# Open Science and Data Management Plan (OSDMP) Guidance

Proposers are recommended to review the guidance below (blue text, red comments), save a copy, then delete the guidance text. Template contents by Marshall Styczinski and Daniel Crichton. Please direct questions about this template to Daniel Crichton at [daniel.j.crichton@jpl.nasa.gov](mailto:daniel.j.crichton@jpl.nasa.gov).

Proposals to most ROSES-2023 program elements must include an “Open Science and Data Management Plan” (OSDMP), formerly known as the Data Management Plan, as part of the proposal document. **Always refer to the** [***ROSES Summary of Solicitation***](https://nspires.nasaprs.com/external/viewrepositorydocument?cmdocumentid=919302)**(Section II.c)*,* the** [**OSDMP FAQ page**](https://science.nasa.gov/researchers/sara/faqs/OSDMP)**,and the** [**Science Mission Directorate Open-Source Science Guidance**](https://science.nasa.gov/oss-guidance) **for the most up-to-date, official information, including new requirements for ROSES-23.**

For DAPR program elements, proposers are required to write the Open Science and Data Management Plan section of the proposal document **in an anonymized format** that does not explicitly identify the names of the team members or their institutions. Refer to the program element for any specific OSDMP requirements. General DAPR guidance is provided here: <https://science.nasa.gov/researchers/dual-anonymous-peer-review>.

From the ROSES Summary of Solicitation: “New in ROSES-2023: The requirements regarding archiving of data, software, and publications have been strengthened. In particular: 1) As-accepted manuscript versions of publications that derive from ROSES-2023 awards must be publicly available at the time of publication; 2) Data and software developed using ROSES funding in support of a peer-reviewed publication shall be made publicly available at the time of publication; 3) Scientifically useful data and software developed during the award that was not already published must be made publicly available by the end of the award; and 4) To be eligible to receive funding, PIs and Co-Is must provide their digital persistent identifier (e.g., ORCID) via NSPIRES under Account Management -> Personal Profile.” **The actions your team will take to meet all of these requirements, including specific roles and responsibilities, must be detailed in the OSDMP. Be sure to include time and costs associated with these steps in your work plan.**

*Acceptable Data Repositories*

**Note that “repository” under current policy refers to long-term, immutable storage of information. This is IN CONTRAST with version-control systems like GitHub.**

SMD policy currently requires that repositories used for SMD information have the following properties:

* Make information findable and accessible to the public without fee or restriction of use;
* Compliant with standards for accessibility for all electronic and information technology to people with disabilities;
* Compliant with a principle of non-discriminatory data access so that all users will be treated equally (any variation in accessibility will result solely from the capability, equipment, and connectivity of the user).
* PSD proposals must select an appropriate repository from among those listed at <https://science.nasa.gov/solar-system/archives>.

Additionally, repositories should be capable of maintaining the information for an extended period (e.g., SPD-41a suggests 25 years) and follow [FAIR Guiding Principles](https://www.go-fair.org/go-fair-initiative/).

In a case where no appropriate archive exists for a particular data set, the OSDMP should discuss alternative methods for making the data publicly available.

Proposers considering archiving their data in the Planetary Data System (PDS) should keep in mind that data archived in the PDS are intended for long-term preservation and must be in [PDS4 formats](https://pds.nasa.gov/datastandards/about/) as well as undergo a separate peer review process. Proposers intending to archive data in the PDS must obtain a letter of support from the appropriate PDS Discipline Node confirming that the PDS is willing to accept their submission. This letter must be included in the proposal package and placed in a section for Statements of Commitment and Letters of Support, Feasibility, and Endorsement (see [*Table 1 of the ROSES-2023 Summary of Solicitation*](https://nspires.nasaprs.com/external/viewrepositorydocument?cmdocumentid=913341&solicitationId=%7b274C8365-A038-339F-A3AE-8F5BFE178312%7d&viewSolicitationDocument=1)) Letters must be requested at least one week prior to the program element due date and proposers are strongly encouraged to initiate conversations with the relevant PDS Discipline Node several weeks ahead. See the [Information for Data Proposers](https://pds.nasa.gov/home/proposers) and [Guidelines for Archiving](https://pds.nasa.gov/home/proposers/archiving/) sections of the [PDS website](https://pds.nasa.gov/) for more information.

Some program elements require that the open science and data management plan be integrated in the page-limited S/T/M section of the proposal instead of being placed in a separate 2-page section. Please read the program element carefully to determine: if an OSDMP is required; what elements are required in the OSDMP; where the OSDMP should be placed; and OSDMP page limits.

**Delete this section if the program element requires the open science and data management plan discussion to appear in the S/T/M section of the proposal instead of here in a standalone section.**

Please read the program element appendix that you are responding to carefully. Depending on the call, there may be other data management plan requirements that will need to be addressed.

# Open Science and Data Management Plan (OSDMP)

Blue text is instructional and is intended to aid in writing your OSDMP. Read blue text and delete before submission. Black text is an example case—edit the black text to fit your proposed project. For most program elements, the OSDMP is limited to 2 pages. For questions on this OSDMP template, contact Daniel Crichton at [daniel.j.crichton@jpl.nasa.gov](mailto:daniel.j.crichton@jpl.nasa.gov).

This template is intended to encompass both SMD-wide and division-specific guidance for the OSDMP and has been developed from the available templates produced by the divisions:  
Astrophysics (APD) template not released yet — Biological and Physical Sciences [(BPSD) template](https://science.nasa.gov/researchers/templates-biological-and-physical-sciences-division-appendix-e-roses-proposals) (coming soon) — Earth Science [(ESD) template](https://www.earthdata.nasa.gov/engage/data-management-guidance#osdmp) — Heliophysics [(HPD) template](https://science.nasa.gov/researchers/templates-heliophysic-division-appendix-b-roses-proposals) — Planetary Science [(PSD) template](https://science.nasa.gov/researchers/templates-planetary-science-division-appendix-c-roses-proposals).

Check your funding solicitation for specific guidance on the OSDMP, which supersedes the general guidance provided here. General guidance from SMD on the OSDMP is also available in the [SMD Open-Source Science Guidance](https://science.nasa.gov/oss-guidance) and on the [ROSES OSDMP web page](https://science.nasa.gov/researchers/sara/faqs/OSDMP).

All ROSES-23 proposals must comply with the SMD [open science policy, SPD-41a](https://science.nasa.gov/science-red/s3fs-public/atoms/files/SMD-information-policy-SPD-41a.pdf). Each SMD division has created their own policy to complement SPD-41a, which may be found at the links below (as of 6/13/23):

[APD policy](https://science.nasa.gov/science-pink/s3fs-public/atoms/files/APD_Data_Policy_v1.8.pdf) — [BPSD policy](https://science.nasa.gov/science-pink/s3fs-public/atoms/files/BPS_Scientific_Data_Management_Policy_Baseline_TAGGED.pdf) — [ESD policy](https://www.earthdata.nasa.gov/engage/open-data-services-and-software) — [HPD policy](https://science.nasa.gov/science-red/s3fs-public/atoms/files/HPD_Data_Policy_Final_20230202_0.pdf) — [PSD policy](https://science.nasa.gov/science-red/s3fs-public/atoms/files/PSD%20Information%20&%20Data%20Management%20Policy.pdf)

## Data Management Plan

DMPs must include the elements listed in the headings below.

Division-specific guidance for DMPs:

APD: <https://science.nasa.gov/astrophysics/astrophysics-data-centers>

BPSD: <https://science.nasa.gov/biological-physical/data>

ESD: <https://www.earthdata.nasa.gov/engage/dmp-earth-science>

HPD: <https://science.nasa.gov/heliophysics/heliophysics-data>

PSD (if archiving with PDS): <https://pds.jpl.nasa.gov/home/proposers/> and <https://pds.jpl.nasa.gov/home/proposers/archiving/Individual-Proposers-Archive-Guide-v5_clean_20190424.pdf>

Type your text here; use Body\_RS style.

### Expected data types, formats, volumes, and standards

See the Table under “Description of the data” at <https://www.earthdata.nasa.gov/engage/dmp-earth-science> for example data types, formats, and standards (cross-division).

Data types: What the data pertains to, such as “Jupiter orbital magnetic measurements” or “digital elevation maps of Greenland”

Formats: File format(s), i.e. what is contained in the file extension, such as “columnar ASCII text (.csv)”, “JPEG images (.jpg)”, or “HDF5 data (.mat)”

Volumes: Estimate how much data your project will produce, consistent with your proposal STM. ESD guidance specifically states that volumes are not necessary; however, showing reviewers you have thought about the volume can help convince them you understand the needs of your project well, especially when the data volume affects your budget.

Standards (optional): If your data will follow any standards (e.g. PDS4), state what the standards are. Provide any additional information reviewers might need to confirm you will be able to follow (and know) the standards, such as PDF/A descriptions of the data that PDS4 requires. PDS guidance states that “In a ROSES proposal the proposer must demonstrate an understanding of the work involved in preparing data for the PDS.” As the OSDMP is considered as part of proposal merit, this is good general guidance.

License: Data should be released under a Creative Commons Zero (CC0) license. APD and HPD require CC0 if there are no other restrictions on the data.

We will generate labeled imaging datasets for machine learning analysis from the HiRISE instrument. Labeled images will be in the PDS4 format. We will construct a PDS4-compliant local data dictionary (LDD) that contains standard metadata features for annotating images for machine learning classification. The LDD will be used to generate annotated feature labels as required for PDS4 compliance. Labeled images will include approx. 100,000 images and 100 TB of data. The software will generate approximately 10 MB of ASCII PDS4 XML labels. The labels and image data will be released under a Creative Commons Zero license.

Type your text here; use Body\_RS style.

### Repositories and timeline for sharing data

Specify the repository(ies) that will be used to archive and provide public access to data and metadata arising from the activities and the schedule for making data publicly available. Include a description of how data will be archived to enable long-term preservation. Also include any additional prerequisites needed to prove you will be able to use the desired repository, such as a Letter of Support for archiving in a PDS node.

**HPD proposers:**

HPD requires that data must be archived in one of the following:

* Space Physics Data Facility (SPDF, <https://spdf.gsfc.nasa.gov/>)
* Solar Data Analysis Center (SDAC, <https://umbra.nascom.nasa.gov/index.html/>)
* Community Coordinated Modeling Center (CCMC, <https://ccmc.gsfc.nasa.gov/>)

HPD proposals must state which of these is planned for data archiving. The following description is included in the HPD OSDMP template:

Data products based on observations shall be archived and registered at the SPDF for non-solar data or the SDAC for solar data according to accepted heliophysics protocols (<https://spdf.gsfc.nasa.gov/guidelines/archive_newdata_reqt.html>), including the use of standard data formats (e.g., CDF, FITS, NetCDF) and the relevant heliophysics standards associated with these, or the heliophysics standard Event List format (<https://hpde.gsfc.nasa.gov/HPEvent_List_Specification.pdf>, along with standard metadata (SPASE descriptions including parameters; see <http://spase-group.org>). Simulation output will be registered by the CCMC and shall conform to their requirements so that it is accessible and usable by a typical scientist. For simulation output, metadata regarding the model and model run(s) are required. Proposers should work with CCMC to produce model output readers and interpolators compatible with the CCMC Kamodo Open Source Project, (<https://github.com/nasa/Kamodo>).

The HPD template states “Researchers will collaborate with data curation scientists at the archives to develop data products suitable for long-term scientific use.” The HPD policy states that data providers must work with the SPASE Metadata Working Team for DOI generation: <https://hdrl.gsfc.nasa.gov/smwt_home/smwt_index.html>. OSDMPs to HPD program elements must include this collaboration in the timeline described in this section.

HPD also forbids data shared exclusively as part of supplemental information along with a journal publication. ALL data must be shared as part of a DOI-linked archive with Zenodo or equivalent.

***PSD Proposers:***

PSD requires that physical sample management be detailed in the OSDMP, and data pertaining to physical samples be archived in [AstroMat](https://www.astromat.org/). Data must be archived with an [International General Sample Number (IGSN)](https://www.igsn.org/). The PSD policy says:

Physical materials include all astromaterials, synthesized physical materials and biomaterials, analog materials, and associated analytical standards. Any physical materials of scientific value that are collected, purchased, or produced by a PSD-funded research award, and are not consumed during the proposed research, must be made publicly available before the end of the award’s period of performance. Any other materials should be made publicly available when it is practical and feasible to do so, and when there is scientific utility in doing so. All science data collected from these materials utilized in PSD-funded investigations must also be made available to the public as rapidly as possible, not to exceed the end of the award’s period of performance.

Any other data is recommended to be archived in one of the approved archives/repositories listed at <https://science.nasa.gov/solar-system/archives>, or otherwise justified in the OSDMP. Mission data must be archived in PDS except when it pertains to physical samples. Data shared as supplemental information along with a journal article must be publicly accessible without the need for a journal subscription or institutional credentials.

We will archive our data in the Imaging Node of the Planetary Data System (PDS). A Letter of Support from the Imaging Node manager is included in <Section 8> of this proposal. The delivery process to the PDS Imaging Node will begin in the third quarter (Q3) of year 2 (Yr2) of the project and will be completed by the end of the funding period. PDS archives are publicly available, include a persistent identifier (DOI), and support long-term preservation.

Type your text here; use Body\_RS style.

### Description of data types that are exempt from data sharing requirements

Optional: This section should only appear in your OSDMP if your proposed project will include ITAR-, EAR-, HIPAA-, or other restricted data, or if you are requesting a variance (exception) from SPD-41a from the Program Officer for your solicitation. **Cite the relevant laws, regulations, or policies that generate the exclusion.** Note the following example is not relevant to the HiRISE image example.

Instrument control data will be generated that are restricted under NASA-STD-1006A. These data are unable to be shared.

For Dual-Anonymous Peer Review (DAPR) proposals, if the exception is due to a preexisting contract, like the NASA/Caltech Prime Contract, include the following text: A variance is requested from the Program Officer and is detailed in the non-anonymized portion of this proposal.

Type your text here; use Body\_RS style.

### Roles and responsibilities for data archiving

Describe who will be responsible for handling each task required to accomplish the plan detailed above. Reuse identifiers from elsewhere in the proposal, such as Co-I-1.

Note: This section may instead be separated into a final major section that details roles and responsibilities for both data and software together.

The PI will be ultimately responsible for the archiving of data products in accordance with this OSDMP. All team members will generate data. Co-I-2 will maintain the data before it is delivered to the PDS for archiving. Co-I-1 will work with the PDS Imaging Node to deliver the dataset.

Type your text here; use Body\_RS style.

## Software Management Plan

A software management plan is required for all SMD-funded activities that are expected to produce software. If the activity is not expected to produce software, include a statement such as: “No software development is anticipated for this effort. If software is created, it will be made publicly available to the extent legally permitted per the Scientific Information Policy for the Science Mission Directorate.”

ESD and PSD both do not offer specific guidance or requirements for the software management portion of the OSDMP. Templates are not yet available for APD or BPSD.  
HPD suggests GitHub and Zenodo or equivalent services for software development and archiving, respectively. Apache Version 2.0 is strongly recommended for HPD—if you will be using a different license, explain why. For models using the CCMC, refer to the onboarding guide (<https://ccmc.gsfc.nasa.gov/model-onboarding/>). The software management plan detailed in this section must include a description of how the proposed project will complete the steps listed in the model onboarding process. The HPD template recommends (but does not require) inclusion of documentation such as a README giving instructions on how to use the software. Describe your intended documentation, process for writing it, and who will be responsible for its delivery. HPD requires that best practices be followed as much as possible, as detailed in <https://arxiv.org/pdf/2106.01477.pdf>. State that you will do so, and explain any clear departure from these guidelines.

Type your text here; use Body\_RS style.

### Expected software types

Describe the software expected to be produced including its purpose, language(s), in what manner it will be developed, and new features or updates for existing software. Note that a code of conduct and guidelines for contribution are specifically required for software repositories under SPD-41a.

Optional: If your software will follow a specific coding standard (e.g. Apache Coding Standards for Java, PEP for Python), name the standard and describe how you will follow it. This is a recommendation from SPD-41a, so it can improve the merit of your project to include, but it is not strictly necessary.

We will produce a Python software package that writes PDS4 labels as an output of a machine learning image classification model of a planetary object. The software will be developed on GitHub, initially in a private repository. Supporting documentation (ASCII text files) and sample test data (a subset of the data described above), including a training set, will be included in the GitHub and archive.

We will develop the software from Q1 Yr1 as required for the proposed research project. We will follow the style guide for the Python Enhancement Proposal (PEP), which describes best practices for open-source development of Python code. Software will be documented in-line according to these standards and a descriptive README will provide detailed instructions for how to run and use the software. A Code of Conduct and Guidelines for Contribution will be included in the GitHub. We will also apply a Continuous Integration (CI) development process to identify bugs and security issues as early as possible.

Type your text here; use Body\_RS style.

### Repositories and timeline for sharing software

Specify the repository(ies) that will be used to archive software arising from the activities and the schedule for making software publicly available. This should include the license under which the software will be made available. If there are no other restrictions, the software should be released under a permissive license. See above for HPD-specific guidance.

The software will be released under the NASA-Planetary-Science organization page (<https://github.com/NASA-Planetary-Science>) by the time the associated data are archived in the PDS, Q3 Yr2, under an Apache License Version 2.0. After release, software development will continue only in the public GitHub repository. The package will be archived with a persistent identifier (DOI) on Zenodo at the time of release. Documentation will include instructions for how to provide feedback. After conclusion of the proposed project, feedback will be incorporated on a best-effort basis. The GitHub page and a citation to the Zenodo archive will be included in the associated scientific publication.

Type your text here; use Body\_RS style.

### Description of software that are exempt from software sharing requirements

Optional: This section should only appear in your OSDMP if your proposed project will include ITAR-, EAR-, HIPAA-, or other restricted software, or if you are requesting a variance (exception) from SPD-41a from the Program Officer for your solicitation. **Cite the relevant laws, regulations, or policies that generate the exclusion.** Note the following example is not relevant to the HiRISE image example.

Instrument control software will be generated that are restricted under NASA-STD-1006A. This software is unable to be shared.

For Dual-Anonymous Peer Review (DAPR) proposals, if the exception is due to a preexisting contract, like the NASA/Caltech Prime Contract, include the following text: A variance is requested from the Program Officer and is detailed in the non-anonymized portion of this proposal.

Type your text here; use Body\_RS style.

### Roles and responsibilities for software release and archiving

Describe who will be responsible for handling each task required to accomplish the plan detailed above. Reuse identifiers from elsewhere in the proposal, such as Co-I-1.

Note: This section may instead be separated into a final major section that details roles and responsibilities for both data and software together.

The PI will be ultimately responsible for archiving and public release of software in accordance with this OSDMP. All team members will develop software and write documentation. The PI will manage the private GitHub repository, including implementing the Code of Conduct, Guidelines for Contribution, and public release on the NASA-Planetary-Science organization page along with the timeline above. Co-I-1 will archive the software release on Zenodo.

Type your text here; use Body\_RS style.

## Publication of Results

Describe the types of publications that are expected to be produced from the activities (e.g., peer reviewed manuscripts, technical reports, conference materials, and books).

Conference abstracts and a publication in a peer-reviewed journal are anticipated as an output of this work. We will present the work at the American Geophysical Union (AGU) Fall Meeting in Yr2. In addition, we anticipate a paper in *IEEE Software* describing the architecture and approach. All publications will have as-accepted drafts deposited in a NASA-designated manuscript repository by publication date; this step will be completed as part of the proposing institution’s manuscript submission process.

Type your text here; use Body\_RS style.

## Other Open Science Activities

Optional: This section should only appear if additional open science activities are included in the proposed project. Describe additional open science activities associated with the project, as these activities may improve the merit of the proposed project. If any described activity will have clear roles and responsibilities, describe them and the proposal team member for each role. Such activities may include:

* Holding scientific workshops and meetings openly to enable broad participation
* Preregistering research plans in advance of conducting scientific activities
* Providing project personnel with open science training or enablement (if not described elsewhere in this proposal)
* Implementing practices that support the inclusion of broad, diverse communities in the scientific process as close to the start of research activities as possible (if not described elsewhere in this proposal)
* Integrating open science practices into citizen science activities
* Contributions to or involvement in open-science communities

Prior to the release of the private GitHub, a research plan will be released on the NASA-Planetary-Science GitHub page where the software will eventually be released. Community comments and feedback in response to the public research plan will be considered alongside the proposed project during development. One month of schedule margin is included in the software development timeline in order to account for additional suggested features from the community. The PI will be responsible for releasing the research plan and incorporating feedback into the project plan and timeline.

Type your text here; use Body\_RS style.