86 FR 60021 - Stewardship of Software for Scientific and High-Performance Computing

AGENCY: Office of Advanced Scientific Computing Research (ASCR), Office of Science, Department of Energy.
ACTION: Request for information.

Responses to questions relevant to the work of our organization.

NumFOCUS serves as a 501(c)(3) public charity that promotes open practices in research, data, and scientific computing by serving as a fiscal sponsor for open source projects and organizing community-driven educational programs. NumFOCUS's comprehensive stewardship includes project advocacy, legal and operational support, fiscal administration and grant management, governance support, DEI assistance, community outreach, events, and user education. The following responses are based on NumFOCUS's ongoing support and interactions as a hub for our 45 Sponsored and 51 Affiliated OS scientific and high-performance computing projects.

(5) Challenges in building a diverse workforce and maintaining an inclusive professional environment:

- What challenges do you face in recruiting and retaining talented professionals to develop software for scientific and high-performance computing?
    - Identifying and assessing relevant skill sets, domain knowledge, and technical skills
    - Appropriate pay structures
    - The need and resources for onboarding and mentoring programs

- What additional challenges exist in recruiting and retaining talented professionals from groups historically underrepresented in STEM and/or individuals from underserved communities?
    - Language barriers (ESL) - documentation and code comments
    - Gender bias - double-blind hiring practices
    - Education pipeline - many enter the field of scientific computing through research programs as postdoc or graduate students. The educational pipeline where this highly-specialized domain-specific knowledge is obtained is often lacking diversity especially from groups that include African Americans, Blacks, Hispanics, Native Americans, women, LGBTQ+, and people who are blind or deaf.
    - Volunteering is not a feasible option for many
    - Access to infrastructure and technology
    - Skilled professionals trained in the relevant communities of practice
    - Lack of a targeted mentoring program for onboarding underrepresented groups

- What challenges exist in maintaining inclusivity and equity in the development community for scientific and high-performance computing software?
    - Attracting contributors from broader communities
    - Maintaining a welcoming and inclusive environment in documentation, online communities, and governance models.
    - Identifying and incorporating inclusive language practices into communications and documentation.
    - The global nature of open source translates to projects with a geographically diverse contributor base. This strength can be a challenge when hiring for awards with citizenship requirements.

- What successful strategies have you employed to help overcome these challenges?
    - Assisting our projects in maintaining a welcoming and inclusive environment in documentation, online communities, and governance models.
        - Ensuring entry-points for new contributors are updated and visible
        - Having new-contributor guidelines clearly visible in the repo
    - Identifying and incorporating inclusive language practices into communications and documentation.

(6) Requirements, barriers, and challenges to technology transfer, and building communities around software projects, including forming consortia and other non-profit organizations:

- What are the important characteristics and components of sustainable models for software for scientific and high-performance computing?
    - A consistent funding network made up of diverse sources
    - A transparent and inclusive governance structure that clearly defines roles and responsibilities
    - Transparent and equitable mechanisms for resource allocation and management
    - A well established and engaged community not concentrated at a single entity
    - A mentoring program for new contributors

- What are key obstacles, impediments, or bottlenecks to the establishment and success of these models?
    - Lack of legal, financial, and administrative infrastructure to facilitate resource allocation
    - Lack of human capital to facilitate community stewardship
    - Lack of funding geared toward project management, documentation, and community management

- What development practices and other factors tend to facilitate the successful establishment of these models?
    - Comprehensive and continuously updated documentation
    - Ongoing recruitment and retention strategies for maintainers and developers
    - Ongoing community engagement and outreach
    - Clear and consistently curated contribution workflows

(7) Overall scope of the stewardship effort:

Potential Scope: Scientific software stewardship is multi-faceted, potentially including, but not limited to:

• Training:

Providing training on software-development best practices and the use of core software.

• Workforce support:

Providing outreach and support activities to build and maintain a diverse, skilled workforce with opportunities for professional recognition and career advancement.

• Infrastructure:

Providing infrastructure for software packaging, hosting, testing, and other common capabilities.

• Curation:

Establishing governance processes and standards to enable resource allocation in the most-effective manner balancing stability with the need to satisfy evolving requirements.

• Maintaining situational awareness:

Defining, publishing, and communicating understandable information about relevant software and its dependencies; collecting information from users and deployment requirements from facilities.

• Shared engineering resources:

Providing software-engineering resources to assist with maintenance activities of key projects, including triaging problems from testing and adjusting for new compilers; system software and platform versions; and changing package requirements.

• Project support:

Providing support for the continued development of key projects, including enhancing them to function efficiently on new hardware platforms; take advantage of emerging hardware and software technologies; comply with best practices, and otherwise provide high priority features desired by other users.

- Are there activities that should be added to, or removed from, this list?
    - A detailed DEI strategy or best practices. The DISC (Diversity in Scientific Computing) committee is a NumFOCUS community initiative working in these areas.
    - Fiscal and Administrative support:
        - Providing support for fiscal sponsors and other institutions that provide critical financial, legal, and administrative support for projects
    - Opportunities for cross-project collaboration and networking. NumFOCUS's events team organizes events that offer opportunities for projects to meet as a broader ecosystem and on a domain-specific scale.

- Are there specific requirements that should be associated with any of these activities to ensure their success and maximize their impact?
    - These activities should be required to be informed by project leadership and community feedback. Efforts should be aligned with project goals and objectives as well as community needs.

(8) Management and oversight structure of the stewardship effort:

- What do you anticipate will be effective models for management and oversight of the scientific and high performance-computing software ecosystem, and how would that management structure most effectively interact with DOE and other stakeholders?
    - Projects should be openly led and community-driven but participate in partnerships with entities that can act as an intermediary between DOE and other stakeholders on their behalf. This model ensures that the institution can guarantee compliance with the requirements of DOE and other stakeholders without placing unnecessary administrative burdens on the project leadership or their communities.

- In addition to DOE, who are the key stakeholders?
    - Other U.S. Government agencies rely on well-supported scientific and high-performance computing projects to inform decision making, and strategy, as well as to facilitate the functioning of technical infrastructure and to process data to track the efficacy of programs.
    - Corporations in the private sector rely heavily on these projects in many of the same ways, but also for creating and facilitating the functioning of their products and services.

    - These scientific computing projects are also the backbone of US and global scientific research, critical to universities, scientific labs, think tanks, and independent government agencies including but not limited to NSF and NASA.
    - Non-profit organizations serving as fiscal sponsors of OS scientific computing projects.

- How can the management structure coordinate with DOE user facilities and others to provide access to relevant testbed systems and other necessary infrastructure?
    - This management structure is uniquely positioned to facilitate access to testbed systems and other infrastructure because the institution can act as a gatekeeper in both an upstream and downstream capacity. The institution can coordinate access to testbed systems and other infrastructure to align them with project needs and initiatives. The institution can also communicate goals and challenges of projects to DOE which can inform DOE's allocation of resources.

(9) Assessment and criteria for success for the stewardship effort:

- What kinds of metrics or criteria would be useful in measuring the success of software stewardship efforts in scientific and high-performance computing and its impact on your scientific fields or industries?
    - Success in meeting release dates
    - Completion of roadmap items
    - Project contributor and user growth
    - NumFOCUS sends out an annual project survey that is helpful in determining the success of our efforts and driving new initiatives.

(10) Other:

- What are key obstacles, impediments, or bottlenecks to progress by, and success of, future development of software for scientific and high-performance computing?
    - Lack of available funding for maintenance, testing, and bug fixes
    - Lack of human resources for continued development
    - Lack of project and community management resources
    - Lack of mechanisms to allocate resources to project maintainers