

Response to RFI: Implementation and Changes to Science Policy
Document (SPD)-41: Science Information Policy
from the Laboratory Astrophysics Division executive committee of the
American Astronomical Society

Primary author: Murthy Gudipati (Jet Propulsion Laboratory, California Institute of Technology), submitting in his capacity as Chair of the Laboratory Astrophysics Division (LAD, lad.aas.org) of the American Astronomical Society (AAS), lad.chair@aas.org

Additional authors: the LAD executive committee composed of
Stefanie Milam (NASA Goddard Space Flight Center), Vice Chair, lad.vchair@aas.org,
Phillip Stancil (University of Georgia), Past Chair, lad.pastchair@aas.org,
Christiaan Boersma (NASA Ames Research Center), Treasurer, lad.treasurer@aas.org
Rachel Smith (NC Museum of Natural Sciences/Appalachian State University),
Secretary, lad.secretary@aas.org
Brett McGuire (Massachusetts Institute of Technology), Member-at-Large for Molecules,
brettmc@mit.edu
Dawn Williams (University of Alabama), Member-at-Large for High Energy Particles,
drwilliams3@ua.edu
Edith Fayolle (Jet Propulsion Laboratory, California Institute of Technology),
Member-at-Large for Dust & Ices, edith.c.fayolle@jpl.nasa.gov
Carla Fröhlich (North Carolina State University), Member-at-Large for Nuclear
Physics, cfrohli@ncsu.edu
Joan Marler (Clemson University), Member-at-Large for Atoms, jmarler@clemson.edu
Michael Hahn (Columbia University), Member-at-Large for Plasma,
mhahn@astro.columbia.edu
Silvia Protopapa (Southwest Research Institute), Member-at-Large for Planetary
Science, sprotopapa@boulder.swri.edu

Brief Summary

Our responses to the RFI focus on the SPD-41 general policies and the policies for research and they are relevant to the astrophysics, planetary science, and heliophysics SMD divisions. While we support open science and data sharing, we have concerns about SPD-41 regarding the potential for increased burden on small-team investigators, competitiveness of early-career researchers building codes, the usefulness of the aggregating data from new projects only, and the standing of the preeminent repositories in our subfields. Our concerns (Aspect 1) and requests for clarifications and suggestions (Aspect 2) are detailed below.

Aspect 1: How would the proposed changes to the existing SPD-41 impact research and related activities of LAD members?

Point 1: The proposed modification of SPD-41 may unnecessarily increase the archiving burden on recipients of smaller ROSES grants. Depending on SMD's definition of "scientifically useful data", a large amount of raw data, intermediate processed data, calibrations, standards, codes or spreadsheets that turn raw data into final results may require archiving. This will result in a significant burden on the proposal team and this data might be of extremely limited use to the scientific community and the public at large. Very few small single-PI teams have the institutional resources and infrastructure to meet the proposed requirements, and certainly many, if not all, will lack the funding to achieve compliance on a reasonable timescale.

Point 2: The immediate public archiving of codes developed by scientists, especially early-career, would reduce the incentives to develop these invaluable scientific tools. The rapid public availability of such codes would be particularly damaging to early-career scientists for whom the tools they have developed themselves form the basis for their future collaborations and help them to establish their expertise in a particular area. In our field of laboratory astrophysics, immediate public archiving of codes that took years of building would be equivalent to allowing public on-demand access to experimental set-ups, which is not required and feasible. This puts young scientists who build codes at a disadvantage.

Point 3: If NASA can provide sufficient financial and logistical support to investigators for archiving the vast amounts of useful data generated by laboratory astrophysics researchers, then data repositories following the practices enumerated in Appendix D could prove useful to the community. However, aggregating fundamental data only from new SMD projects would severely hamper the usefulness and limit the value of these repositories.

Point 4: The definition of an acceptable repository as described in Appendix D is too stringent and would prevent adequate archiving and compliance to SPD-41. Preeminent repositories in our subfields as described in Appendix D are often grassroot efforts or managed by non-NASA entities and thus may not follow other criteria listed in that same appendix such as being citable or certified. Changing archiving practices in subfields to repositories fulfilling appendix D characteristics would dilute and decrease the impact of SMD funded work since non-NASA funded colleagues may not embrace the shift proposed by the updated SPD-41. In addition, standard repositories do not exist for some subfields, especially for new ones and/or those that are interdisciplinary in nature, preventing compliance to SPD-41.

Point 5: Some institutions have rules about intellectual property rights that may not be compatible with SPD-41. This could prevent investigators applying for ROSES grants, possibly leading to a diminishing return for the ROSES program overall.

Aspect 2: What support, services, training, funding, or further guidance is/are needed to support the successful implementation of the existing or proposed information policy?

With respect to our Point 1, **the final version of SPD-41 should define “scientifically useful data”, section V. B. to be “at the discretion of the investigators”.** Given the diversity of subfields and types of data generated within the SMD, investigators should have broad discretion to define what is meant by useful data. This discretion can be scrutinized by SMD during the proposal stage in the Data Management Plan and throughout the grant via the Annual Reports. Alternatively, the final guidelines should adopt a limited and specific definition of “scientifically useful data” that balances the needs of the community and the reasonable effort required by the investigators.

With respect to our Point 1, **SMD must provide a prorated increase in proposal funding** to support the extra burden on the investigators and to allow subcontracting data archiving professionals. That need will depend on the definition of “scientifically useful data” in section V. B (see previous paragraph).

With respect to our Point 1, **SMD should provide a list of archiving services to all SMD funded proposers;** all costs associated with archiving should be specified and fully covered by NASA.

With respect to our Point 2, **a proprietary period for codes should be allowed on the PI’s request.** In the case of codes generated as part of a PhD student thesis, this period should extend to a minimum of 6 years (i.e., two postdoctoral appointments) past the student’s matriculation. Failure to do so could result in students being unable to establish themselves in their fields with their own work, as larger, established teams can simply annex their code and bypass the junior researchers.

With respect to our Point 3, we suggest that SMD **support efforts to mine the literature and include datasets from earlier SMD projects and other science fields.** This would help data users by grouping similar fundamental physical and chemical data into a single platform, prevent unnecessary and wasteful duplication by data providers, and allow for identification of gaps in the literature which would guide future research.

With respect to our Point 4: **SMD should identify and sustainably fund data repositories that are considered preeminent in subfields,** with decisions on which repositories to financially support made by the communities themselves. NASA/SMD should take responsibility to ensure the quality and ease of availability of such

repositories, as this would be a bottleneck to achieve the goals of the SPD-41 initiative. The minimum funding horizon for repositories should be at least 15 years, with advanced notice of at least 5 years before any termination or decrease in funding.

With respect to our Point 4, **SMD should publish and update a list of acceptable repositories and it should assist investigators in identifying the adequate repositories for data archiving upon the investigators' request.** Regarding the fundamental data generated by laboratory astrophysics, we urge NASA to identify compliant repositories that have less data formatting and archiving constraints than, e.g., the Planetary Data System in its existing form. While such repositories might be appropriate for observational data sets, they are not suitable for fundamental physical and chemical data. In fact, this is the main reason why laboratory planetary scientists hesitate to submit proposals to the PDART ROSES program. Archiving in unnecessarily complex formats would be one of the biggest hurdles and time commitments for small-budget investigators producing extensive experimental data.

With respect to our Point 5, SMD should work with investigators and institutions to ensure that SPD-41 requirements do not infringe on institutional intellectual property rights, especially pertaining to code upgrades.