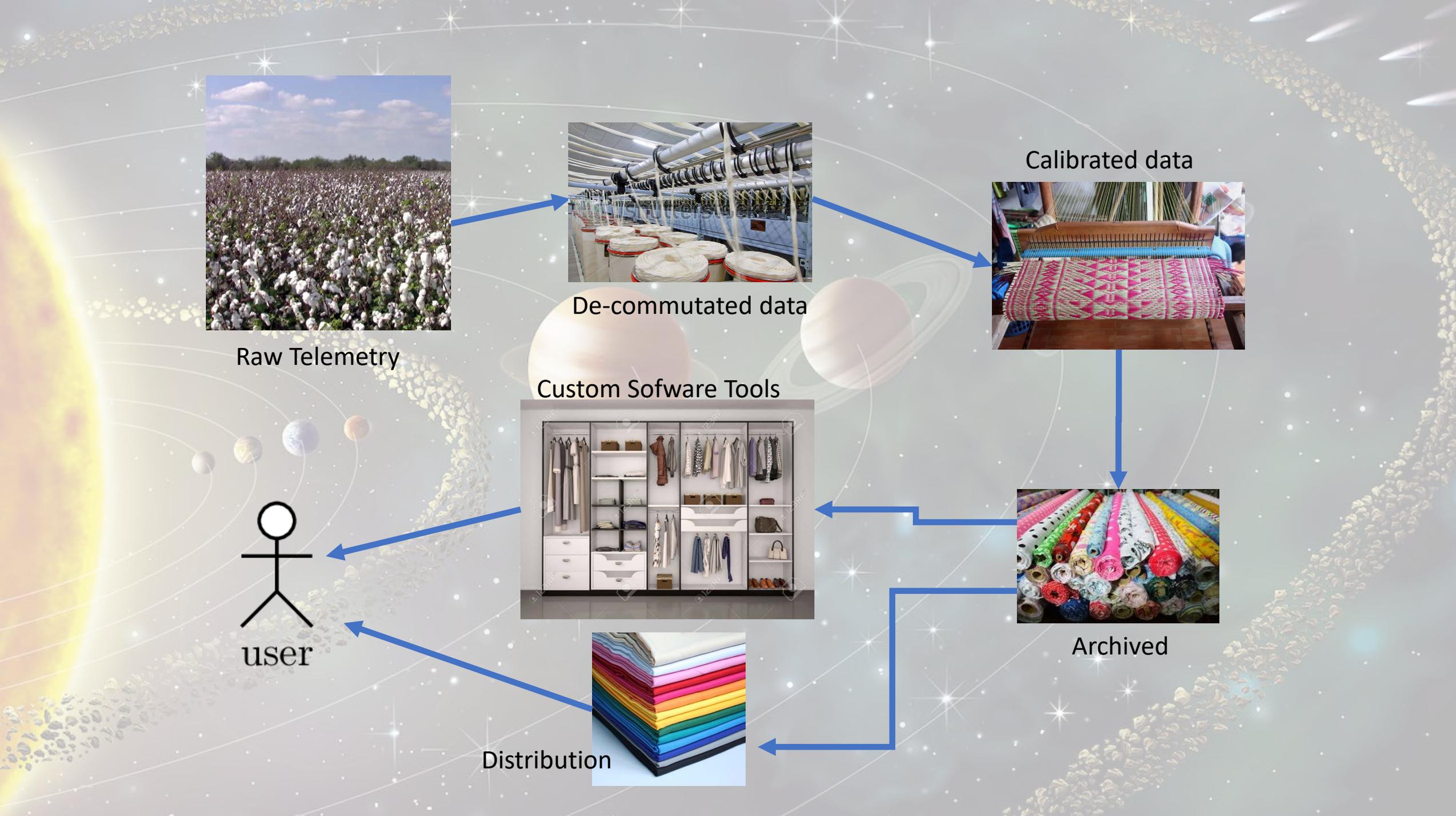
Custom Software Tools



Distribution



user



The Planetary Environments Lab Solution



- **Key Innovation:** We include the hooks in flight software to aid in post processing.
 - Allows the post-processing software to determine the context of the data
 - Crucial for detail trending and performance analysis

The Planetary Environments Lab Solution



Example: A single experiment on Mars contains many mass spectra. Some of the mass spectra are calibration, some are background collection, some are from an actual samples.

Markers allow us to quickly identify where the mass spectra come from.

The Planetary Environments Lab Solution

Key Point: We considering post-processing needs at *the lowest level* and early in development.





Raw Telemetry



De-commutated data



Calibrated data



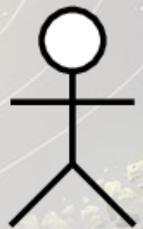
Archived



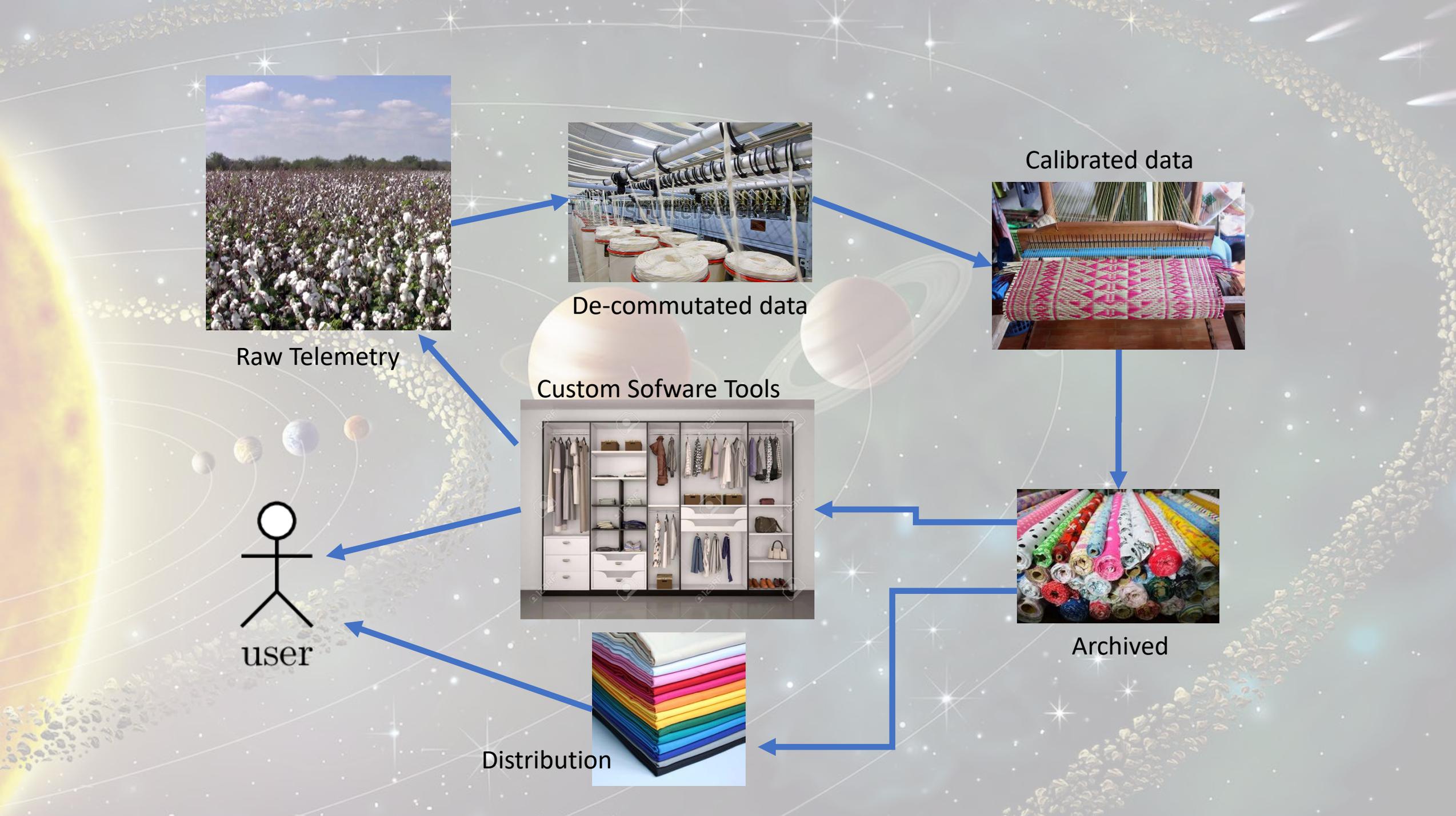
Distribution



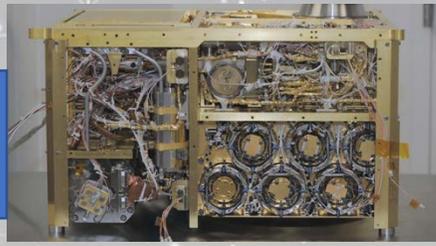
Custom Software Tools



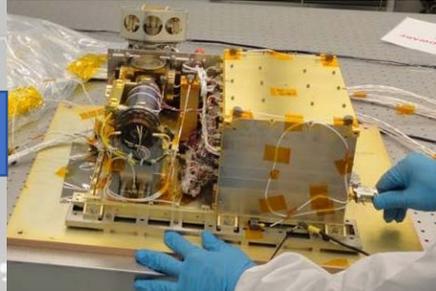
user



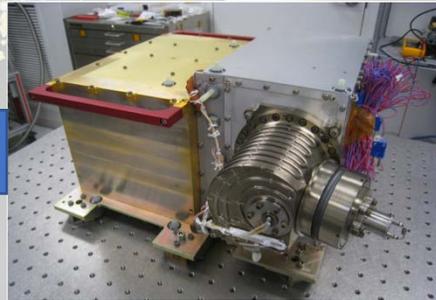
Sample Analysis
at Mars



LADEE NMS



MAVEN NGIMS



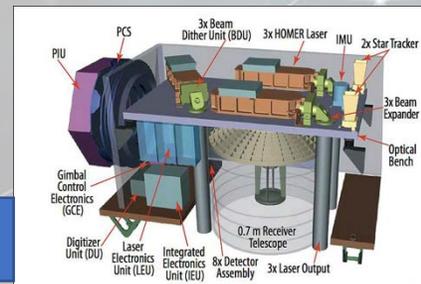
ExoMars MOMA



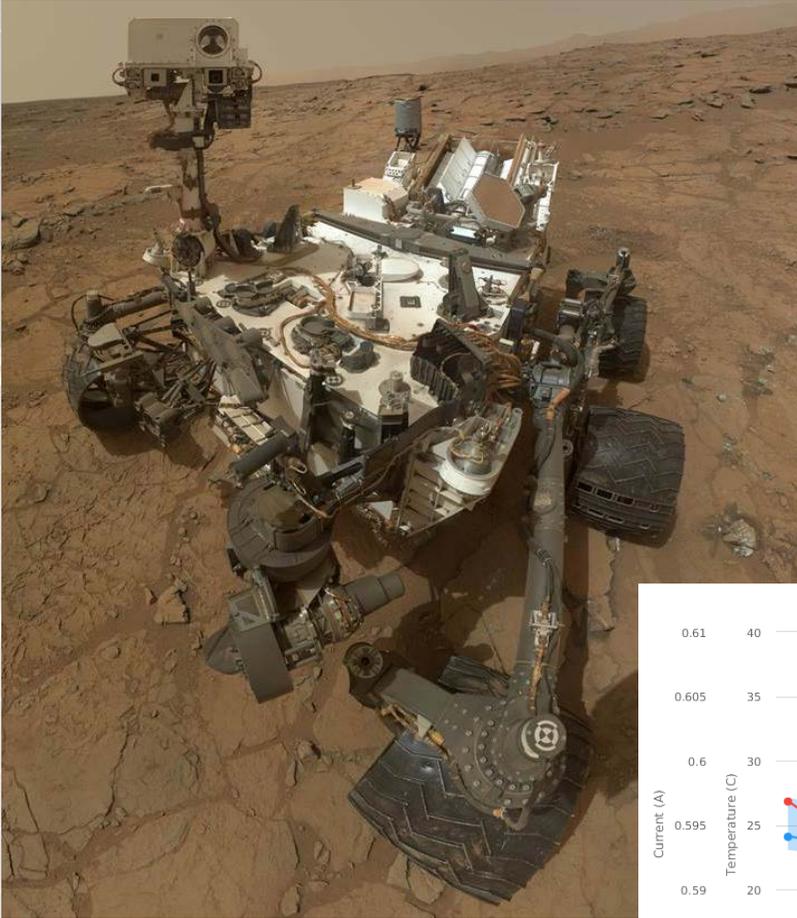
GEDI

Answering Questions

- How do we know SAM is healthy on Mars?
 - How much longer will it last?
- Is everything still working after landing?
- Has instrument performance degraded?
- How do the latest chromatograms compare to calibration tests?
- How much contamination is there in the instrument?
- Is the contamination reducing?
- How close similar is the new calibration target from the previous target?



Question: Is SAM Healthy?

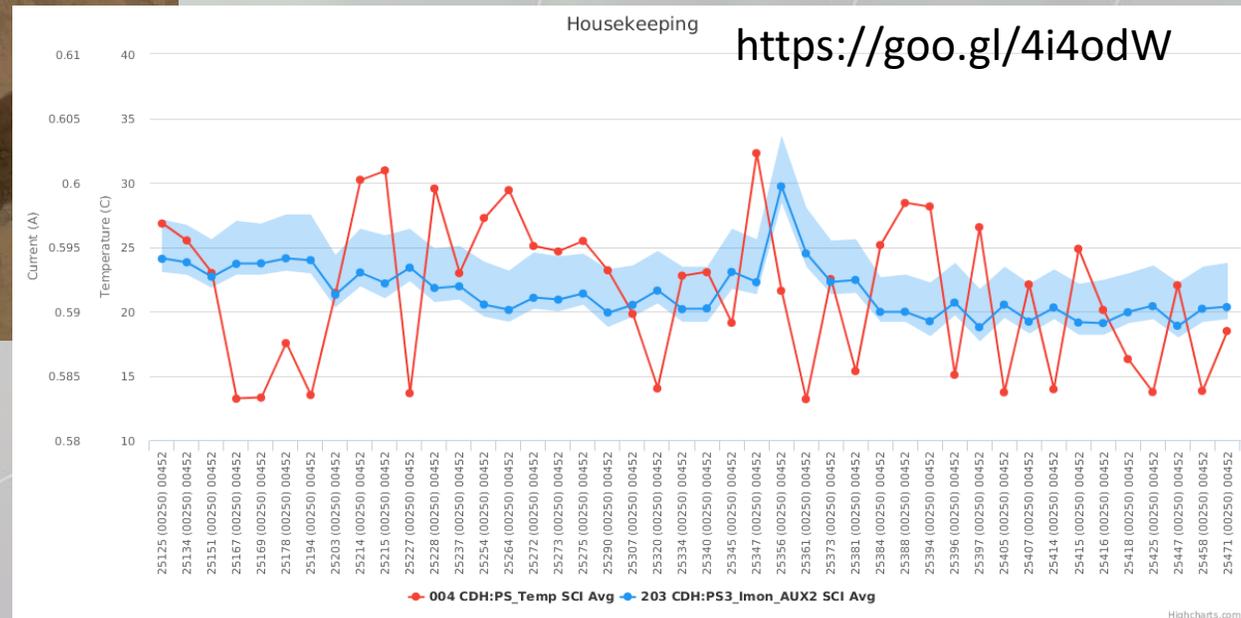


Sample Analysis at Mars instrument performs regular electrical baseline tests.

44 have been performed on Mars in 6.5 years since landing.

Each of the 44 test checks over 1200 data points.

XINA Web interface provide trend of every point.



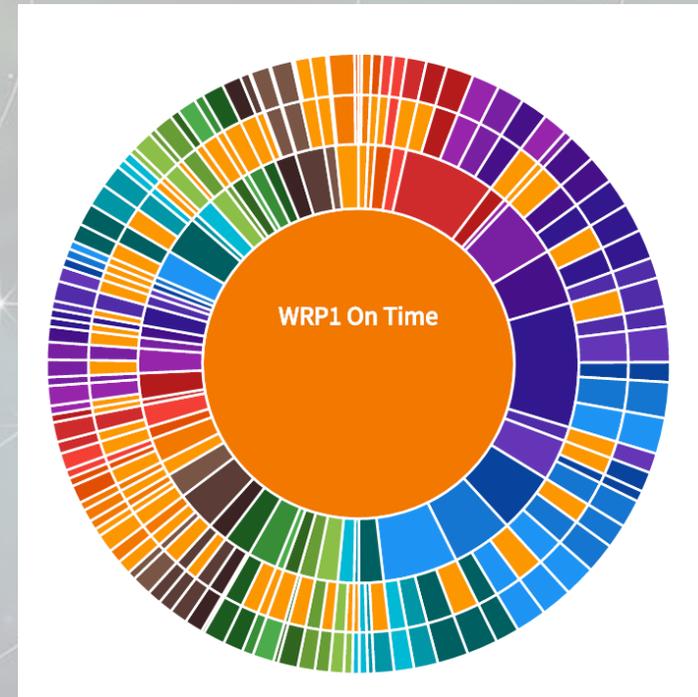
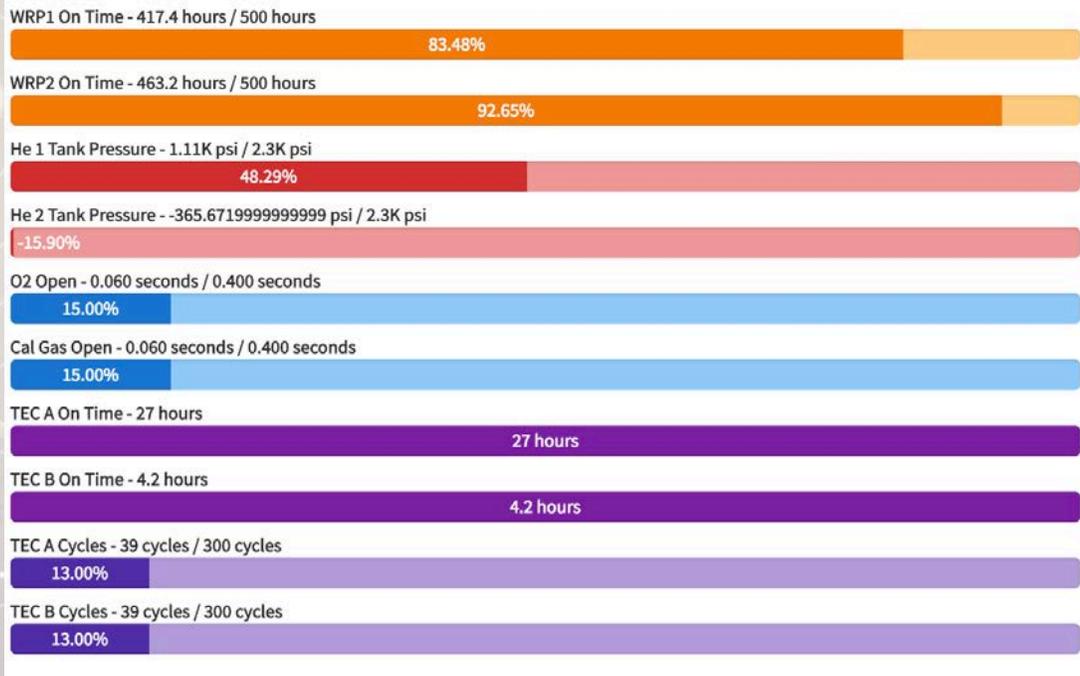
Question: How Much Longer Will Sam Last?

SAM has several limited-lifetime items:

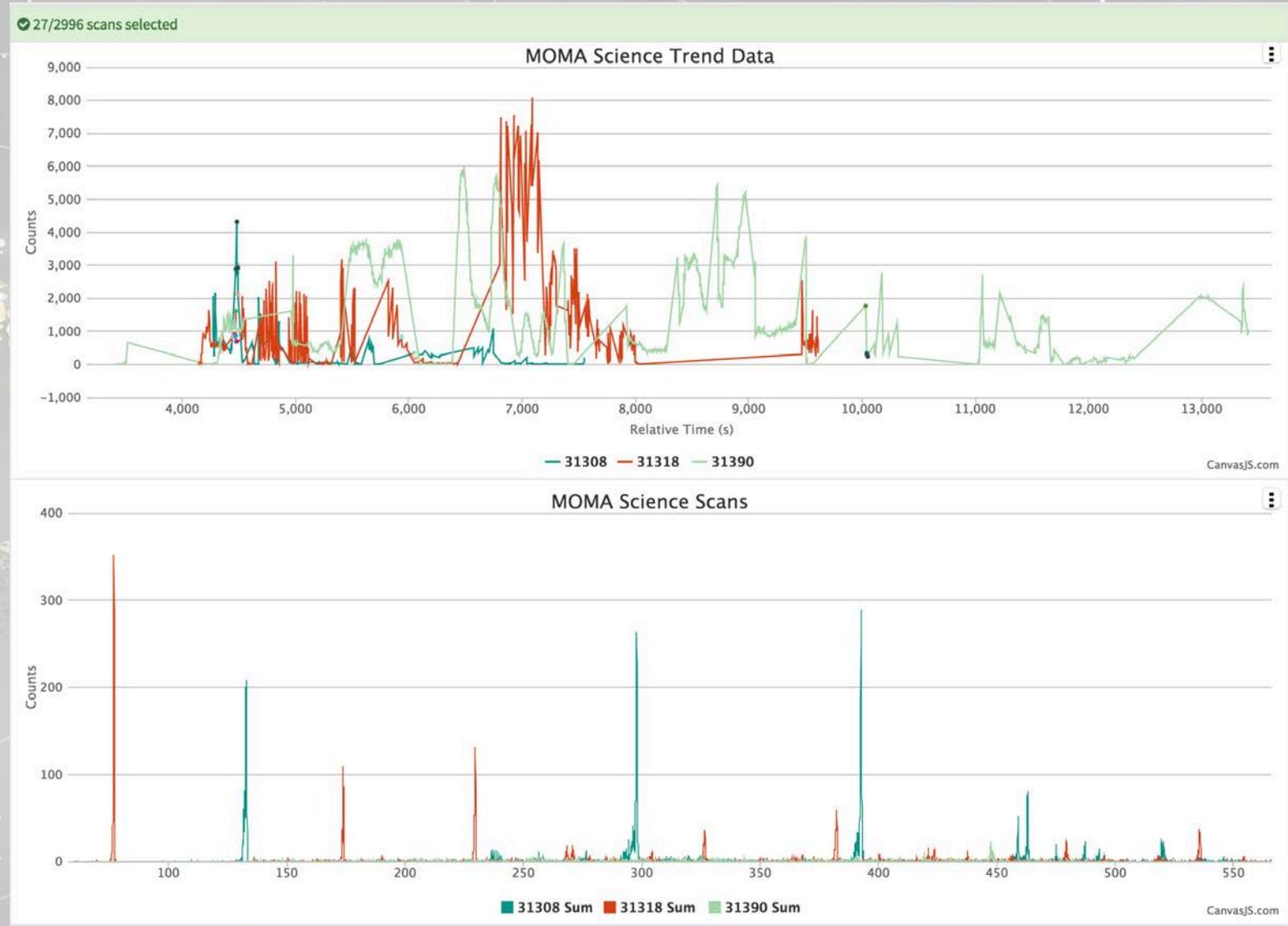
- Wide-range Pump
- Helium, Oxygen and Calibration Gas
- Thermo-electric coolers

XINA Tracks each consumable

<https://goo.gl/kU9f6D>



Question: How to the MOMA Laser Desorption Calibration scans compare?



Question: How do the latest chromatograms compare?

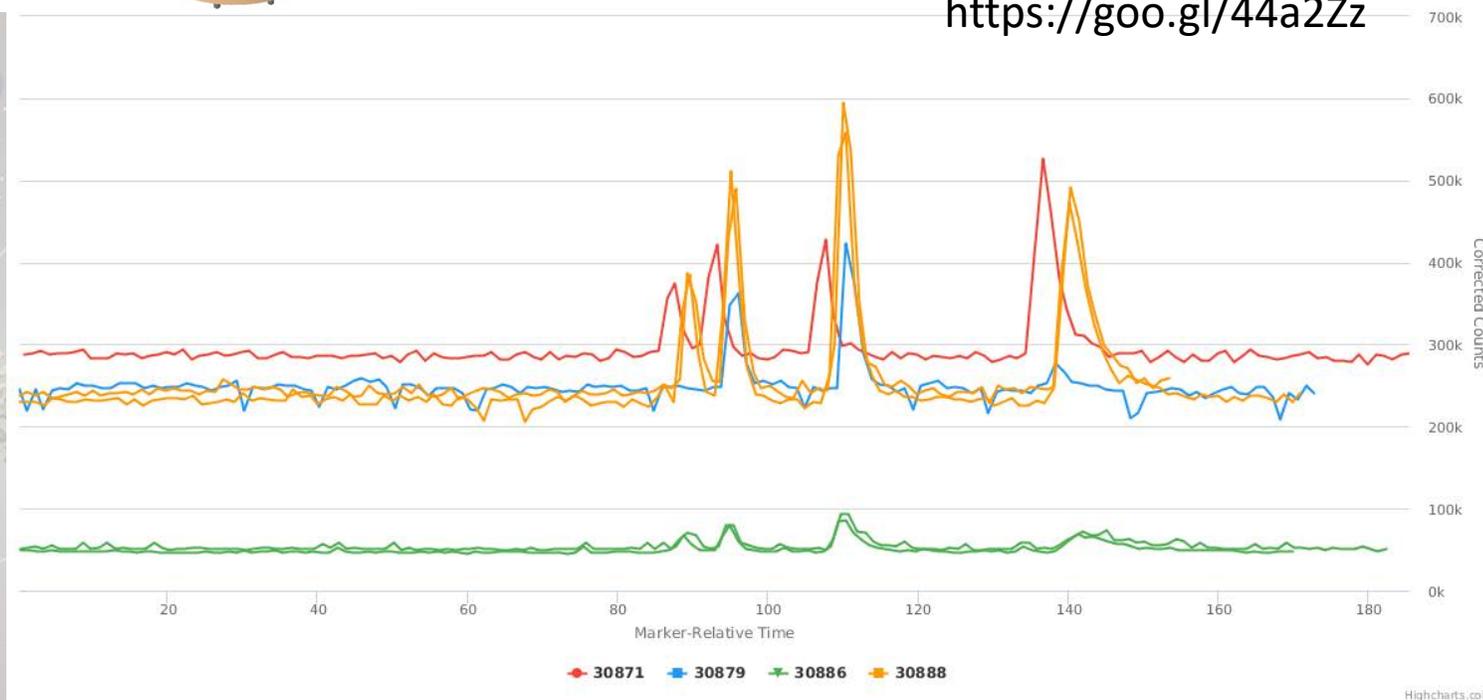


MOMA performs gas chromatograph pyrolysis injections with known gases.

XINA allows comparison of injections of the same gas at different times.

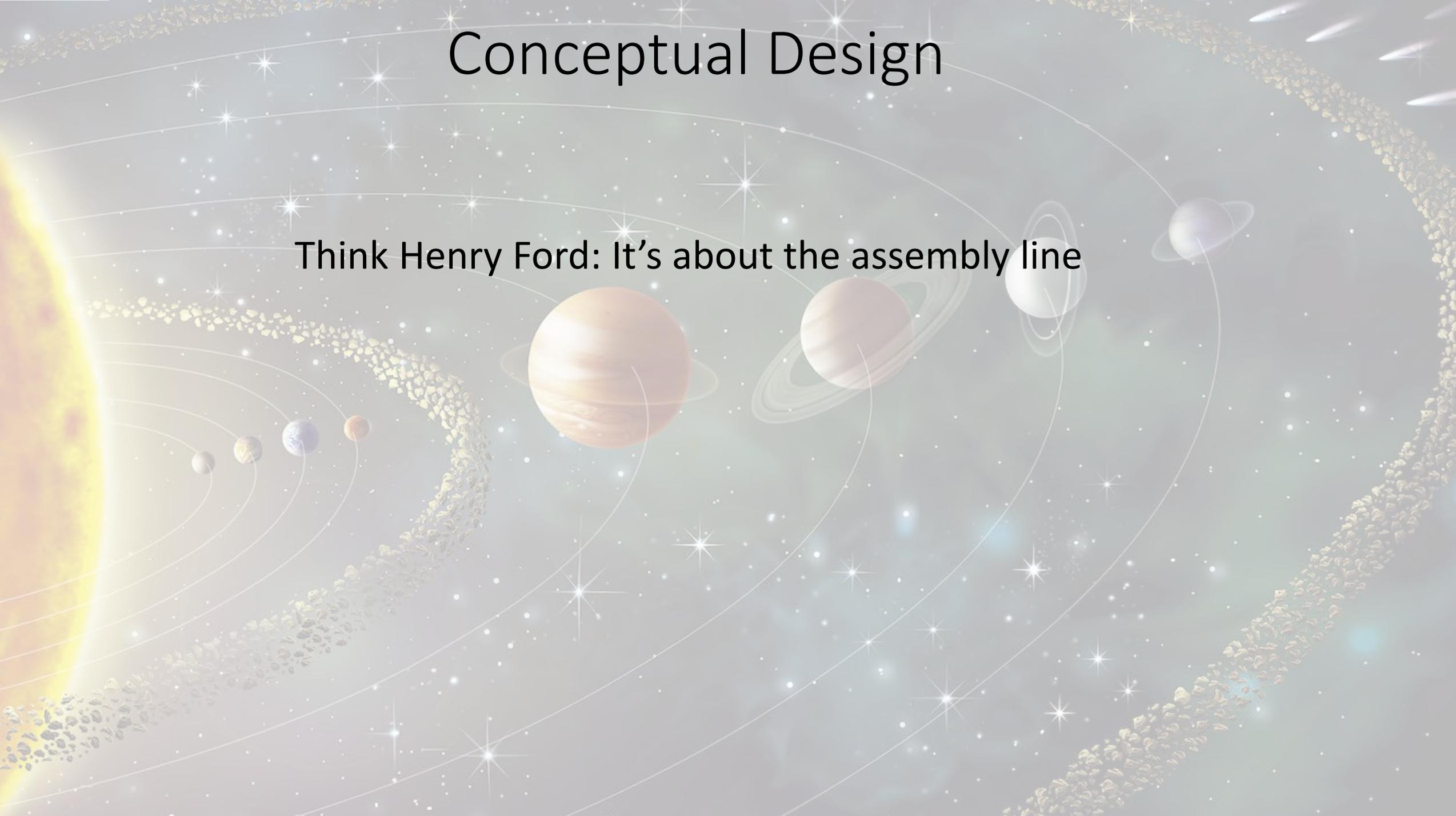
MOMA Chromatogram

<https://goo.gl/44a2Zz>

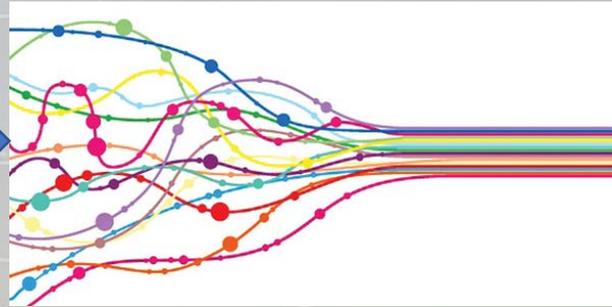
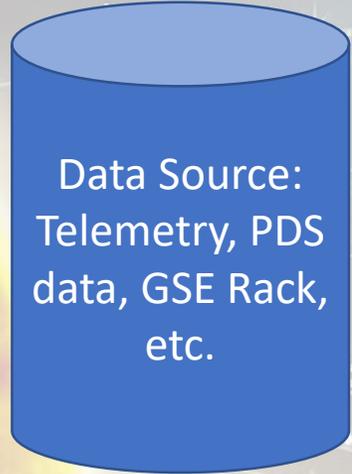


Conceptual Design

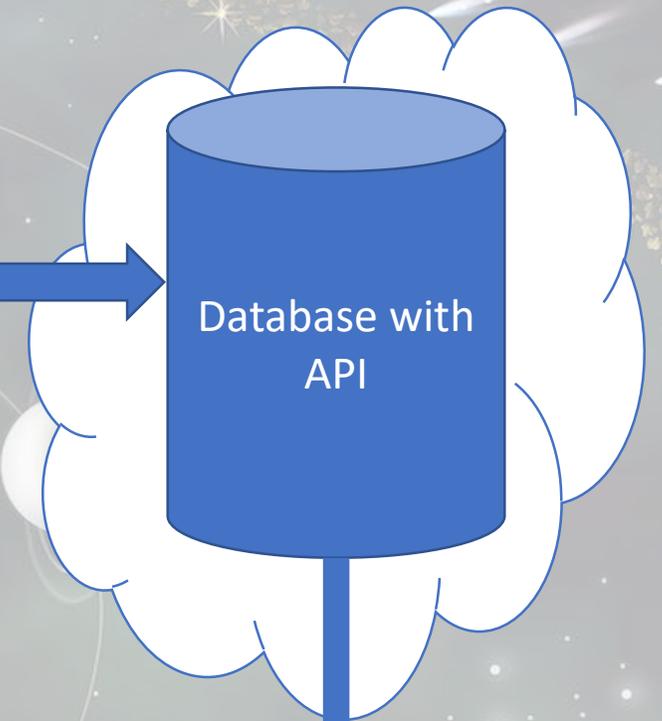
Think Henry Ford: It's about the assembly line



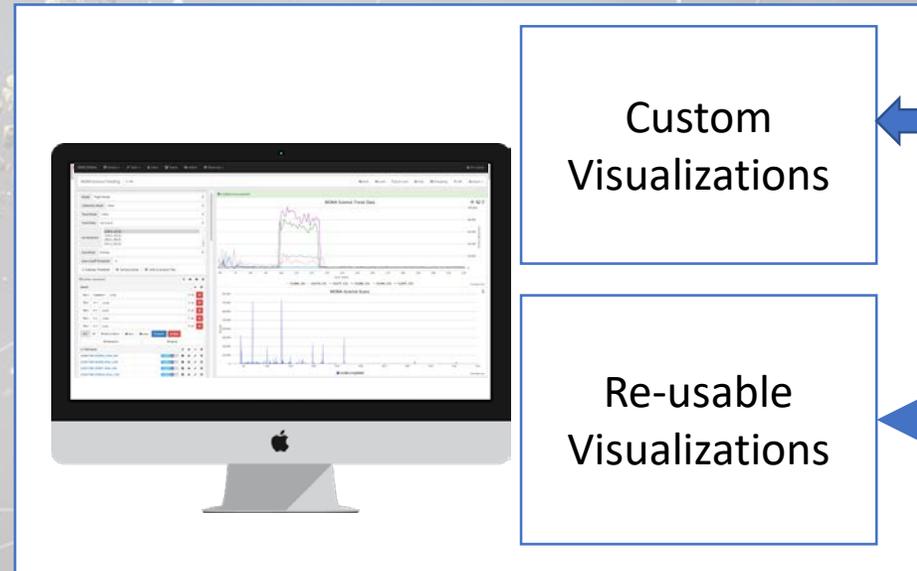
Conceptual Design



Software read generic
files and perform
additional calculations
to add value



JavaScript applications in the
browser provide many
visualizations without
modifications. New
visualization easily fit into
design.





Raw Telemetry



De-commutated data



Calibrated data



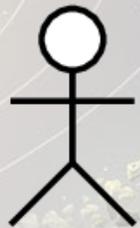
Archived



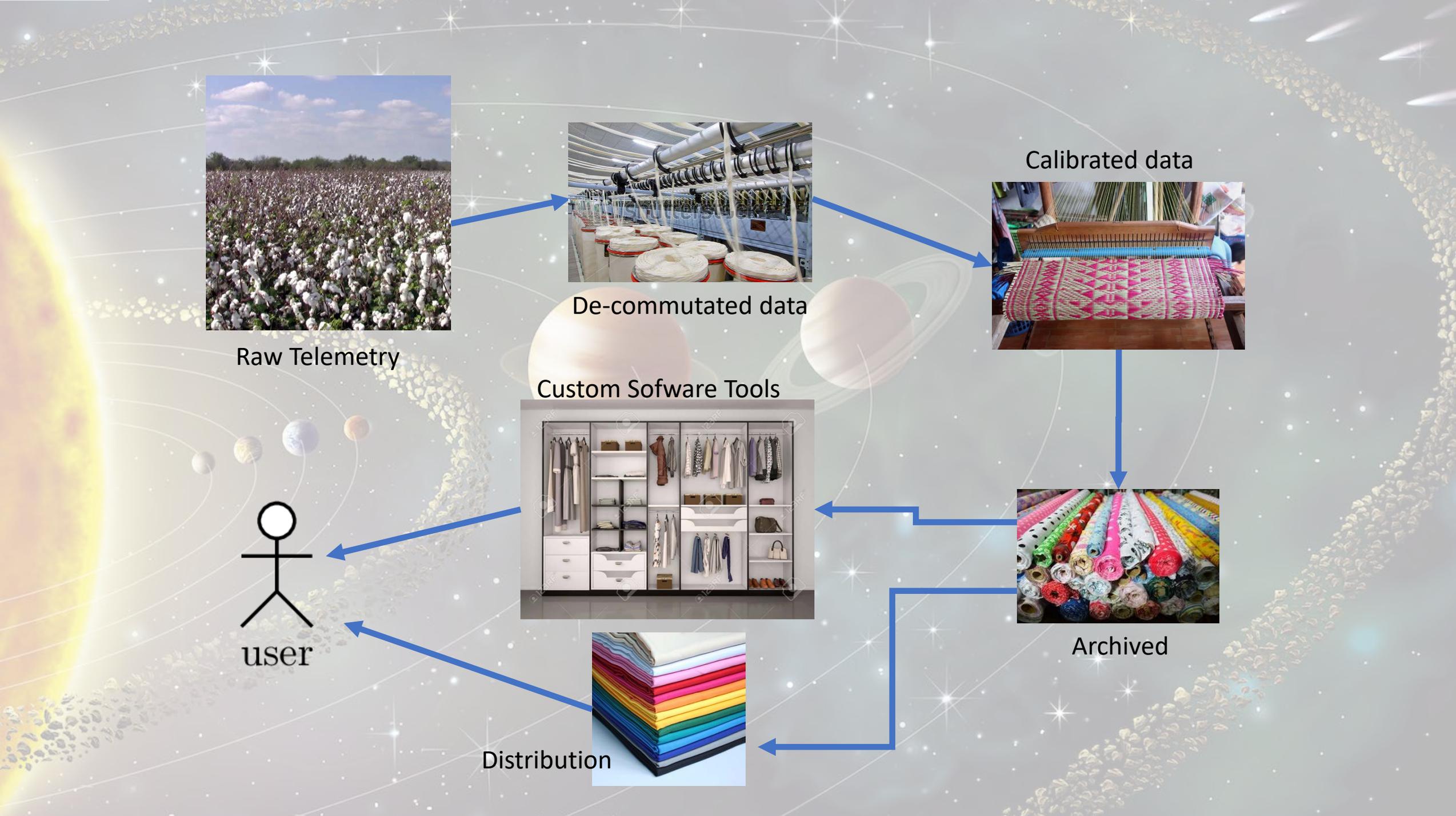
Custom Software Tools



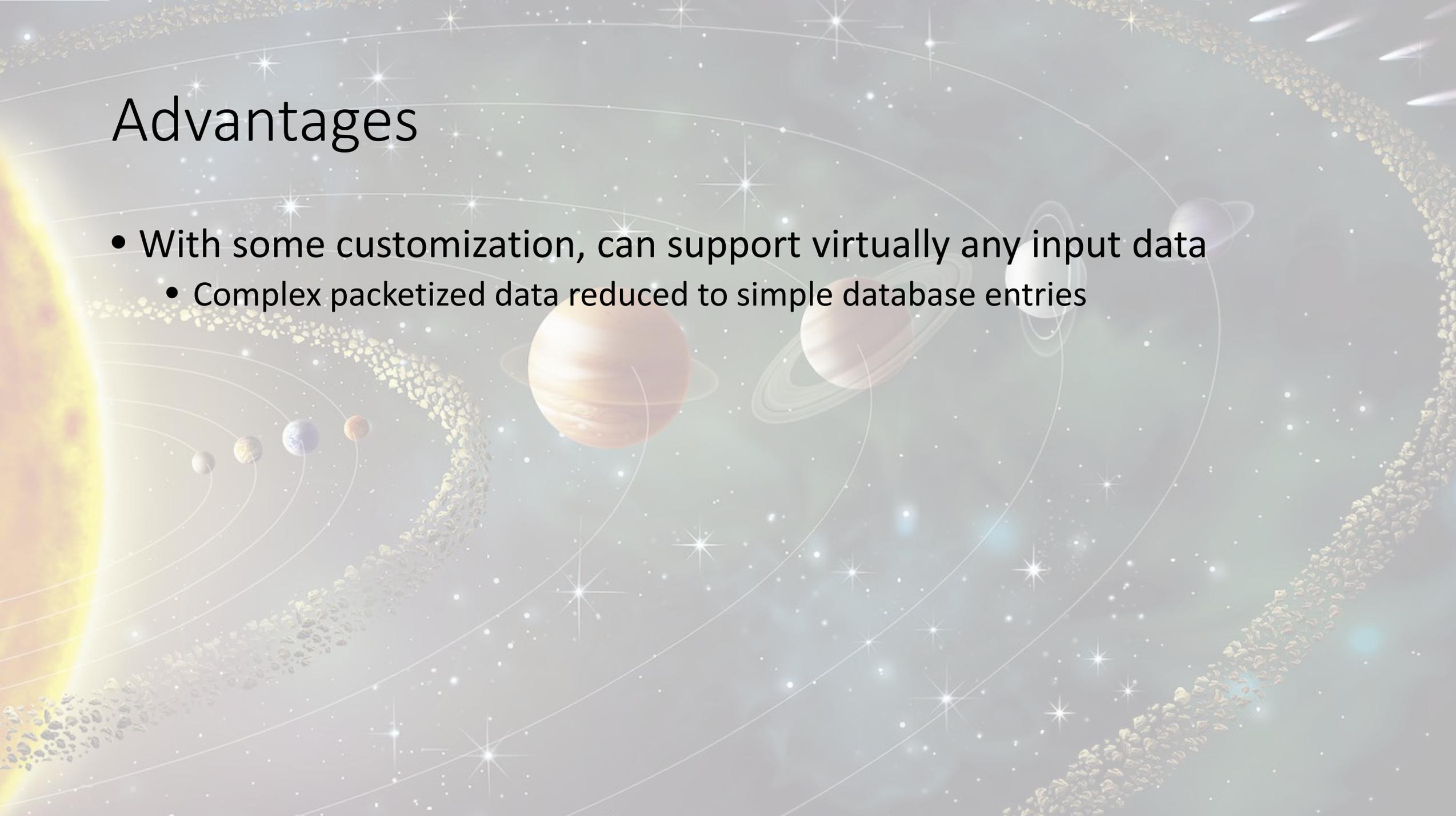
Distribution



user



Advantages

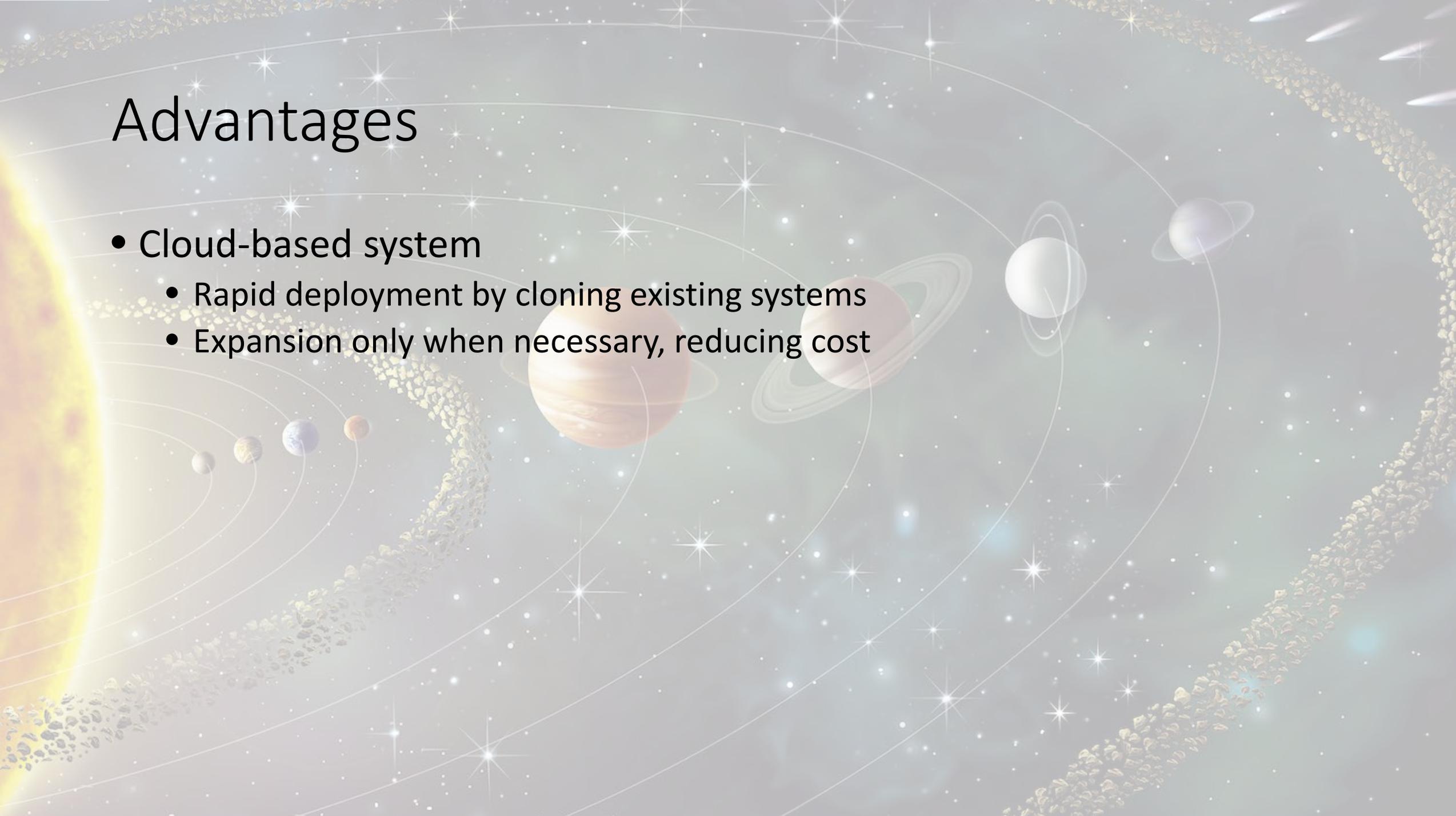


- With some customization, can support virtually any input data
 - Complex packetized data reduced to simple database entries

Advantages

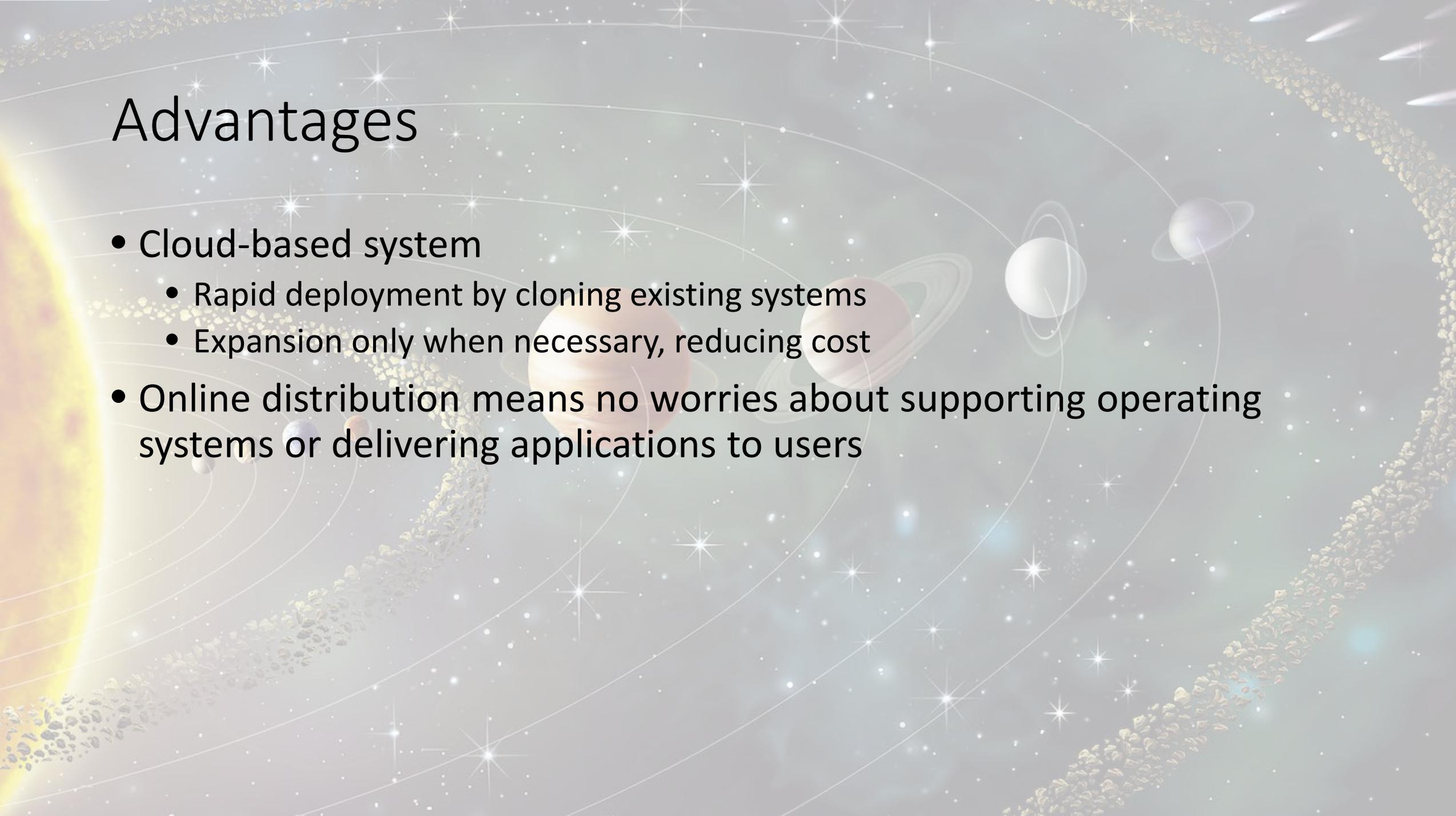
- With some customization, can support virtually any input data
 - Complex packetized data reduced to simple database entries
- Database reduces the problems of visualization and distribution to a computer science problem
 - Most software developers don't need to understand mechanics of spacecraft data.
 - Young programmers directly out of university can develop code
 - Reduces cost
 - Allows scalable work force

Advantages



- **Cloud-based system**
 - Rapid deployment by cloning existing systems
 - Expansion only when necessary, reducing cost

Advantages



- Cloud-based system
 - Rapid deployment by cloning existing systems
 - Expansion only when necessary, reducing cost
- Online distribution means no worries about supporting operating systems or delivering applications to users

Advantages

- The interface is modern and developed according to 2018 software development processes
 - Regular, seamless updates
 - The latest web-based look and feel
 - Automatically scales for mobile devices
 - Attractive to the newest, brightest developers
 - Comfortable interfaces for users who expect Amazon- or Google-like experiences
- NPR 7150.2 Compliant

Near Future

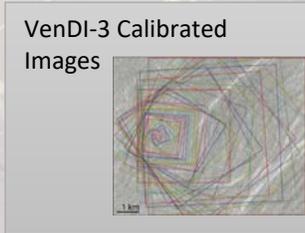
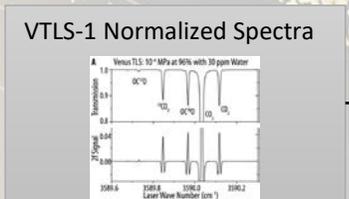
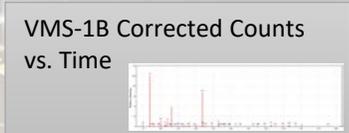
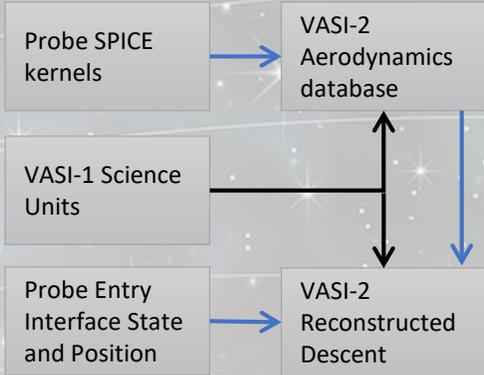
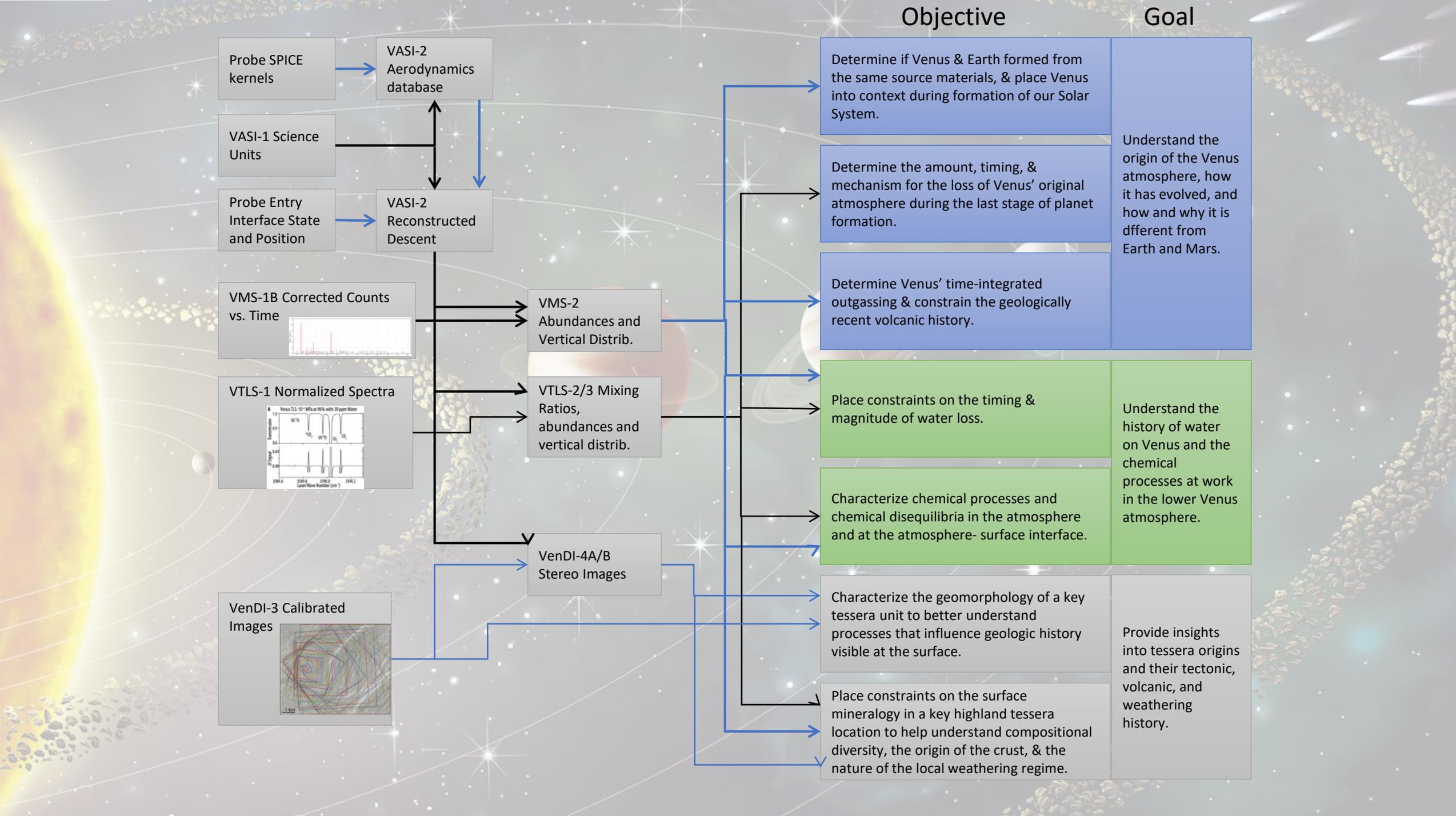
- Integrate MOMA data with other ExoMars instruments to include their products
- Sample-based interface that allows users to select the solid sample then view each instrument's analyses of that sample
- Ingest more PDS data from MSL mission
- Support new missions, such as the Ocean Color Instrument
- Track flight hardware subsystems as well as complete instruments
 - For instance, record life test data developed in a lab for a pump

Part 4: The Future – Deductive Approach

- One innovation PEL had was to place hooks in flight software and scripts to aid in post-processing.
 - This made it possible to answer the questions the engineers and scientists had about the instrument
- Should PDS be considering the questions about the planets and influencing the data collection such that the questions can be answered?

Part 4: The Future, A Deductive Approach?

- Each mission aims to answer *specific questions* from the decadal survey *Visions and Voyages*.
- Each mission documents the instruments and data required to answer each question
- A Deductive Approach *starts* with the question
 - Example for a Venus mission: Did Earth and Venus form from the same source material?
 - What happened to the original atmosphere of Venus?



VMS-2 Abundances and Vertical Distrib.

VTLS-2/3 Mixing Ratios, abundances and vertical distrib.

VenDI-4A/B Stereo Images

Determine if Venus & Earth formed from the same source materials, & place Venus into context during formation of our Solar System.

Determine the amount, timing, & mechanism for the loss of Venus' original atmosphere during the last stage of planet formation.

Determine Venus' time-integrated outgassing & constrain the geologically recent volcanic history.

Place constraints on the timing & magnitude of water loss.

Characterize chemical processes and chemical disequilibria in the atmosphere and at the atmosphere- surface interface.

Characterize the geomorphology of a key tessera unit to better understand processes that influence geologic history visible at the surface.

Place constraints on the surface mineralogy in a key highland tessera location to help understand compositional diversity, the origin of the crust, & the nature of the local weathering regime.

Understand the origin of the Venus atmosphere, how it has evolved, and how and why it is different from Earth and Mars.

Understand the history of water on Venus and the chemical processes at work in the lower Venus atmosphere.

Provide insights into tessera origins and their tectonic, volcanic, and weathering history.

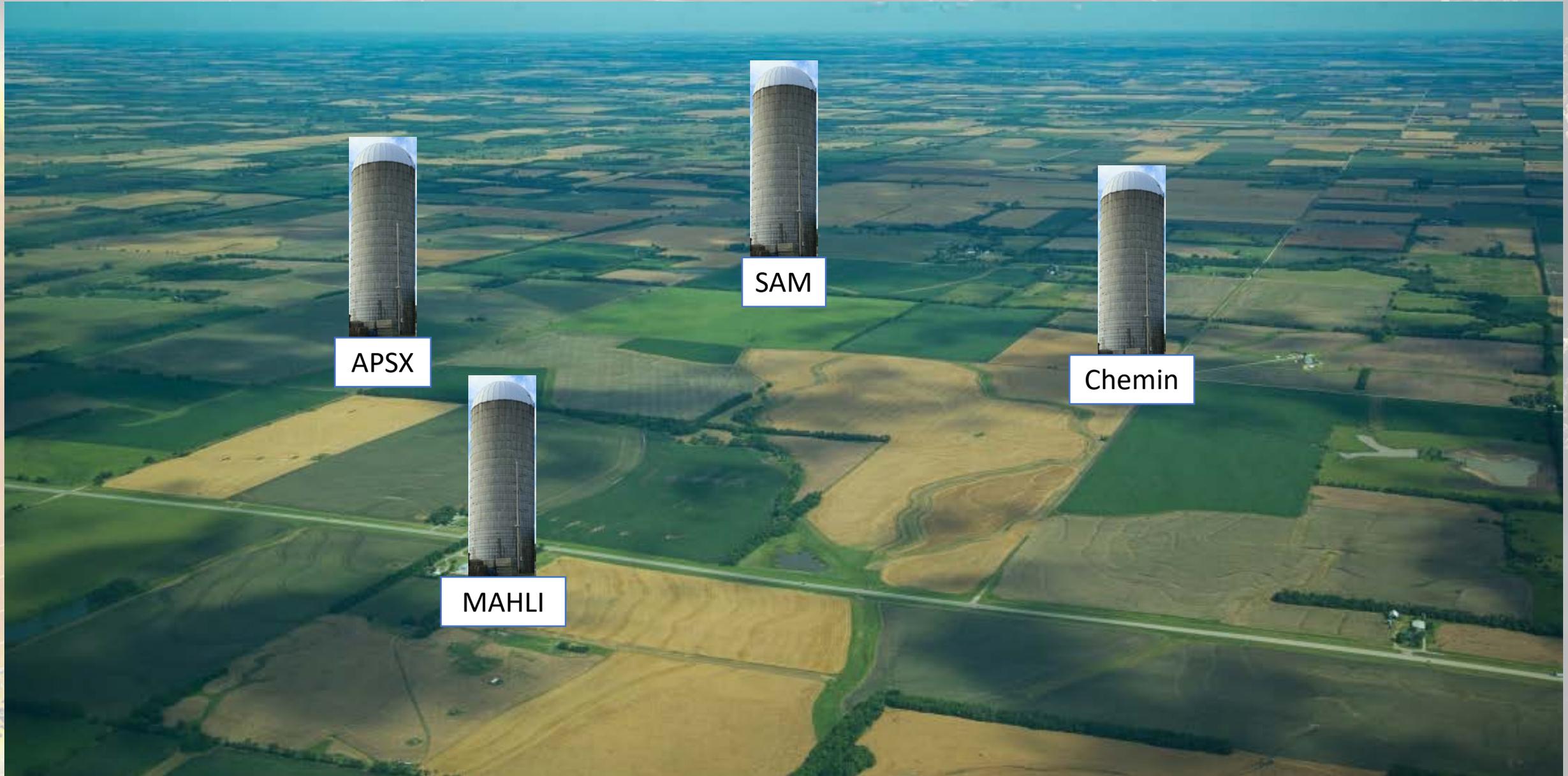
Part 4: The Future, A Deductive Approach?

- Instruments and missions don't just identify data products, they also identify what question it helps to answer
- A Mission Data Management leader, cutting across instrument teams, could help make this happen

Back to the Future



The current landscape



The current landscape, zoomed out

Opportunity



InSight



MSL



Better Landscape

Opportunity



InSight



MSL

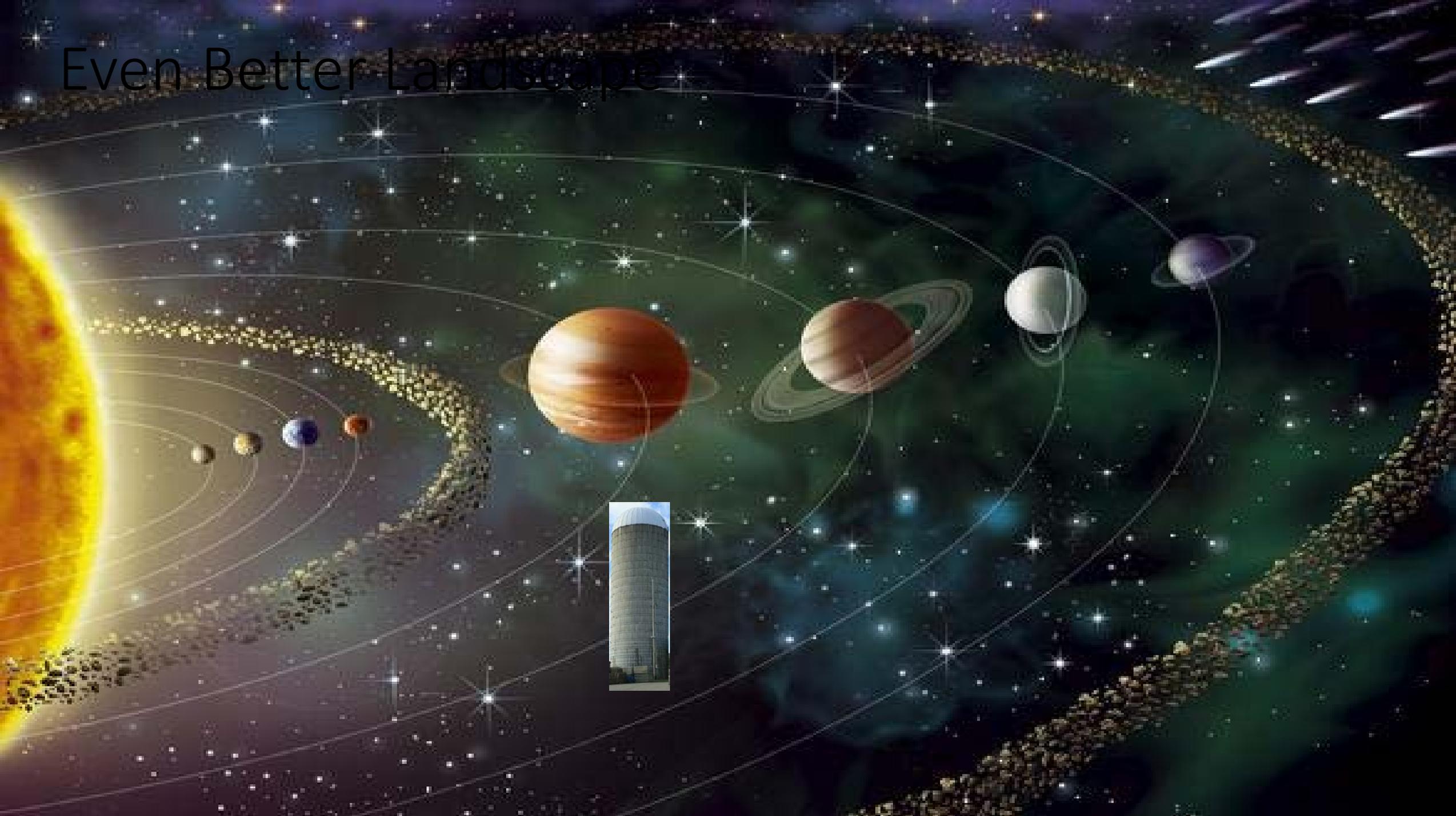


Even Better Landscape

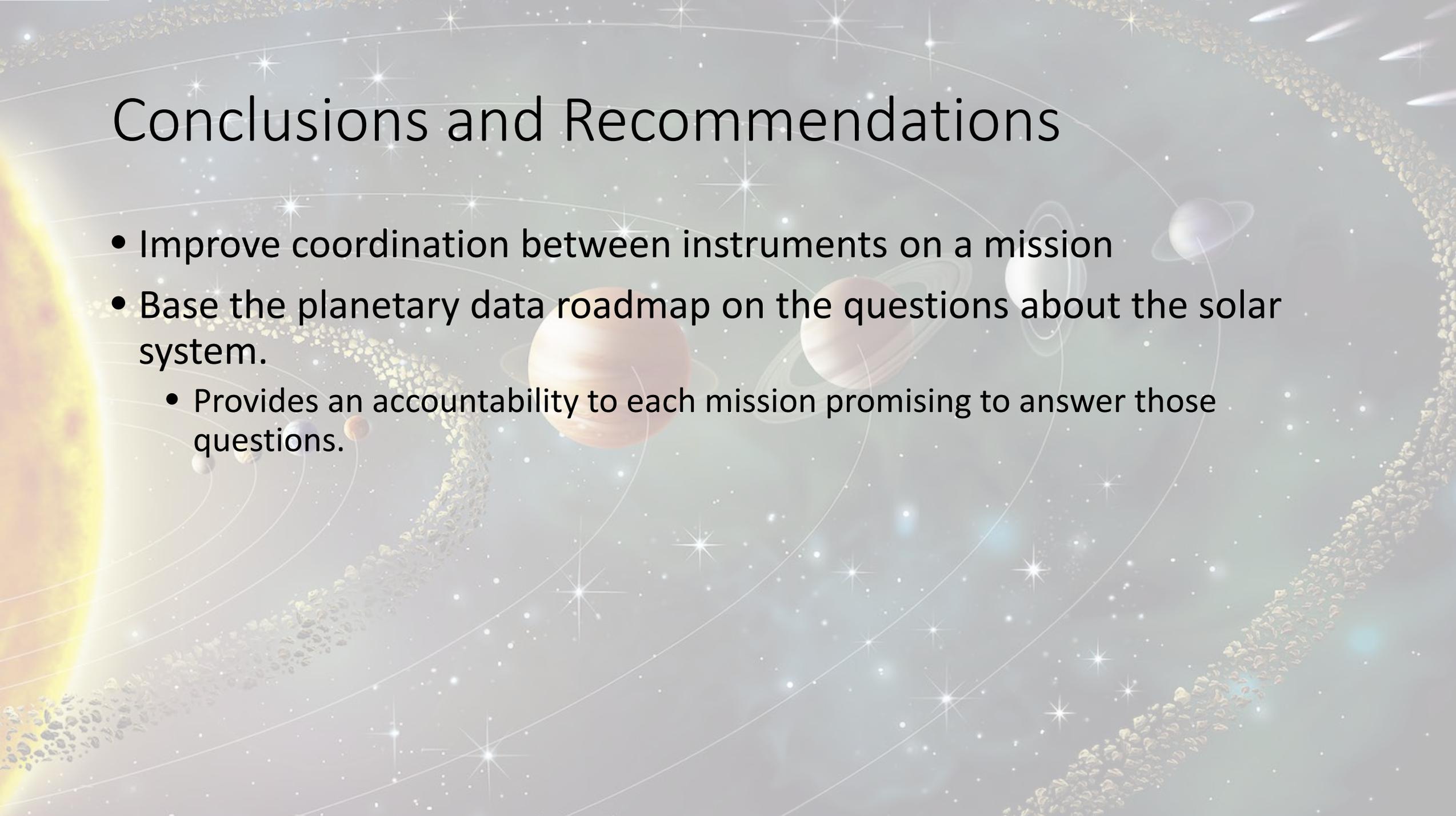
Mars



Even Better Landscape



Conclusions and Recommendations



- Improve coordination between instruments on a mission
- Base the planetary data roadmap on the questions about the solar system.
 - Provides an accountability to each mission promising to answer those questions.