

National Aeronautics and Space Administration



EXPLORE SCIENCE

**Request for Information:
Strategic Plan for
Scientific Data and
Computing**

Summary Report

December 2018

www.nasa.gov

All studies, findings and recommendations in these deliverables have been submitted to the Strategic Data Management Working Group (SDMWG) and to officials at NASA HQ. The opinions expressed in these materials do not reflect NASA's concurrence, approval, or indicate steps to implementation.

Executive Summary

The National Aeronautics Space Administration (NASA) leads our nation on a great journey of discovery, seeking out new knowledge and a deeper understanding of planet Earth, our Sun, solar system and the farthest reaches of the universe – back to its earliest moments of existence. NASA's Science Mission Directorate (SMD) and the nation's science community use space observatories to conduct scientific studies of Earth, visit and return samples from other bodies in the solar system and peer out into our Galaxy and beyond.



In the course of its organizational mission, SMD produces, examines and catalogs a large amount of data. The body of data SMD manages has grown rapidly in recent years, and is expected to continue on this growth pattern for the foreseeable future. This creates new opportunities for scientific studies, but also opens the door to significant data management and sharing challenges. SMD's Strategic Data

Management Working Group (SDMWG) is tasked with developing ways to effectively address and overcome these challenges.

In September of 2018, SMD's SDMWG released a Request for Information (RFI) to solicit suggestions in the development of an all-encompassing Strategic Plan for Scientific Data and Computing. To inform the Strategic Plan, SMD chose to gather input from key stakeholders, including but not limited to members of the scientific community, academic institutions, other government agencies, the private sector, professional societies, advocacy groups, the general public and international collaborators. Over the course of 7 weeks, these stakeholders were encouraged to respond and think broadly about the future capabilities and needs of NASA Science's data and computing capabilities. A total of 67 RFI responses were successfully submitted. Across all responses, the most common themes were: open-source tools and data, cloud computing, data management, metadata standardization as well as leveraging machine learning and artificial intelligence. This report will summarize and provide an analysis based on the responses received.

Introduction

SMD depends upon data. Each division within SMD produces, examines and catalogs significant amounts of data to fulfill scientific objectives and provide scientific findings to millions of people. All SMD divisions are predicted to generate more than 100PB of data annually by 2023. The expected increase in the Directorate's archival needs presents opportunities for cutting-edge scientific discovery as well as significant challenges in data analysis, management and access in the future.

SMD aims to answer some of humanity's most profound scientific questions:

How and why are Earth's climate and the environment changing?

How and why does the Sun affect Earth and the rest of the solar system?

How do planets and life originate?

How does the universe work, and what are its origin and destiny?

Are we alone in the universe?

Historically, SMD's management of data and computing resources has been conducted based on the specific needs of each mission or division, which led to a limited consideration for enabling inter-disciplinary research. SMD's prospective Strategic Plan for Scientific Data and Computing will change that trend, and enable data sharing across the Mission Directorate's four Science divisions as well as with other research organizations and the public.

The five-year Strategic Plan has four goals:

1. Improve discovery and access of all SMD data for immediate benefit to science data users and the overall user experience.
2. Identify researchers and use cases that are both large-scale and cross-disciplinary to inform future science data system capabilities.
3. Champion robust theory programs that are firmly based on NASA's observations.
4. Modernize science data and computing systems to improve efficiency and enable new technology and analysis techniques for scientific discovery and commercial use.

SMD released a RFI to the public to gather recommendations from stakeholders on how to develop the Strategic Plan for Scientific Data and Computing. The RFI encouraged respondents from members of all fields of NASA science, data science, engineering, information technology to think broadly about future capabilities and needs. This report summarizes the overall RFI effort and includes an analysis of the RFI responses submitted.

Methodology

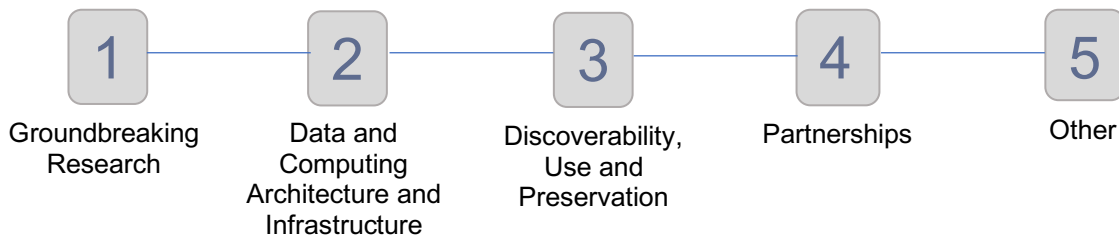
Data Collection

The SMDWG hosted the RFI for the Strategic Plan for Scientific Data and Computing on the NASA Solicitation and Proposal Integrated Review and Evaluation System (NSPIRES). It opened on September 18 and closed on November 2, 2018. The RFI consisted of two parts.

The first part of the RFI collected the following participant information:

- Author name(s)
- Company/Organization Affiliation (or “No Affiliation”)
- Email
- “Summary” that included the author(s)’ Research domain(s), discipline(s) and industry

The second part of the RFI, Program-Specific Data, solicited participants to answer a series of questions organized into the following topical categories:



The full text of the RFI is available in the appendix.

Conducting Analysis

The SMDWG tasked a team from Booz Allen Hamilton (Booz Allen) with conducting the analysis of the RFI responses. The Booz Allen team assessed and analyzed each RFI response in consultation with Booz Allen analysts and members of SMDWG. The team used a variety of documents, spreadsheets, templates and automated tools to organize and analyze the RFI and develop a full report.

The Booz Allen team conducted the analysis of responses to the RFI in three phases:

- Phase One: Collection and Automation
- Phase Two: Assessment
- Phase Three: Summarization

During the collection and automation phase, the team extracted raw data from the 67 valid RFI responses submitted to NSPIRES in a .csv format. The team developed a Python script

which automatically pulled and summarized all the individual responses using a variation on the TextRank algorithm.¹

TextRank attempts to determine the relative importance of different sections of text, in this case individual sentences, within a full document. TextRank searches for and identifies key sentences in a document and similarities between them. The algorithm uses an iterative approach, first analyzing the entire target document, determining which sentences are “important” based on similarity to other sentences, and then repeating the analysis using “importance” scores derived from the previous analysis. After repeated iterations, the “important” sentences are the only sentences that remain, and these can be used to develop a summarization of the original response.

To develop a summarization of the responses using TextRank, the team simply chose an arbitrary reduction in document size, in this case 90%. Thus, the team’s analysis identified the 10% of sentences in the response documents with the highest importance score.

In the Assess phase, the team engaged two Subject Matter Experts (SMEs) to read each response in full. The SMEs split the responses in half and noted recurring topics of discussion across all responses. After comparing notes between the various responses, the team identified the five key themes that appeared the most often. These themes are listed in the *Common Themes in Responses* section of this summary report.

In the third Summarization phase, the SMEs manually summarized the RFI responses using the TextRank summaries as a guide. The SMEs created two separate summaries for each response.

The first set of summaries were those the SMEs developed manually. In these two-page documents, the SMEs made an effort to capture all the issues and recommendations put forth by the respondent. The summary was broken down at the topical category level.

The SMEs then extracted one to three sentences that best summarized the purpose and intent of each response. In addition, the SMEs assigned each response between zero and five of the “main themes” identified in the Assess phase. A main theme was assigned to a response only when the SME determined that it was a significant talking point within the response – e.g., an answer to a question was centered around it, or it was something the author(s) mentioned repeatedly. In most (but not all) cases the authors advocated for the specific approach or tools associated with the main theme(s) that SMEs assigned to their responses.

Quantitative Analysis

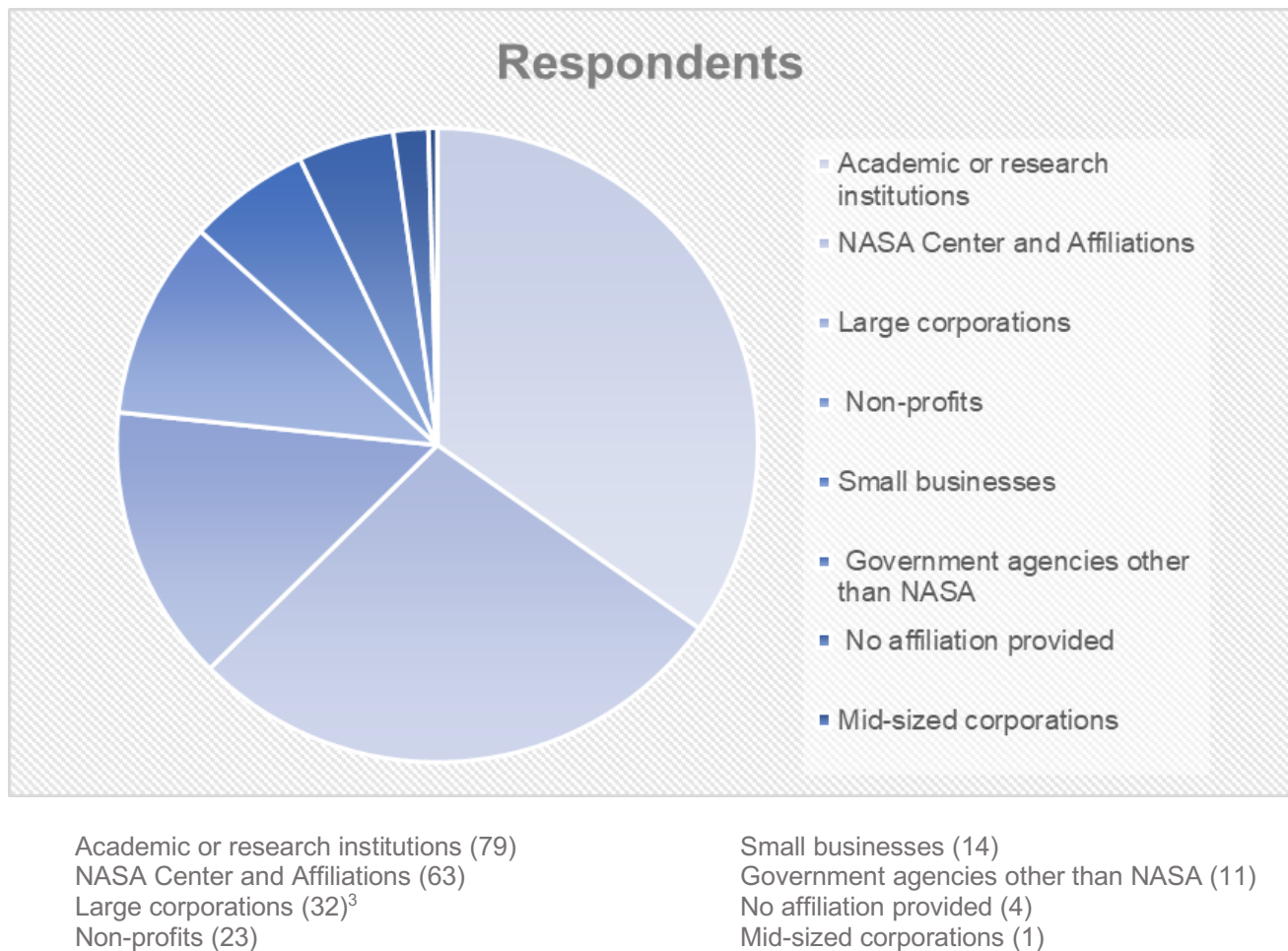
Respondent Demographics

While the RFI was open, it received a total of 68 responses, 67 of which were valid.²

¹ <https://web.eecs.umich.edu/~mihalcea/papers/mihalcea.emnlp04.pdf>

² A total of 68 responses were submitted through the NASA Solicitation and Proposal Integrated Review and Evaluation System (NSPIRES). The first submission, N8-

227 respondents contributed to RFI submissions. According to answers provided in the participant information section of the RFI, the respondents included data scientists, scientific researchers, educators, large and small business leaders and the public. The group affiliations of the respondents (from most common to least common) were:



The RFI respondents have diverse areas of expertise, which is exemplified by the assorted responses summarized in the *Common Themes in Responses* section of this report.

Common Themes in Responses

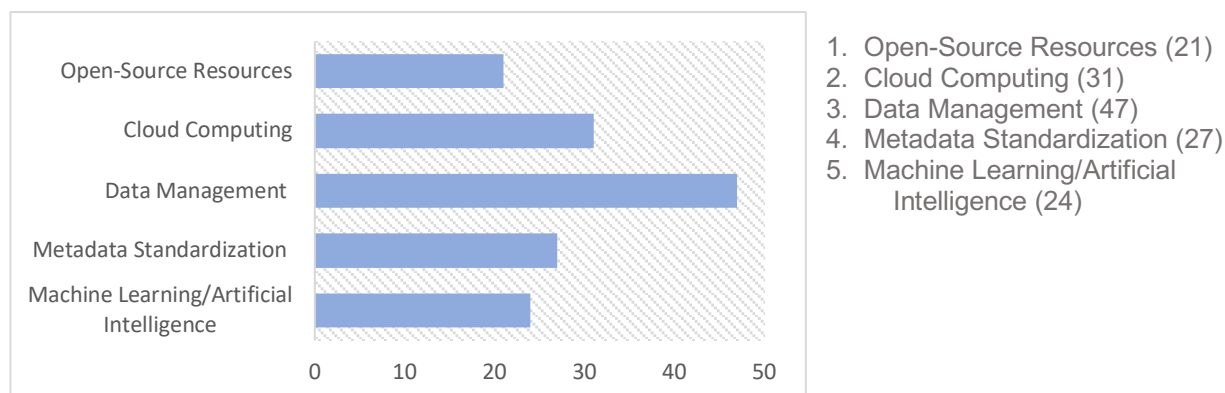
During the Assessment phase of analysis, the analysts noted recurring topics of discussion across responses and selected the five that appeared most often. It is important to note that a main theme(s) being assigned to a response does not necessarily imply that the response was in support of, for example, cloud computing or using open-source tools – although this was the case with most responses. The themes and short descriptions follow:

RFISDC18-0001, was not completed by the principal investigator and was resubmitted under N8-RFISDC18-0002.

³ Company sizes are based on how each company listed by a respondent describes itself on its website.

1. **Open-Source** – As defined by the Open Source Initiative, open source software is software with source code that anyone can inspect, modify and enhance. Responses discussed about open source data or code and open science.
2. **Cloud Computing**⁴ – The National Institute of Standards and Technology (NIST) defines cloud computing as "...a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction." Discussions of whether NASA should have a cloud computing strategy, discussions of using external vs. in-house industry vs. in-house cloud capabilities were amongst the responses.
3. **Data Management** – Data management encompasses the ways in which data is stored and shared by an organization; it is an integral part of research planning as it has the potential to increase the pace of scientific discovery. Discussions of how NASA should arrange, store and provide access to its data, including through cloud computing services.
4. **Metadata Standardization** – Discussions about the level at which metadata should be standardized (SMD divisions vs. SMD as a whole), and which standards should be used that meet the needs of interoperability between independent standardization communities.
5. **Machine Learning (ML)/Artificial Intelligence (AI)** – Machine learning is defined as the field of study that gives computers the ability to learn without being explicitly programmed. Artificial intelligence is defined as machines learning the basic and most critical programs to gain experience and respond to demands and perform human-like tasks. Respondents centered questions around whether and how NASA should be investing furthering ML/AI technology and capabilities, both in terms of developing in-house and supporting external research.

The following list shows the number of reports that addressed each of the five most common themes (with some addressing multiple themes):



⁴ This theme addresses using cloud resources for compute only. Any discussions of using cloud computing for storage were categorized under the Data Management theme.

Qualitative Analysis

The following subsections summarize the RFI responses that pertain to each of the five common themes.

An Emphasis on Open Source

21 of the responses discussed (and almost all strongly recommended) using open-source software and tools in SMD's data management going forward. Responses in this category were almost uniform in their support for keeping NASA data as open and accessible as possible. There was also a large contingent advocating for enhancing the accessibility and usability of data using tools and techniques featured in the other four main themes.

NASA researchers frequently use open-source software to support their work, and many of the submissions that discussed open-source software recommended that the Agency continue supporting this practice. Some argued that NASA should make a concerted effort to transition away from proprietary data analysis tools, like Microsoft Excel, and towards Python, R, and other programming languages that can easily interface with immense amounts of data; and instead use open-source software wherever possible. Leveraging open source technologies has the potential to reduce NASA's risk of "vendor lock-in," i.e., becoming dependent on a single vendor's proprietary software and formats.

Many responses that touched on open-source software pointed out that NASA cannot rely solely on tools developed by other organizations: the Agency will have to develop its own tools for mission-specific data reduction and analysis. Respondents encourage NASA to make any such tools developed in the future available for the wider research community. In a similar vein, some respondents asserted that any software developed through NASA grant-funded efforts should be made freely available to anyone who wants to use them. Ultimately, the minimal costs associated with acquiring and developing open-source tools is their most attractive property according to respondents.

Cloud Computing

31 responses addressed the potential role of cloud computing in managing and making SMD's scientific data more accessible and usable. Nearly half of all submissions in this category recommend supplementing or replacing NASA's current high-performance computing (HPC) capabilities with cloud-based computing. Some respondents in this group advocate for using commercial cloud computing services (e.g., Amazon Web Services, Google Cloud, Microsoft Azure, and NVIDIA GPU Cloud). These respondents note two key benefits of using commercial services:

1. In using a commercial cloud computing service, NASA can pay only for resources it actually uses, rather than paying to build and maintain in-house capacity that it may or may not utilize completely.
2. Many commercial cloud services provide virtual environments equipped with pre-installed popular open-source tools such as Jupyter Notebooks.

Respondents opposed to the use of commercial cloud services note volatile pricing, security threats and the potential, albeit temporary, disruption caused by transitioning cloud computing and storage. Instead, some suggest that NASA should further develop its own internal HPC infrastructure and make that available through the cloud. Most respondents agreed that, while most NASA data is technically freely available, it is difficult to use NASA data because NASA does not currently support large-scale analyses. NASA's current HPC capabilities have been key to advancing science but are becoming oversubscribed. Whether through NASA, a commercial cloud service or a blend of the two, ensuring that the community has access to adequate HPC resources is important to many respondents.

Most supporters of cloud computing also suggest that analytics be performed server-side to mitigate the financial and temporal costs associated with moving large data sets.

Improving NASA Data Management

47 responses addressed the current state of NASA data management, with most suggesting potential improvements. The overarching theme across these suggestions is ensuring that NASA data is rooted in the FAIR principles: findable, accessible, interoperable and reusable (FAIR). According to FAIR data principles⁵, data should be:

1. Easily locatable in a centralized, computer-searchable, SMD-wide data portal.
2. Easily accessible through a user interface or a standard set of application program interfaces.
3. Stored in standardized inter- and intradisciplinary formats.⁶
4. Usable for purposes other than the missions for which the data was intended collected.

The community envisions a single, web-accessible service for discovering and requesting data across SMD's divisions. Cloud-based data storage is highly favored by the community as a means of providing FAIR data, whether it be through commercially available data management platforms, NASA's internal infrastructure or a combination of both.

Metadata Standardization

27 of the responses discussed metadata standardization. One of the most important ways that NASA ensures that collected data can be used to its full potential is by collecting robust and standardized metadata alongside data. Metadata is essential to making both "raw" data and products of data analysis easily accessible and useful. Respondents believe that

⁵ The international Future of Research Communication and e-Scholarship organization, FORCE11, which was founded in 2011, propose that for data to be reusable the infrastructure should follow the Findable, Accessible, Interoperable, and Reusable (FAIR) guiding principles. These principles have been adopted by the American Geophysical Union (AGU) and other partners have just started a project to Enable FAIR Data across the Earth and space sciences.

⁶ This overlaps to some degree with both the "Open-Source Resources" and "Metadata Standardization" themes.

scientific discovery and collaboration between SMD disciplines would increase if interdisciplinary metadata standards were established, promoted, and potentially incentivized by NASA. Proper metadata collection and management could potentially mitigate data siloization, make data more easily machine-searchable and greatly reduce the difficulty associated with creating an SMD-wide data management platform.

There were two main points of contention regarding metadata standards among the respondents. The first is which standards to use for benchmarking purposes. For example, the Planetary Data Science 4 (PDS4) standards were mentioned frequently. The second is whether there should be one NASA-wide unified set of metadata standards, or whether standards should be set at a lower level, perhaps by division.

Machine Learning (ML) and Artificial Intelligence (AI)

24 of the responses addressed ML/AI. NASA is already exploring uses for ML/AI in the Frontier Development Laboratory. Overall, respondents see two general use cases for ML/AI with regards to SMD data.

The first involves using ML techniques in their most common setting: to gain information about data. Examples include performing cluster analysis for mineral classification, nuclear system design optimization and autonomous instrument control. The second use case involves using ML techniques to address the challenges of organizing and managing the ever-growing amount of data being collected. For example, ML can facilitate data management through automated cataloging, cross-referencing and natural language processing of raw data.

In both cases, the ML and AI barrier to entry can be high for researchers. If NASA chooses to pursue increased use of ML and AI tools, they can lower this barrier by educating researchers on cutting edge data science/ML techniques and incorporating these techniques into an end-to-end data analytics environment.

Conclusion and Next Steps

In conducting the analysis of the 67 responses to the RFI, SMD has gained a broad understanding of the advantages and challenges inherent in the current state of NASA Science data. The five key themes gemmed from the RFI responses assist the SMD in building a pathway forward to meet the needs of our missions, but most importantly the needs of our researchers. The next step will be to evaluate these challenges and proposed recommendations and develop a vision and plan for the future state of SMD data management and computing.

Appendix: Request for Information for Inputs to the Science Mission Directorate Strategic Plan for Scientific Data and Computing

Solicitation Number: NNH18ZDA017L

Release Date: September 18, 2018

Response Date: 11:59 PM Eastern Time November 2, 2018

1. Summary:

This Request for Information (RFI) invites comments and suggestions to assist NASA's Science Mission Directorate (SMD) in the development of a new Strategic Plan for Scientific Data and Computing. Over the next five years the plan will be used to guide the evolution of the array of data and computing systems supporting research across four science areas: Astrophysics, Earth Science, Heliophysics and Planetary Science. This notice is published to solicit input from all stakeholders, including but not limited to members of scientific community, academic institutions, other agencies, the private sector, professional societies, advocacy groups, the general public, and international collaborators. Information gathered through this RFI will solely be used for strategic planning purposes and program development.

2. Principles:

As NASA's Science Mission Directorate develops its Strategic Plan for Scientific Data and Computing, the following four principles underlie this work:

1. Continued free and open access to scientific data for any use
2. Improved ease of use and discoverability
3. Enhanced science applications and new use cases
4. Incorporates best practices and "state of the art" through partnerships

3. Background:

Consistent with NASA's 2018 Strategic Plan,⁷ the SMD seeks to expand human knowledge through new scientific discoveries in order to understand the Sun, Earth, Solar System, and Universe. SMD, in partnership with the Nation's science community, conducts scientific studies of the Earth and Sun from space, returns data and samples from other bodies in the Solar System, and peers out into the vast reaches of the Universe. This work seeks to address three core contexts that span the breadth of SMD's activities:⁸

- Discover the secrets of the Universe
- Search for life elsewhere

⁷ https://www.nasa.gov/sites/default/files/atoms/files/nasa_2018_strategic_plan.pdf

⁸ <https://science.nasa.gov/about-us>

- Safeguard and improve life on Earth

SMD organizes its work into four broad scientific pursuits: Astrophysics, Earth Science, Heliophysics, and Planetary Science. Each of these pursuits is managed by a Division within the Directorate and has its own science sub-goals. In addition, SMD manages satellite acquisition on behalf of other federal agencies, most notably the National Oceanic and Atmospheric Administration, through its Joint Agency Satellite Division. Additionally, through the Science Engagement and Partnerships Division, SMD's results are broadly disseminated through a variety of media to learners of all ages.

Each of the four Science Divisions within SMD generates, analyzes, and archives large amounts of data to support unique science objectives and delivers data and scientific results to millions of users around the world. In the past, management of data and computing resources has been conducted based on the specific needs of each mission or Division, with limited consideration for enabling inter-disciplinary research.

Currently SMD stores over 100 Petabytes (PB) of observational and model data. For example, the Earth Science Division archives contain nearly 30PB of observational data and more than 70PB of model output. By 2023, the Earth Science archives will contain over 150PB of observational data with an annual growth rate of nearly 50PB per year.⁹ Within 5 years, nearly 100PB of data will be collected by instruments or generated by models each year across SMD. This anticipated growth of SMD's science data presents unique opportunities for new scientific discovery as well as significant challenges for data management, curation, access, analysis, maintenance of provenance and computing.

In addition to its archives, SMD supports two major supercomputing centers which provide more than 10 Petaflops peak capacity for the Agency's scientific research and engineering workloads.¹⁰ These centers mainly support research in areas of Astrophysics and Heliophysics theory development and validation, weather and climate modeling and data assimilations, and large-scale data synthesis and analysis.

4. SMD Data and Computing Systems:

Data and computing systems supporting the Science Mission Directorate have embraced a system-of-systems approach. These loosely coupled systems have allowed each Science Division flexibility to support unique science requirements and user needs of the communities. Please visit the following URLs for more information.

- Astrophysics Data System – <https://science.nasa.gov/astrophysics/astrophysics-datacenters>
- Earth Observation Data and Information System – <https://earthdata.nasa.gov>
- Heliophysics Data Environment – <https://hpde.gsfc.nasa.gov/index.html>
- High-End Computing – <https://www.hec.nasa.gov/>
- Planetary Data System – <https://pds.nasa.gov/>

⁹ <https://earthdata.nasa.gov/about/eosdis-cloud-evolution>

¹⁰ <https://www.nas.nasa.gov/hecc/resources/environment.html>

5. Development of a Strategic Plan for SMD Data and Computing:

Responses to this RFI will inform development of a Strategic Plan for Scientific Data and Computing for NASA's Science Mission Directorate. The objective of the Strategic Plan is to articulate a whole-of-SMD five-year strategy that has four overall goals:

- Improve discovery and access for all SMD data for immediate benefit to science data
- users and to improve the overall user experience
- Identify large-scale and cross-disciplinary/division science users and use-cases to inform future science data system capabilities
- Enable strong theory programs that are firmly based on NASA's observations
- Modernize science data and computing systems to improve efficiency and enable new technology and analysis techniques for scientific discovery and commercial use

6. How to respond to this RFI:

The Strategic Plan for Scientific Data and Computing is intended to span the range of activities within the Science Mission Directorate, including opportunities for cross-disciplinary science investigations. Respondents are therefore encouraged to think broadly about future capabilities and needs, and we encourage members of all fields of NASA science, engineering, industry, and academia to respond to this RFI in order to ensure a range of views.

NASA may use your response to aid in programmatic decisions about future investments for data and computing, including partnerships with private sector and philanthropic organizations. This RFI is not to be construed as a commitment by the Government nor will the Government pay for information solicited. No proposals will be awarded funding as a result of this RFI.

Responses will be accepted only if submitted as a Notice of Intent (NOI) via NSPIRES at <https://nspires.nasaprs.com/external>; see Section 9 for instructions. Responses provided by email, mail or other means will **not** be accepted.

Each response to this RFI should provide the following information:

- I. Cover Page that includes:
 - a. Title of the response
 - b. Author name(s)
 - c. Company/Organization or specify "No Affiliation"
 - d. Email
 - e. "Summary" that includes the author(s)' Research domain(s), discipline(s), industry and summary of the response ('abstract'; 1500 characters maximum)
- II. Answers to the following questions in the Program Specific Data section of the NOI:
 1. *Groundbreaking Research:*

- a. What large-scale and inter-disciplinary/divisional science investigations could be enabled by advanced data and computing capabilities? (4000 characters maximum)
 - b. Describe pressing research questions within or among SMD-supported science areas, the current state-of-the-art and how new data and computing capability could lead to groundbreaking science. (4000 characters maximum)
2. *Data and Computing Architecture and Infrastructure:*
 - a. Describe any limitations, including absence of existing science data tools or computational capabilities, which must be addressed to: (4000 characters maximum)
 - b. Support groundbreaking research
 - c. Support fundamentally new approaches to science (including machine learning, deep learning and artificial intelligence)
 - d. Manage and provide open access to nearly an exabyte of data by 2025.
 - e. How does SMD data and computing architecture and infrastructure need to evolve to support groundbreaking science? (4000 characters maximum)
3. *Discoverability, Use and Preservation:*
 - a. How can SMD harness advances in data science and modern computational platforms to provide seamless search, discovery and use of data and computing across Divisions? (4000 characters maximum)
 - b. Describe standards, access methods, analysis tools and visualization capabilities that will increase the use of NASA science data by researchers, the public and commercial users. (4000 characters maximum)
 - c. Describe use cases and how to improve the user experience and user interfaces, including for citizen scientists. (4000 characters maximum)
 - d. How can data from multiple locations/sensors/sources be identified as it is collected? As data is analyzed, how can data be identified for repeatable science investigations? (4000 characters maximum)
 - e. How can SMD continue to provide free and open access to scientific data and maximize availability and discoverability while also maintaining data integrity? How can SMD prevent unauthorized usage, disruption, modification or destruction? (4000 characters maximum)
4. *Partnerships:*
 - a. How can SMD partner with academic, nonprofit, commercial and international organizations to improve access to, analysis of and long-term preservation of scientific data? (4000 characters maximum)
 - b. How can SMD partner with existing academic and commercial high performance computing networks to improve large scale analysis of data and models? (4000 characters maximum)
5. *Other: (4000 characters maximum)*
 - a. What else should SMD consider with regards to data, computing, storage, and networking, software, algorithm release, software portability, development practices/approaches, policies and open data, etc. that was not previously addressed? (4000 characters maximum)

7. Submission Information:

Due **November 2** by 11:59 PM Eastern Time

Submissions will be used to inform development of the Science Mission Directorate's Strategic Plan for Scientific Data and Computing. As such, please think broadly about responses to ensure that they are in alignment with the cross-SMD nature of this activity. Information gathered through this RFI will be solely used for strategic planning purposes and program development. Respondents shall not provide any information that is considered proprietary and/or confidential as all information collected through this RFI will be used to develop a public plan.

It is emphasized that this RFI is neither a Request for Proposal, nor an Invitation for Bid. This RFI seeks information for planning purposes only; therefore, NASA does not plan to respond to the individual RFI responses. As stipulated in FAR 15.201(e), responses to this notice are not considered offers and cannot be accepted by the Government to form a binding contract. Pursuant to FAR 52.215-3, entitled Request for Information or Solicitation for Planning Purposes, this information is being made available for market research, information, and planning purposes and to allow potential proposers the opportunity to verify reasonableness and feasibility of the requirements, as well as promote competition. This RFI is subject to review or cancellation at any time and is not to be construed as a commitment by the Government to enter into a contract. The Government will not pay for the information submitted in response to this request, and respondents will not be notified of the results.

Please do not request a copy of the solicitation, as no solicitation exists at this time. If a firm requirement is developed and a solicitation is issued, the solicitation will be made available through NSPIRES (<https://nspires.nasaprs.com/external/>). It is the responsibility of Offerors and interested parties to monitor the internet sites for the release of the solicitation and amendments, if any, and they will be responsible for downloading their own copy of the documents. NASA Clause 1852.215-84, Ombudsman, is applicable. The Center Ombudsman for potential acquisitions can be found at http://prod.nais.nasa.gov/pub/pub_library/Omb.html.

8. Contact:

Questions concerning this RFI may be addressed to Kevin Murphy, Science Mission Directorate, NASA, Washington, DC 20546; Email: hq-smd-data-computing-rfi@mail.nasa.gov and Ellen Gertsen, Science Mission Directorate, NASA, Washington, DC 20546; Email: hq-smd-datacomputing-rfi@mail.nasa.gov.

9. NSPIRES Instructions:

All submissions will be in the form of a Notice of Intent (NOI) submitted via NSPIRES. The individual submitting the NOI must have an NSPIRES account, but does not require an affiliation with an NSPIRES organization. See registration instructions at <http://nspires.nasaprs.com> (select "Getting Started"). Once submitted, an NOI cannot be edited or deleted. Any changes should be made by submitting a "revised" NOI as a second submission.

1. Log in to your account at <http://nspires.nasaprs.com/>.
2. Select "Proposals/NOIs" from your account page.
3. Select "Create NOI" from your proposals page.
4. Click "Continue" on the next page.
5. Select "Request for Information: Inputs to SMD Strategic Plan for Scientific Data and
6. Computing (NNH18ZDA017L)" from the bulleted list of announcements.
7. Click "Continue".
8. Enter RFI response title ("NOI title" field will be shown).
9. Select "do not link at this time" for submitting organization page or choose an organization to link this to. Note that once a link it made it cannot be changed.
10. Click "Save" on next page.
11. It is not necessary to complete the following "NOI Details" sections: Business Data;
12. Budget; Team members. HOWEVER, it is imperative that you complete the "Summary" (this is the abstract mentioned in "I.e." above) and answer all "Program Specific Data" questions. These answers make up your RFI response and reflect the content described in Section 6 above.
13. Click Submit NOI button. NOTE that this does not complete the submission process.
14. Ignore any warnings about incomplete NOI elements. Click "Continue".
15. Click "Submit". This will take you to the NOI submission confirmation page, which provides you with the NOI/RFI number for your records.

Please note: You may delete and change the answers to the Program Specific Data questions any time before the submission deadline, however once your RFI is submitted, it cannot be deleted.

NASA

A vertical strip of cosmic imagery. At the top, there are blue and purple nebulae and galaxies. A bright comet streaks across the upper part. Below that, the planet Saturn is visible with its rings. Further down is a smaller planet. At the bottom of the strip is a large, glowing orange sun. A black silhouette of a person with arms raised stands in front of the sun. The entire scene is set against a dark background with faint vertical lines.

EXPLORE
with us