

White paper in response to request by the committee on “**Best Practices for a Future Open Code Policy for NASA Space Science**”

Title: Comments for Open Code Policy for NASA SMD

Authors:

V. G. Merkin¹ (slava.merkin@jhuapl.edu, 240-228-1756)

K. Sorathia¹

L. Daldorff¹

A. Ukhorskiy¹

M. Sitnov¹

J. Lyon²

¹The Johns Hopkins University Applied Physics Laboratory

²Dartmouth College

Short Summary:

The comments below are based on many years of scientific code development and use by the authors, primarily funded by NASA and NSF research grants.

Comments to the committee:

1. While we sympathize with the idea of making research codes open source, we would like to point out that preparing codes for public release requires significant resources, i.e., funding, which historically has never been allocated. In other words, the open source requirement must not be issued as a mandate without funding, in which case it would be destined to fail.
2. To elaborate on why dedicated funding is required for open-sourcing a scientific code, an open source code means: sufficiently detailed documentation, build and configuration system, auto-testing, clean code and good programming practices. These features require additional funding not only for scientist developer time but for software engineers and system administrators.
3. Such open source code development can only be funded via contracts – not research grants – where the open source code is a specific deliverable. Otherwise, the requirement cannot be enforceable as results of research grants are not known a priori by definition of scientific research.
4. A distinction must be made between large scale production codes developed under dedicated NASA funding for long term service in the community (in which case an open source requirement might be warranted and beneficial), from boundless helper programs, visualization scripts, and all kinds of auxiliary software that each researcher develops in their everyday work, sometimes for limited or even one-time use. Requiring that such programs (e.g., visualization scripts used to produce a specific plot in the paper) be open source will place a strenuous burden on already overworked researchers, ultimately degrading the quality of their science. Requiring that data used for a particular plot be made available upon request is sufficient for reproducibility.

5. Scientific code developers and researchers must be allowed an opportunity to include clauses in the open source license used by their code that will protect their intellectual property going forward. An example might be a clause requiring permission from the authors of the code to publish or propose research resulting from or using the open source code.