NASA SMD Information Policy

Steve Crawford
SMD Science Data Officer

Strategy for Data and Computing for Groundbreaking Science 2019-2024

Vision: To enable transformational open science through continuous evolution of science data and computing systems for NASA's Science Mission Directorate.

Mission: Lead an innovative and sustainable program supporting NASA's unique science missions with academic, international and commercial partners to enable groundbreaking discoveries with open science data. Continually evolve systems to ensure they are usable and support the latest analysis techniques while protecting scientific integrity.

Goal 1: Develop and Implement Capabilities to Enable Open Science	**Goal 2**: Continuous Evolution of Data and Computing Systems	**Goal 3:** Harness the Community and Strategic Partnerships for Innovation
Strategy 1.1: Develop and implement a consistent open data and software policy tailored for SMD	**Strategy 2.1**: Establish standardized approaches for all new missions and sponsored research that encourage the adoption of advanced techniques	**Strategy 3.1:** Develop community of practice and standards group
Strategy 1.2: Upgrade capabilities at existing archives to support machine readable data access using open formats and data services	**Strategy 2.2**: Integrate investment decisions in High-End Computing with the strategic needs of the research communities	**Strategy 3.2:** Partner with academic, commercial, governmental and international organizations
Strategy 1.3: Develop and implement a SMD data catalog to support discovery and access to complex scientific data across divisions	**Strategy 2.3**: Invest in capabilities to use commercial cloud environments for open science	**Strategy 3.3:** Promote opportunities for continuous learning as the field evolves through collaboration
Strategy 1.4: Increase transparency into how science data are being used through a free and open unified journal server	**Strategy 2.4**: Invest in the tools and training necessary to enable breakthrough science through application of AI/ML	

Strategy 1.1: Develop and implement a consistent open data and software policy tailored for SMD

Develop and implement a consistent open data and software policy tailored for SMD building on Agency and U.S. Government guidance. This policy should be informed by the National Academies of Science, Engineering, and Medicine's 2018 report *Open Source Software Policy Options for NASA Earth and Space Sciences*. This SMD-tailored policy is the framework for the open science ecosystem and will be used to inform science data and software policies for system development.

Strategy for Data and Computing for Groundbreaking Science 2019-2024

Core Values in the Development of the policy

- Maximize the openness of useful scientific information produced by NASA SMD funding
- Preserve the scientific information produced by NASA SMD funding
- Support the community in accessing the information that they want
- Minimize the burden in complying with the policy

Why an information policy?

The information produced as part of NASA's scientific research activities represents a significant public investment. NASA holds this information as a public trust to increase knowledge and serve the public good.

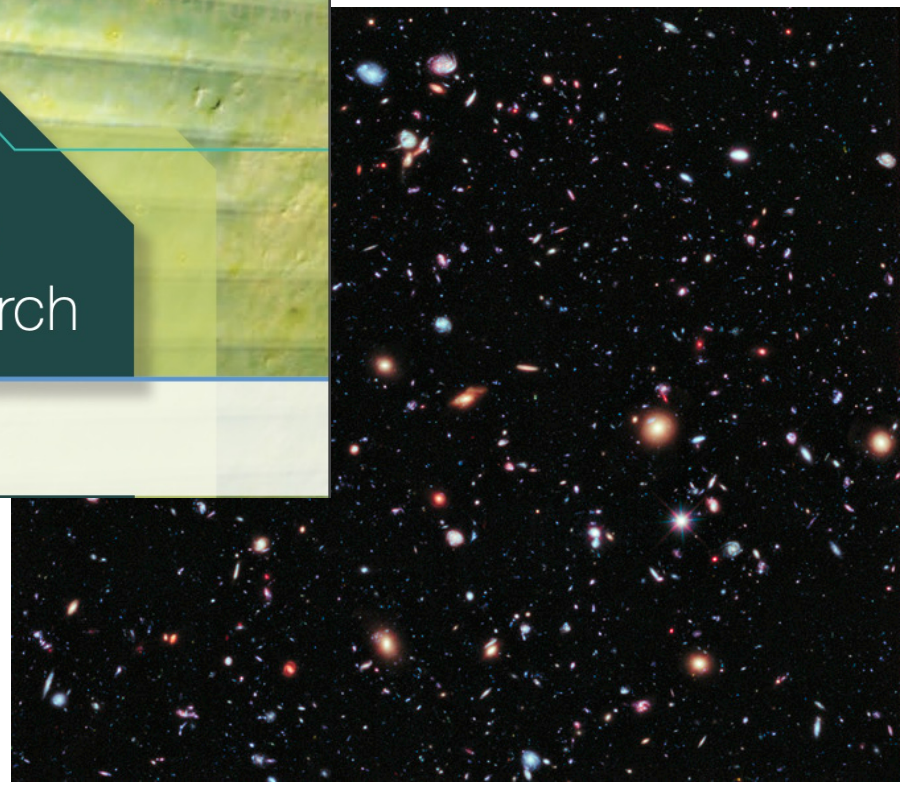
Types of information

- Publications (peer reviewed, conference presentations, technical reports)
- Data (scientific and engineering)
- Software (scientific and operations)

Benefits of an information policy:

- Need to make sure that information produced by NASA SMD funding is shared openly and widely to maximize benefit and reach of the of the information
- Policy also helps consolidate a wide range of applicable laws, guidance, and policies
- Provides clear and transparent policy for our program officers, scientists, and communities.
- Helps support for open science

NASA's commitment to Open Science



NASA has a long history of making its data publicly available and in 2015, released the Plan for Increasing Access to the Results of Scientific Research based on Federal Guidance.

Current Policies on Information

Information Policy

There is a wide range of policies, federal directives, and laws applicable to the information that SMD produces. As such, we are attempting to consolidate this information into a single policy to help clarify what information needs to be made available.

This policy will include:

- Applicability: Explaining the different types of activities to which the policy applies. Unless prevented by other laws or requirements, this policy applies to all SMD funded scientific activities.
- General: Policies that apply to all applicable activities.
- Missions: Policies or deviations applicable to information produced by our Strategic or Flagship missions.
- Research: Policies or deviations applicable to information produced by our Research Activities.
- Definitions of common terms to be used across SMD. For example, this will include a common definition for levels of data.

The policy is currently going through final approvals to be made publicly available. The initial version of the policy is focused on current or applicable policy to SMD and NASA.

Applicable Information

Publications

Scientific and technical documents released through print, electronic, or alternative media.

- This includes peer reviewed manuscripts, technical reports, conference materials, and books.
- This does not include laboratory notebooks, preliminary analyses, drafts of scientific papers or preprints, plans for future research, peer review reports, or communications with colleagues.

Data

Scientific information that can be stored digitally and accessed electronically.

- Information produced by missions includes observations, calibrations, coefficients, documentation, algorithms, and any ancillary information.
- Information needed to validate the scientific conclusions of peer-reviewed publications. This includes data underlying figures, maps, and tables.
- This does not include laboratory notebooks, preliminary analyses, drafts of scientific papers, plans for future research, peer review reports, communications with colleagues, or physical objects, such as laboratory specimens.

Software

Computer programs in both source and object code that provide users some degree of scientific utility or produce a scientific result or service.

Highlights from the General Policy

Current policies include the following:

- SMD-funded publications shall be made publicly accessible.
- SMD-funded data shall be made publicly available without fee or restriction of use.
- SMD-funded software should be released as open-source software.
- All SMD-funded activities shall have data management plans describing the management and release of data to facilitate the implementation of these information policies The DMP should include a description of the software to be used and how it will be managed.

Mission Specific Highlights to the Policy

- SMD shall commit to the full and open sharing of information produced by NASA SMD Missions. This includes observations, calibrations, coefficients, documentation, software, algorithms, technical reports, and any ancillary information or work product related to the Mission.
- There shall be no period of exclusive access to Mission data. A period after the data have been obtained may be allowed for activities such as calibration and validation of the data. This period shall be as short as practical and shall not exceed six months.

Research Specific Highlights to the Policy

- Research data shall become publicly available no later than the publication of the peer-reviewed article that describes it.
  - Information needed to validate the scientific conclusions of peer-reviewed publications that resulted from SMD funding shall be made publicly available at time of publication. This includes the data underlying figures, maps, and tables.
- In order to achieve reproducibility, research software developed using SMD funding and used in support of a scientific, peer-reviewed publication should be released as open source software no later than the publication date.

Proposed Additions to the Information Policy

New Guidance and Additional Policies

Since 2015 there have been a range of new and proposed laws, recommendations, policies and Federal Guidance related to Open Science.

TITLE II—OPEN GOVERNMENT DATA ACT

SEC. 201. SHORT TITLE.

This title may be cited as the "Open, Public, Electronic, and Necessary Government Data Act" or the "OPEN Government Data Act".

Science Mission Directorate's
Strategy for Data Management and Computing for Groundbreaking Science 2019-2024

Prepared by the Strategic Data Management Working Group

Goals of Proposed Additions

- Making SMD data FAIR.
  - This means data should be findable, accessible, interoperable, and reusable (FAIR).
- Maximizing the openness of our software both for missions and research
- Using Persistent identifiers throughout our processing including for funding mechanisms and missions, investigators, and data.
- Recognizing that the value of peer reviewed data and software can be commensurate with peer reviewed manuscripts.
- Prioritizing automation and streamlining processes over compliance, but making sure accountability is included as well.

Planetary Data Ecosystem (PDE)

Many of the IRB recommendations are aligned with the principles of the information policy

R05 [H]: NASA should expand intra- and inter-agency efforts to ensure that best practices, lessons learned, and appropriate technologies are shared and implemented across the PDE

R10 [M]: NASA should prioritize the reuse of data and metadata standards, data format conversion tools, and APIs across other organizations rather than inventing new ones.

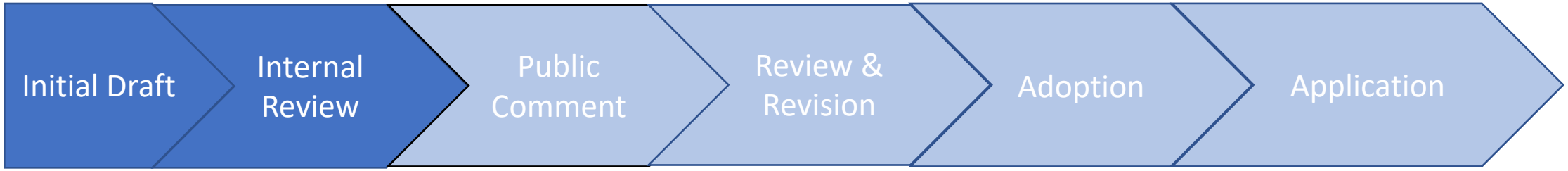
R11 [H]: The PDE should regularly assess FAIR across each PDE element for machine-actionable access to data. This assessment should be used to establish the priorities for Ecosystem management and advisory groups.

R16 [M]: Create a shared, common taxonomy, controlled vocabulary, high level data dictionary, and/or glossary of terms across the PDE. This will substantially advance the machine-actionability of Ecosystem data, and specifically improve interoperability and reusability...

R63 [M]: PSD should adopt a single, coherent, open source software policy across all its activities. Ideally, this policy should be a consistent Science Mission Directorate policy.

Next Step for the Information Policy

Schedule for the policy development



The schedule though is consensus driven and these dates should be viewed as targets and not deadlines. The schedule will shift to give appropriate time for discussion and comment.

- Approval of the proposed additions will occur no earlier than August 2021.
- If adopted, policy will be in place for ROSES and AO calls in 2022.
  - Some requirements already apply.
  - Some solicitations may adopt parts of the policy earlier.
  - Existing Missions and Researchers should adopt the policy as their resources allow.
- Except for variances, researchers funded from ROSES22 will need to be compliant.
  - Funded projects will typically start in 2023.
  - Funded publications will typically be produced in 2024.

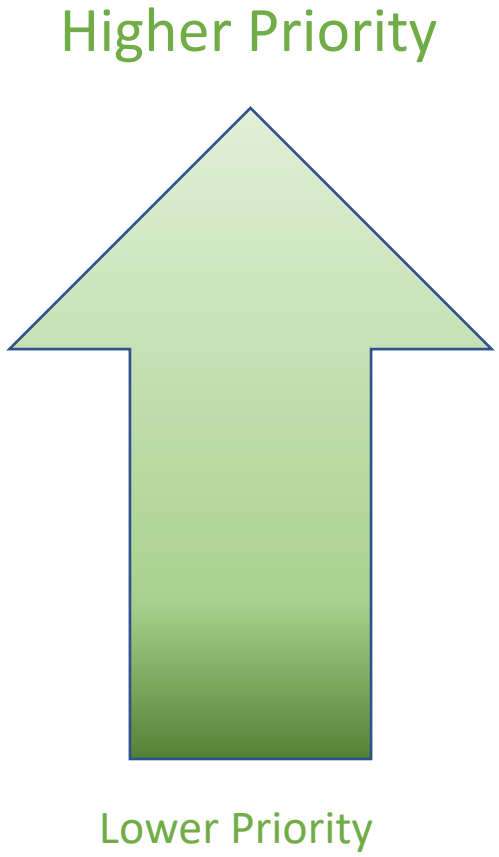
Core Values in the Development of the policy

- Maximize the openness of useful scientific information produced by NASA SMD funding
- Preserve the scientific information produced by NASA SMD funding
- Support the community in accessing the information that they want
- Minimize the burden in complying with the policy

Next steps on the Information Policy

The development of the policy is only an early step in the overall process that will take place over the next 5 years.

1. Identify ways that compliance can be further streamlined and automated both for our community and program officers.
2. Provide further guidance on compliance and what it means. This includes discussions with the community and program officers to identify their needs. Each division will be providing its own information policy.
3. Provide templates and examples of compliance with the policy.
4. Provide training on how to take advantage of existing technologies to make compliance easy.
5. Support and prioritize the development of technologies that can ease and automate compliance. This can include support for data management plans, data curation, and indexing of publications.
6. Identify metrics and technology for assessing compliance.



PubSpace and a New External Portal

RESEARCH ACCESS

What's Happening?

Submission Portal: A new submission portal is being developed for NASA-funded external users (grantees/coop agreements holders/contractors) for Accepted Manuscripts and other Scientific and Technical Information (STI) products.

- Those who currently submit through National Institutes of Health Manuscript Submission (NIHMS) System will submit directly to the STI Program Office
- Login options for the portal will authentic via NASA Launchpad, ORCID, and/or guest.nasa.gov
- The new portal will leverage Persistent Identifiers (PIDs) and funding information
- Target completion date is August 2021

PubSpace: A new PubSpace environment will be created in the STI Repository

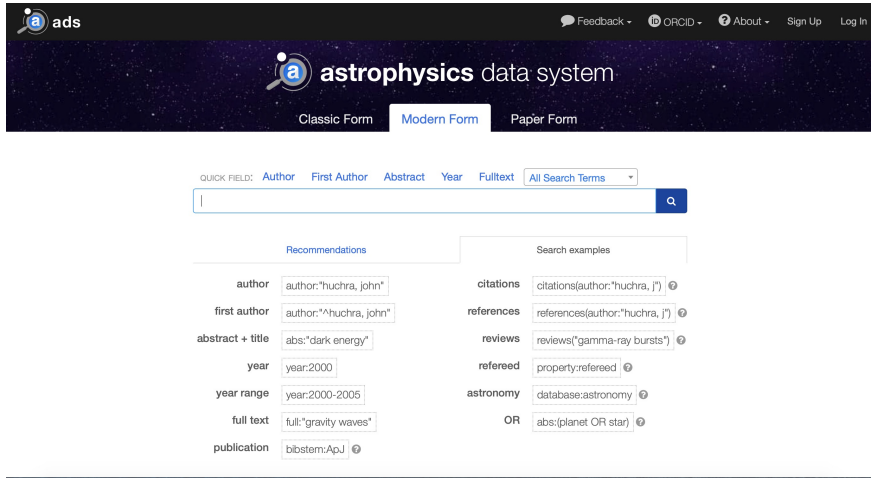
- Currently PubSpace is a subcollection in PubMed Central (PMC)
- Hosting PubSpace in the STI Repository provides greater opportunity to capture content from multiple sources through the external submission portal, legacy STI content, and PMC. Investigating further integration with CHORUS and other Open Access sources as well as journals.

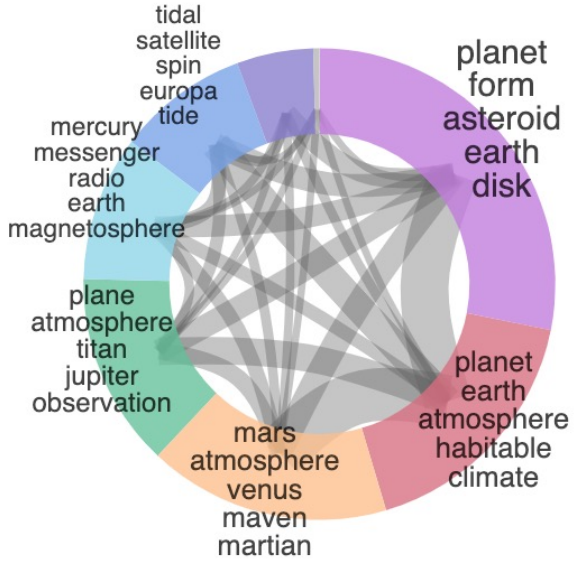
More information and updates will be provided at https://sti.nasa.gov/submit-to-pubspace

NASA Astrophysics Data Systems extending support for Planetary and Heliophysics

As part of the overall *Strategy for Data Management and Computing for Ground Breaking Science 2019-2024* and aligned with the policy development, NASA ADS is looking to expand its holdings and integrations in the Planetary and Heliophysics.

It already provides indexing and search for many of the publications in these fields, but it will provide greater indexing and search for a wider range of materials.





Public Comment on the Policy

The policies will be release soon. With their release, there will be an RFI to requesting information on:

- How will the proposed changes to the existing information policy impact the research activities of your communities?
- What support, services, training, funding, or further guidance is needed to support the successful implementation of the existing or proposed information policy?

Back up material

Development of the Policy

There are already a significant number of NASA policies, government directives, and laws that govern the availability of the information that NASA produces.

- Federal laws such as the Open Data Act and America Competes Act
- Federal Guidance on open data, open source software, and publications
- NASA policies
- National Academies of Science, Engineering, and Medicine studies
- Recommendations and requirements from professional societies and journals
- Community best practices

This policy tries to consolidate this information and provide direction that is relevant to NASA SMD.

Why support Open Science?

Here are some reasons to support open science:

- Helps produce science that is more reproducible
- Increases transparency to the scientific process
- Increases the accessibility by lowering barriers through open sharing of knowledge
- Increase efficiency: the reuse of data and software allows scientists to focus on the next step rather than building the same tool
- Enables new science that was not originally envisioned
- Improves the quality of the science
- Compliance with Federal guidance