Science Mission Directorate
Earth Science Data Systems

OSSI SMD AI activities

Manil Maskey, Ph.D.

OSSI Workshop
October 14, 2021

OSSI SMD AI team

Srija Chakraborty
AI SME/Post Doc

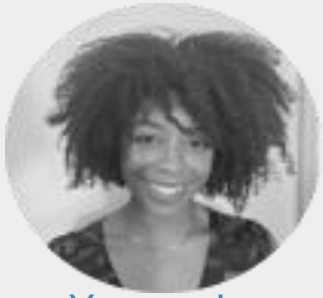
Dan Duffy
HECC

Lika Guhathakurta
Heliophysics

Mike Little
HECC

Megan Ansdell
Planetary Science

Roopesh Ojha
Astrophysics



Yvonne Ivey
SDMWG

Sylvain Costes
Biological and
Physical Science

Evan Scannapieco
Astrophysics

Manil Maskey
Earth Science/Lead

NASA Science Mission Directorate (SMD) Artificial Intelligence and Machine Learning (AI/ML) Initiatives

SMD's Strategy for Data Management and Computing for Groundbreaking Science 2019-2024 Report identified that AI/ML has yet to be fully appreciated and understood by SMD and science disciplines

Strategy 2.4: Invest in the tools and training necessary to enable breakthrough science through application of AI/ML

Recommendation 11: SMD should make investments to incentivize and educate the community on how to use AI/ML to approach science in new ways. Hands-on training can be achieved through expansion of hackathons, competitions, and grant programs. Science results and lessons learned about the use of AI/ML will be shared at community meetings to increase awareness of the potential of these techniques.

Activities:

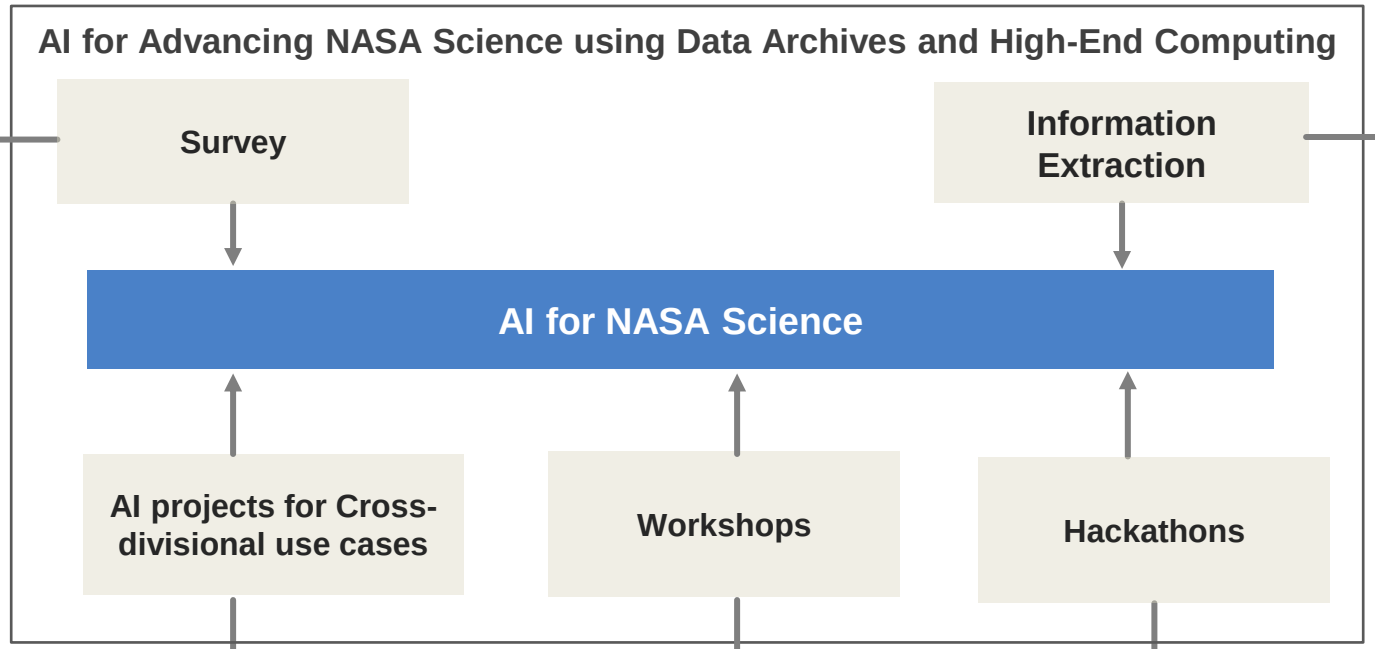
- Identify areas of natural collaborations on AI/ML across SMD
- Conduct expert workshop on AI for science
- Explore industry partnership
- Develop a roadmap to leverage **large volumes of data and computation** to accelerate AI/ML

SMD SDMWG AI/ML

Goal of determining the Current state of the art, Workforce, Policies, Datasets, Computation, Needs for AI/ML

Participants: NASA and NASA funded Researchers

Goal of building a knowledgebase of AI activities within NASA science objectively; understanding trends in AI adoption, challenges, datasets, recommendations using scientific literature, proposals, projects



Goal of demonstrating early win projects in utilizing AI for cross divisional use cases within NASA science, understanding challenges,

Participants: NASA funded researchers

Goal of addressing challenges in AI for NASA science from data and computation perspective, discussing tools and framework, strategy for making NASA data AI-ready, benchmarking, and open sharing for competitions and re-use, path forward for accelerating AI research and applications

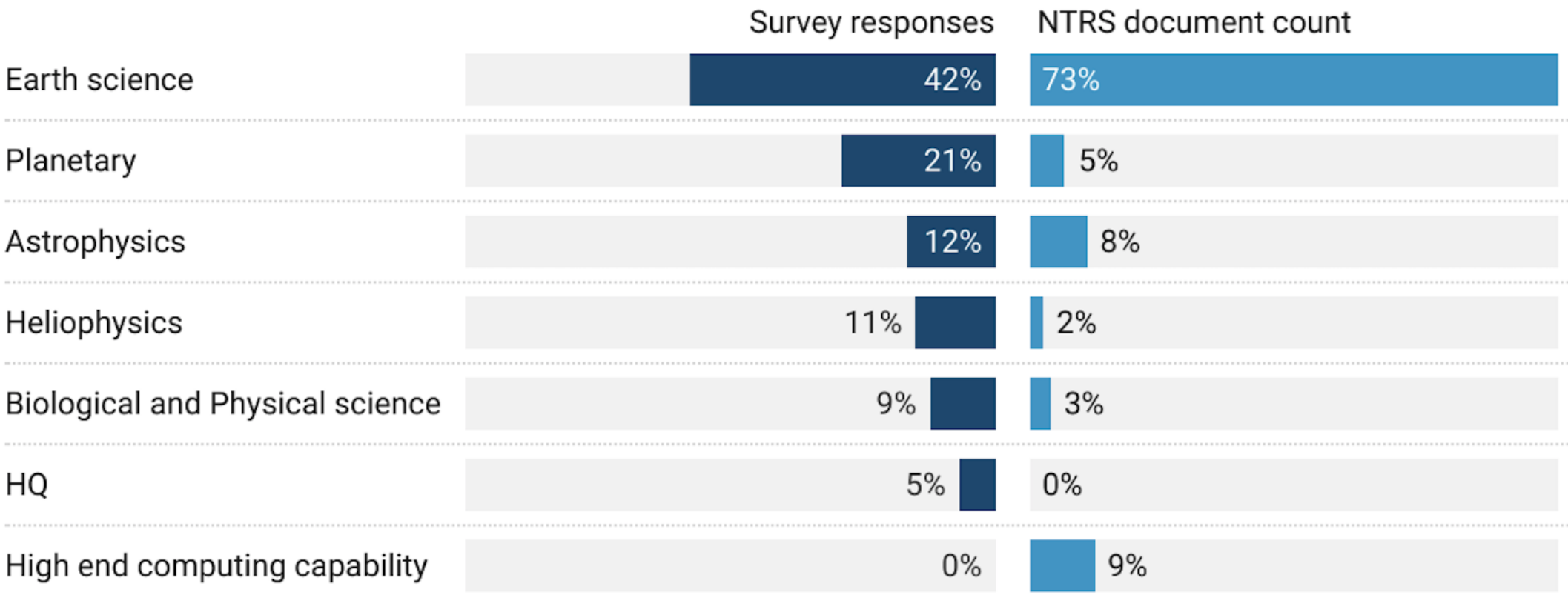
Participants: Industry, Academia and NASA AI experts

Goal of providing AI implementation experience, fostering collaboration – Team Science, utilizing existing benchmark NASA science datasets,
utilize computational resources, solve cross-divisional problems, identifying challenges

Participants: Anyone (in US)

AI within NASA SMD Divisions/Programs

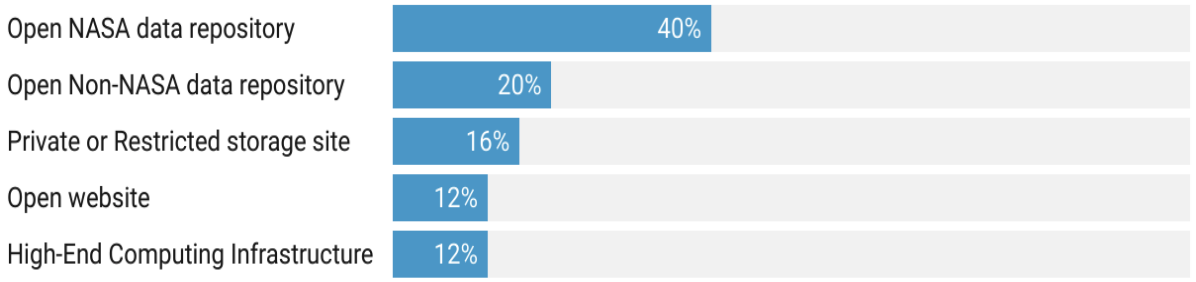
SMD AI Survey Responses and NTRS Documents



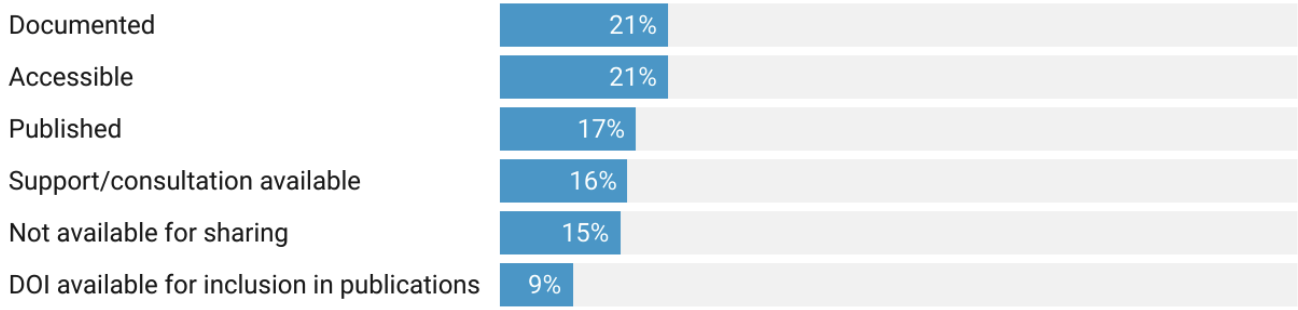
560 total survey responses 8317 total NTRS documents

Survey response - AI and data

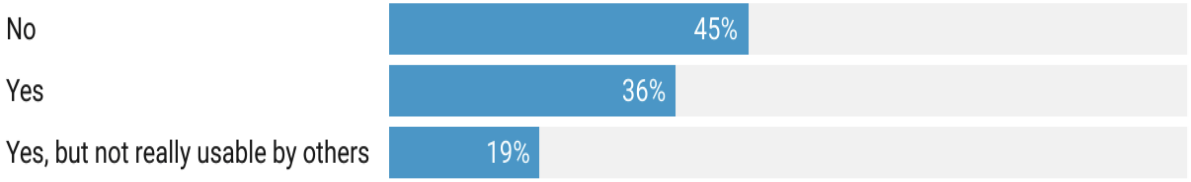
Source of data used



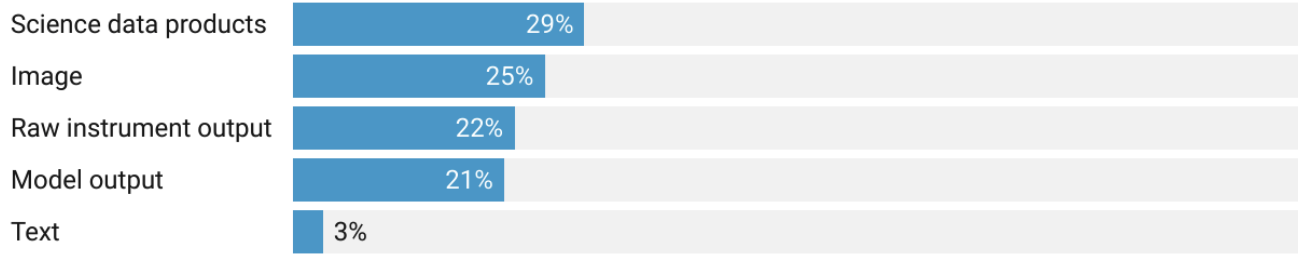
How re-usable is your training data?



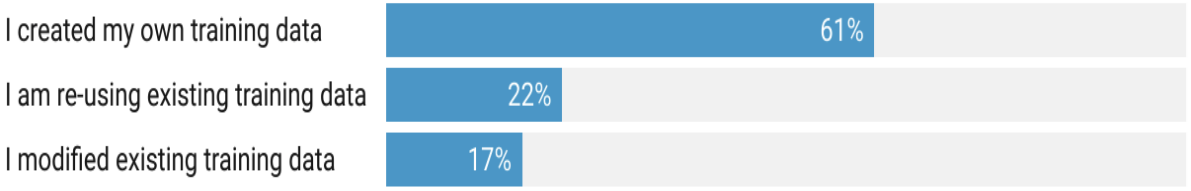
Is there a catalog of training data for your use?



What type of data do you use for AI?

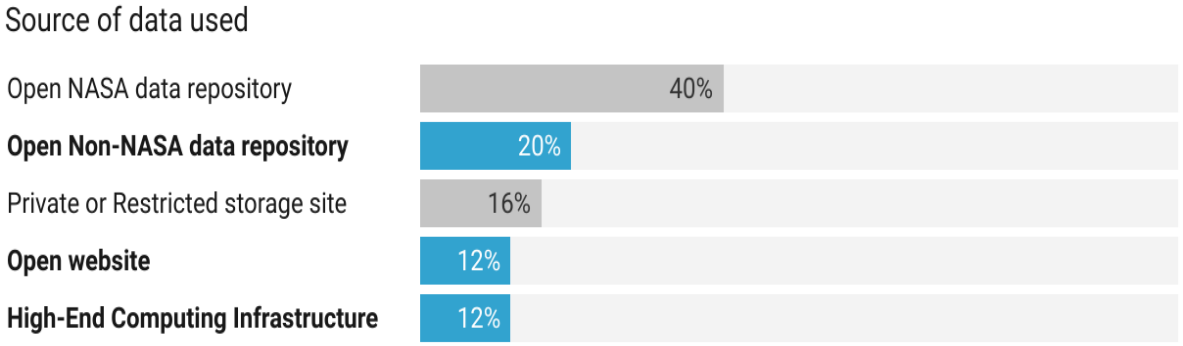


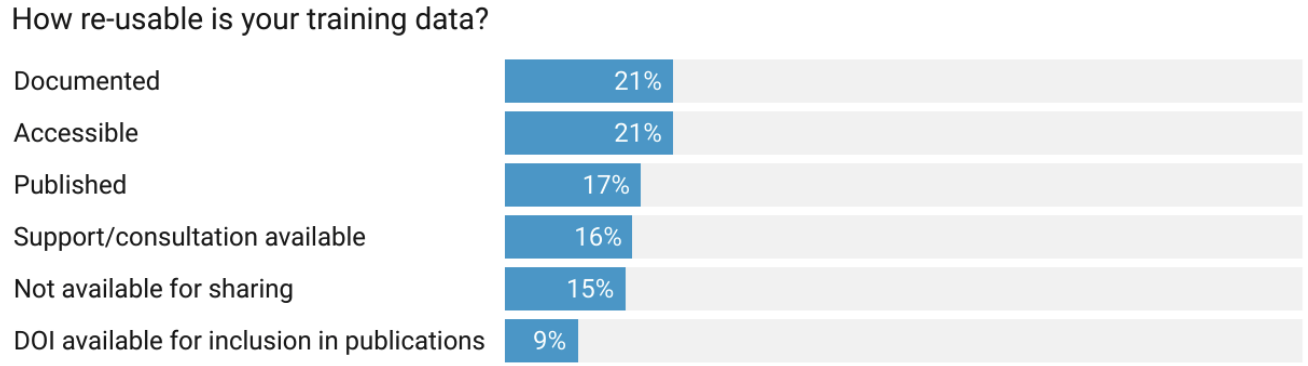
How did you construct training data?

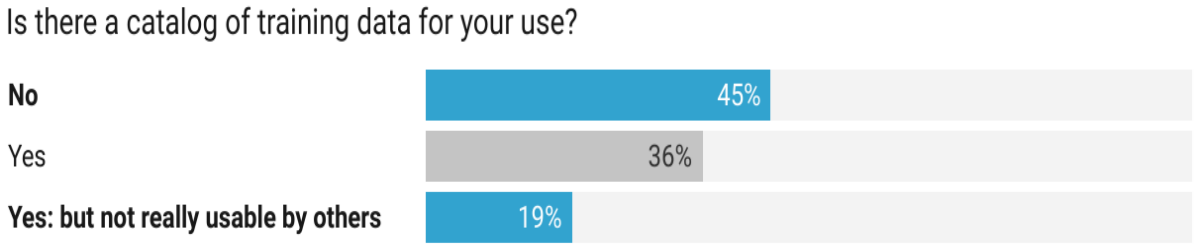


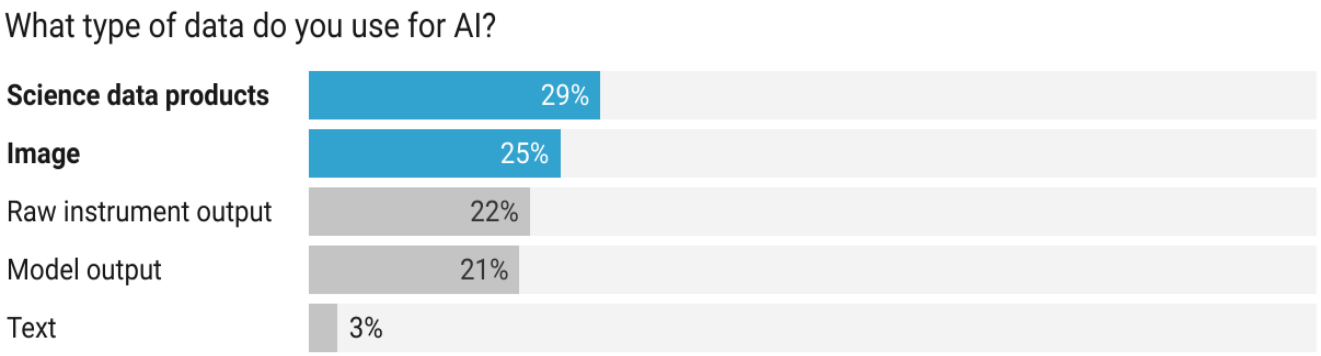
Amount of effort required to prepare data for AI?

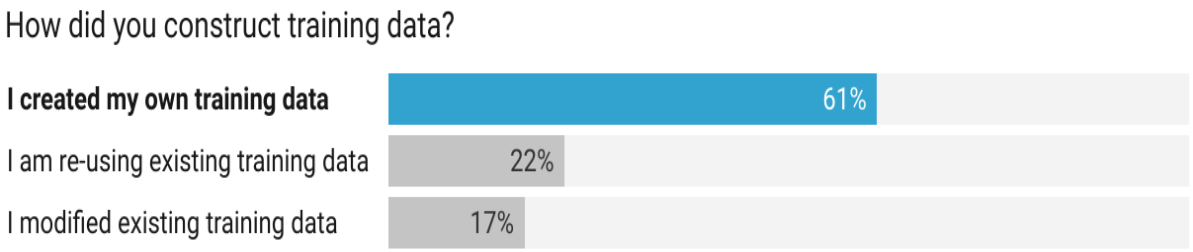
Survey response - AI and data

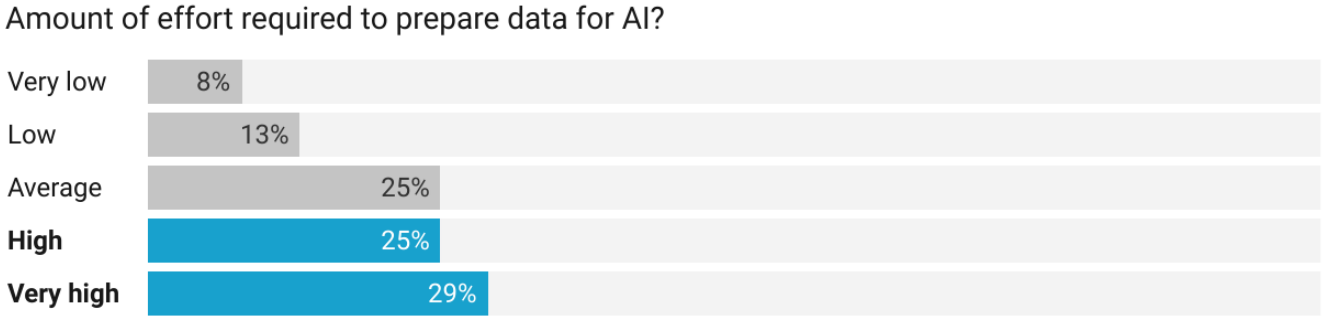












Cross-divisional use cases - Prototype AI projects

Domain-Agnostic Outlier Detection in Science Datasets - *Earth science + Astrophysics + Planetary*

Leveraging AI to Perform Pixel-Level Extraction, Classification, and Segmentation of Astrophysics and Earth Science Imaging Data - *Earth science + Astrophysics*

Petabyte scale search on multi-spectral unlabeled data to rapidly curate annotated datasets + Search By Image for NASA Science - *Earth science + Astrophysics*

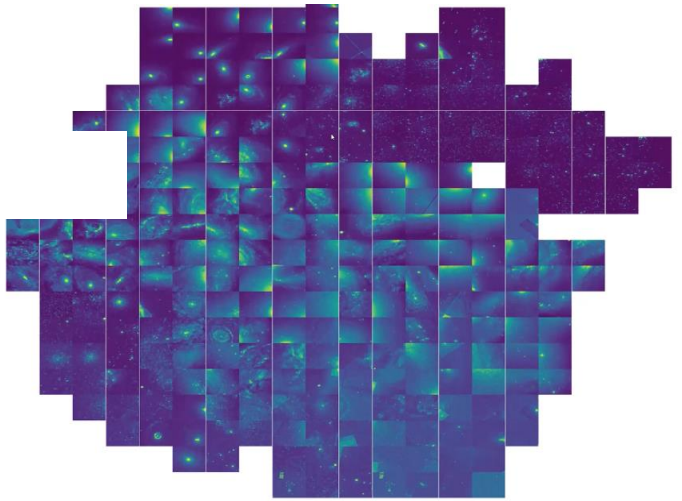
Enhancing NASA's Science using Physics Informed Deep Learning - *Heliophysics + Planetary*

A Bayesian View of the Solar Wind Impact on Mars' Magnetic Environment - *Heliophysics +*

Cross-divisional use cases - Prototype AI projects

Petabyte scale search on multi-spectral unlabeled data to rapidly curate annotated datasets + Search By Image for NASA Science - *Earth science + Astrophysics*

PI: A. Koul & J. Peek

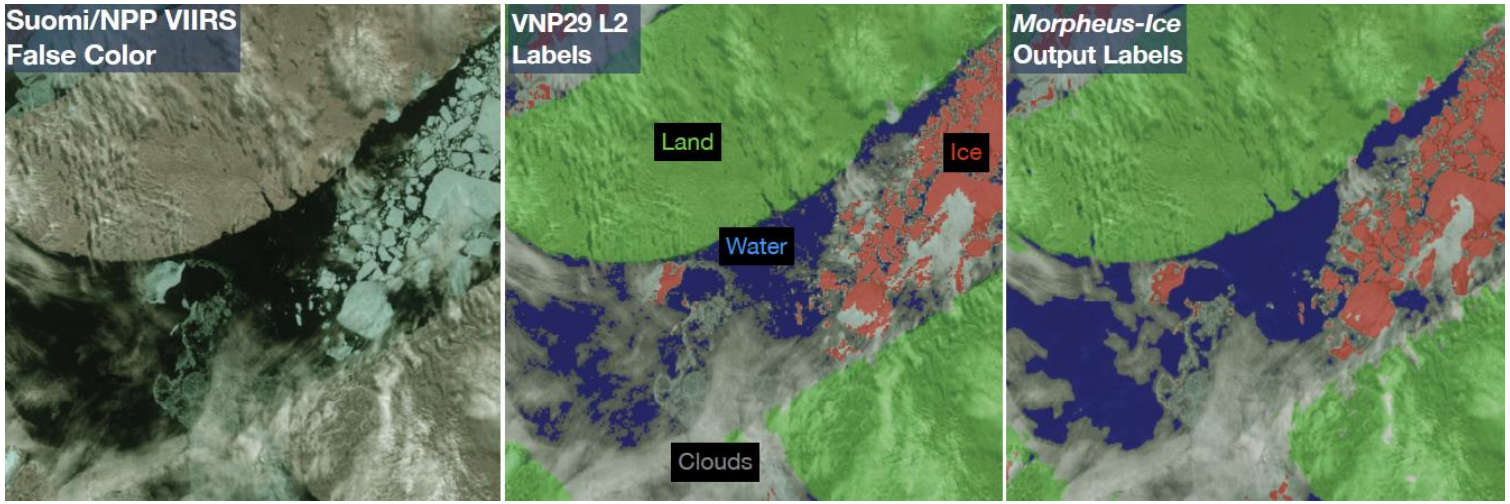


Cross divisional AI project: Self supervised learning approach developed for GIBS archives applied to Hubble telescope data

Cross-divisional use cases - Prototype AI projects

Morpheus-Ice: Pixel-Level Classification of Imaging Data from Astrophysics to Earth Science

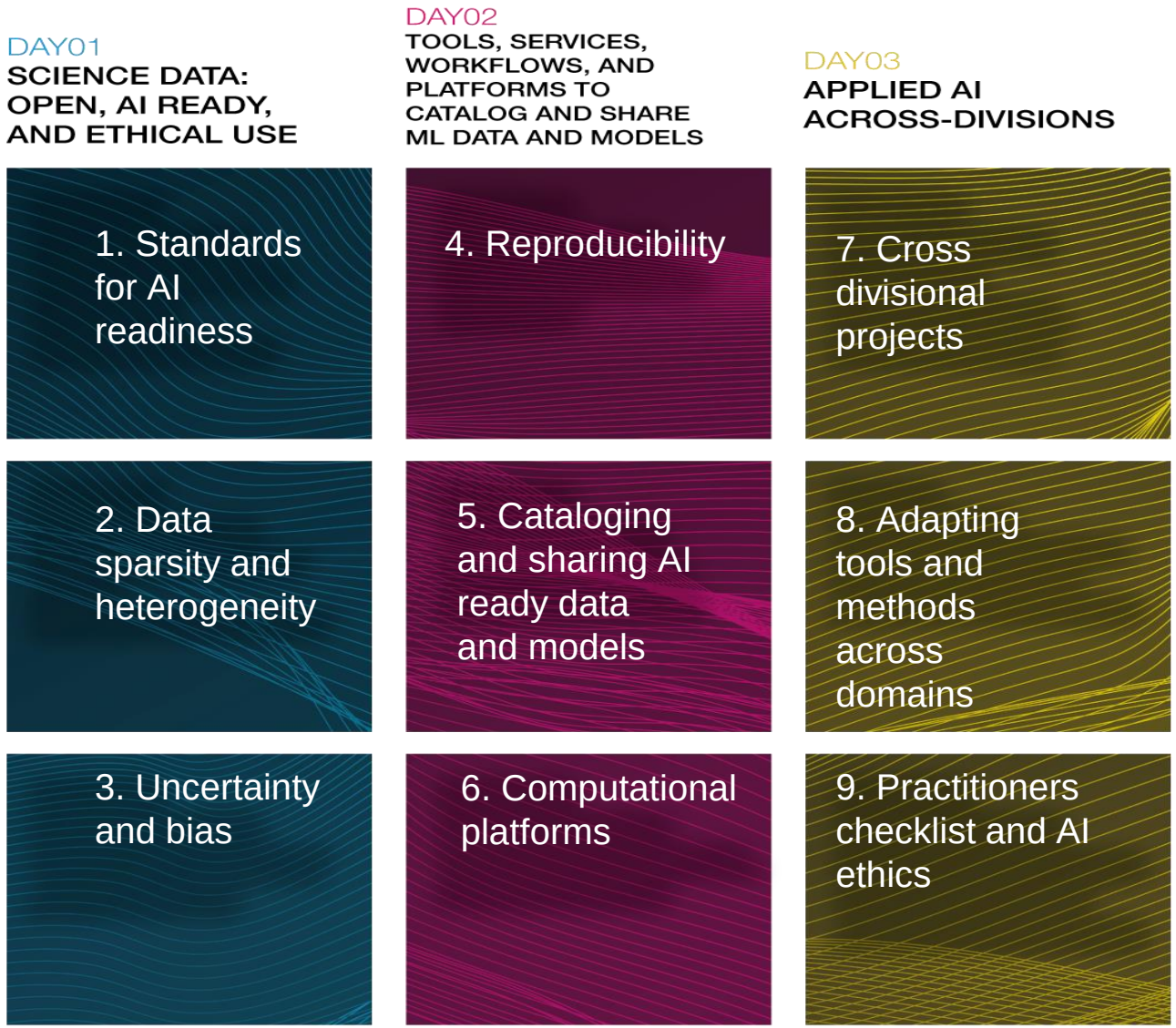
P. B. Robertson



Level 2 data products (VNP29) available from NSIDC that label land, water, clouds, sea ice, night, and bad pixels, created by identifying the relative reflectance of these features at different wavelengths.

Directly supply VIIRS reflectance images and labels to Morpheus, retrain with no architecture changes

SMD AI Workshop



Workshop main findings

Development of SME informed AI-ready data standard

Development of reusable AI-relevant data management tools

Development and publication of labelled training data and models for AI applications and benchmarking

Lowering the barrier to entry to access computing resources for AI

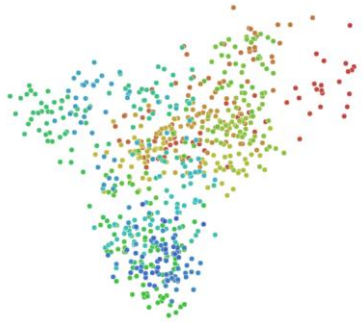
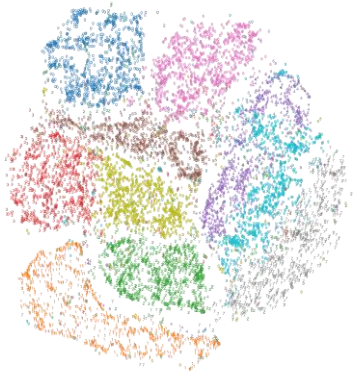
Supporting cross domain collaborations and sharing

Training and education for AI skill development

Incentives for reproducibility and open sharing of AI artifacts

Embedding of ethical considerations of AI into science research processes



Thank you.