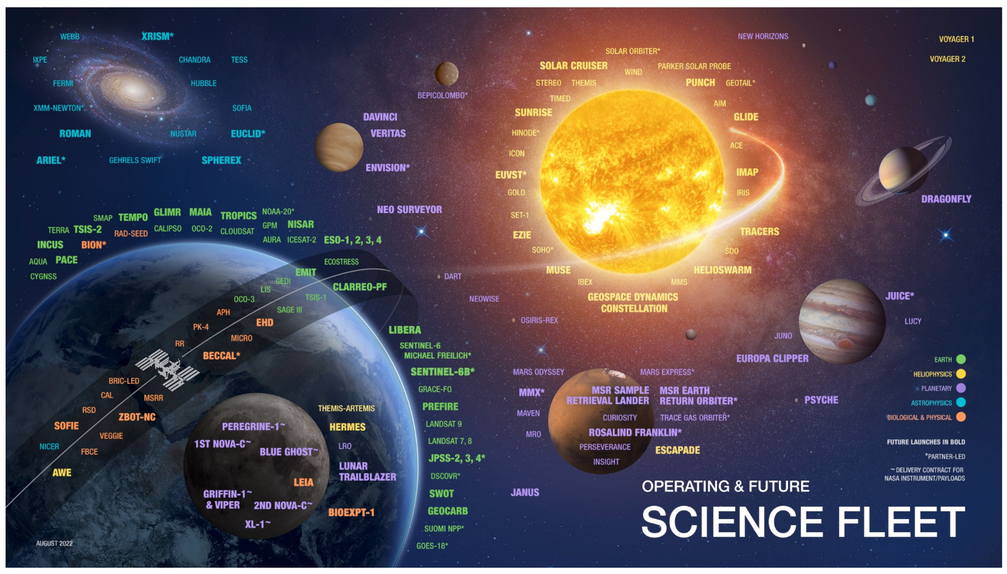
Data Science at MSFC ST

Dr. Rahul Ramachandran
Dr. Manil Maskey
NASA IMPACT/ST11

Data Science

There is a challenge of managing overwhelming data and distilling it into meaningful knowledge



"Data science is a broader concept involving principles for data collection, storage, integration, analysis, inference, communication, and ethics appropriate for this new data-driven era": NAS 2018

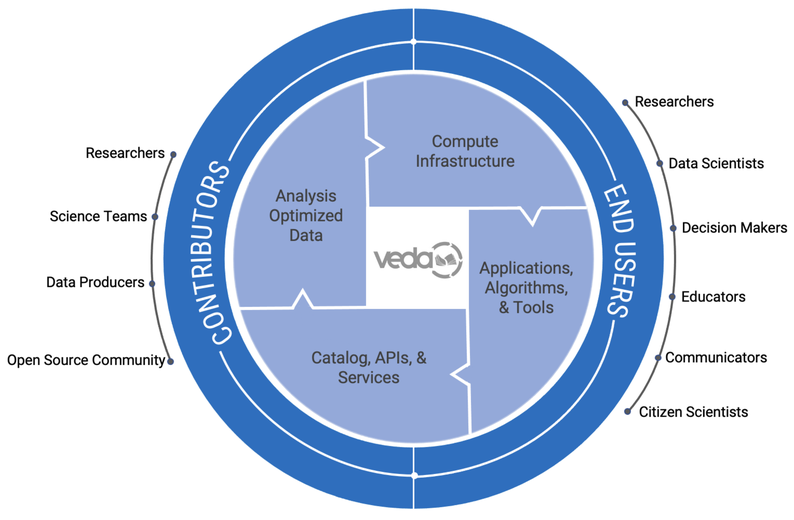
Data Science Thrust Areas

1. Science Cyberinfrastructure
2. Data Management & Stewardship
3. Artificial Intelligence
4. Partnerships & Outreach

1. Science Cyberinfrastructure

Enabling new science at scale through analytics platforms

Visualization, Exploration, and Data Analysis (VEDA) Platform



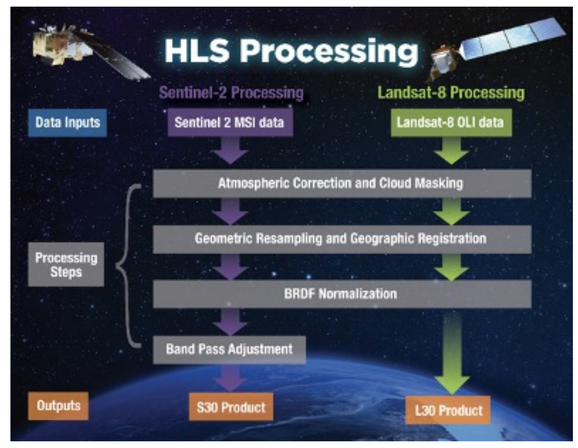
Example application: US Greenhouse Gas Center

2. Data Management and Stewardship

Managing the complexities of big data effectively to support modern scientific endeavors

Production of Harmonized Landsat Sentinel-2 dataset

- Merging Sentinel-2 and Landsat data streams to provide 2–4-day global coverage
- Collaboration among NASA GSFC, UMD, ESA
- Cloud-native data processing for scalability

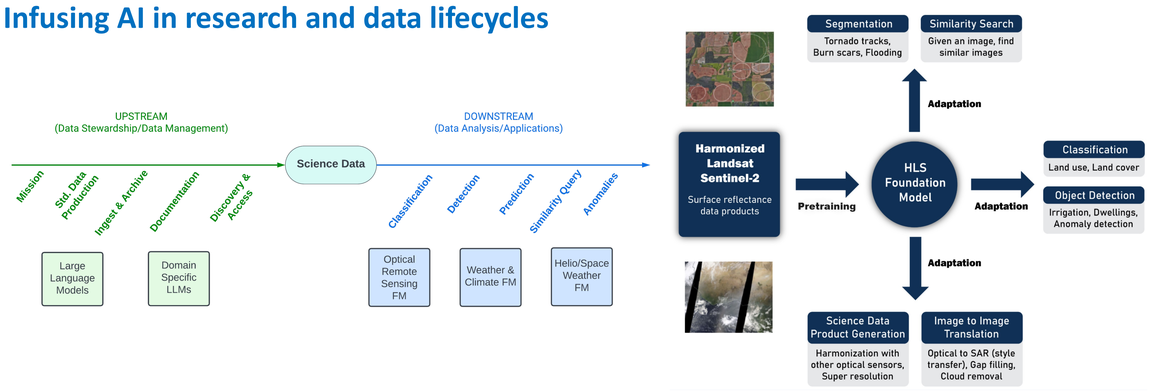


Impact

- ~8000 citations
- 2 PB+ data transferred to users
- 2nd most downloaded data from NASA archive
- Featured on weather.com, nature.com, ibm.com

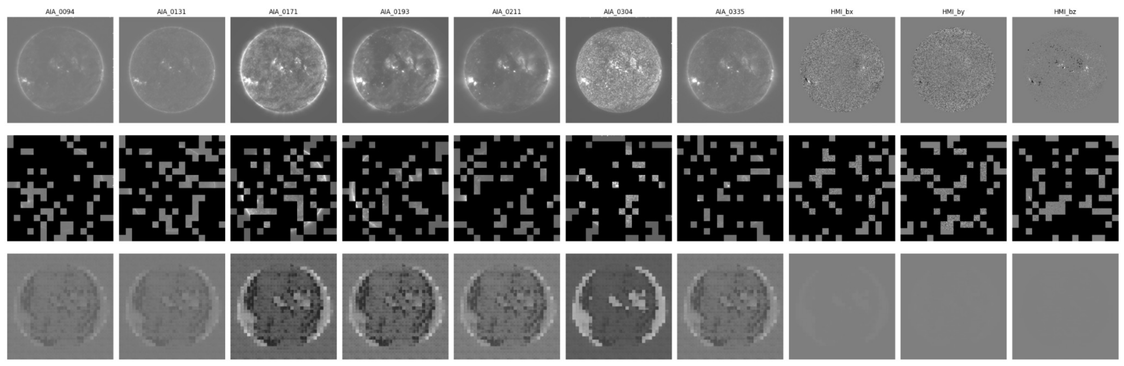
3. Artificial Intelligence

Infusing AI in research and data lifecycles



AI adoption both upstream and downstream

Geospatial foundation model



Foundation model for Heliophysics using SDO data

4. Partnerships and Outreach

Fostering innovation via strategic partnerships and empowering science community to use new tools and methodologies

IEEE Summer School on High Performance and Disruptive Computing in Remote Sensing

Takeaways

- With upcoming high-data-rate missions, the challenge of scale intensifies.
- Data science is poised to be an integral component in science research, essential for managing and deriving insights from large science datasets.

Supported by:
Office of Chief Science Data Officer, NASA/SMD
Earth Science Data Systems Program, NASA/ESD

Contact:
rahul.ramachandran@nasa.gov
manil.maskey@nasa.gov