Office of the Chief Science Data Officer

NASA's groundbreaking science and exploration missions increasingly rely on the efficient use of large-scale data, advanced computing, and highperformance analytics. The Office of the Chief Science Data Officer (OCSDO) is modernizing data and computing systems for science and engineering across NASA in support of efficiency, sustainability, security, and scientific integrity.

By managing the agency's High-End Computing Capability (HECC) portfolio, the OCSDO provides advanced computational resources for NASA's cutting-edge science and engineering missions. Within the Science Mission Directorate (SMD) OCSDO ensures that the vast amounts of scientific data generated by NASA's missions are accessible and preserved for future analysis. The OCSDO enhances the value of NASA's science data through the evolution of scientific data and computing systems, accelerating discoveries through innovative data science techniques, and fulfilling NASA's strategic goal of ensuring scientific data are readily available to all.

Evolving NASA's Data and Computing Systems

The OCSDO is transforming NASA's computing and data infrastructure to enhance efficiency, interdisciplinary collaboration, and scientific impact. The agency-wide HECC program provides strategic leadership and operational management for NASA's high-end computing capabilities, ensuring optimized computational resources that address complex scientific and engineering challenges.

Within SMD, the OCSDO is implementing a unified cloud infrastructure that enhances interdisciplinary collaboration and simplifies access to computing and data resources. By improving data management strategies, the OCSDO strengthens the discoverability, accessibility, and usability of NASA's



An NVIDIA Grace Hopper superchip installed at the NASA Advanced Supercomputing Facility at NASA's Ames Research Center. (Credit: Don Story/NASA)

scientific archives, maximizing the long-term impact of research investments. These advancements ensure that scientific findings remain accessible, preserved, and ready for future analysis.

To further streamline access to NASA's vast scientific resources, the OCSDO develops enterprise-level data, analysis, and user interface services that complement existing SMD platforms. For example, NASA's Science Discovery Engine (SDE) unifies search capabilities, making scientific data more accessible and discoverable.

Innovating Data Science

The OCSDO advances scientific discovery through innovative applications and partnerships in data science, advanced analytics, and artificial intelligence (Al).



Top data science projects for the OCSDO include robust Al foundation models for science data, providing data discovery tools, establishing Al infrastructure, and facilitating broader scientific collaboration. These tools and initiatives aim to revolutionize how scientists analyze data, build scientific models, and integrate diverse datasets for comprehensive scientific insights.



Burn scar mapping after wildfires in Sikanni Chief, Canada on June 1, 2023, generated by NASA and IBM's open-source Prithvi Geospatial artificial intelligence foundation model. (NASA/OCSDO)

5+1 Strategy on Al

Through strategic partnerships with industry, the OCSDO is creating powerful open-source foundation models that harness AI to process vast datasets quickly, shortening the time between data collection and discovery. The OCSDO is providing the AI expertise, system engineering, infrastructure, and training to facilitate the development of five foundation models, one for each division (Astrophysics, Biological and Physical Science, Earth Science, Heliophysics, and Planetary Science), and one large language model for all divisions of the NASA Science Mission Directorate.

Sharing Scientific Data

NASA is committed to the sharing of scientific data, promoting transparency and efficiency in scientific research. The agency ensures that scientific data, research findings, and new discoveries are accessible to the public without unnecessary barriers. Sharing scientific data is about trustworthiness and accountability. To the maximum extent possible, the raw data collected from NASA missions (like satellite imagery, planetary observations, or climate measurements), as well as the peerreviewed studies and technical insights derived from them, are shared freely online or through public repositories. This empowers citizens, educators, amateur researchers, and innovators to explore, build upon, or even challenge these research findings, making new discoveries of their own in the process.

The OCSDO promotes the public sharing of scientific data and research through training, guidance, and funding. The Open Science 101 curriculum teaches researchers how to work in a collaborative and transparent environment to ensure reproducibility and trust in scientific research. The updated Scientific Information Policy (SPD-41a) ensures that NASA-funded science, including mission data, software, and publications, are publicly shared as rapidly as possible. Additionally, NASA's Research Opportunities in Space and Earth Sciences (ROSES) includes opportunities to support the development and sustainment of existing high-value, opensource tools, frameworks, and libraries with significant impact to SMD.

For more information about the OCSDO and it's programs, projects, and initatives, visit: https://science.nasa.gov/about-us/OCSDO



National Aeronautics and Space Administration Washington, D.C. science.nasa.gov/about-us/ocsdo

www.nasa.gov