CORSA: Supporting open source research software projects via foundations and metrics

CHAOSSCon-EU
January 30, 2025

Daniel S. Katz[1], Gregory R. Watson[2], Addi Malviya Thakur[2],
Elaine M. Raybourn[3], Bill Hoffman[4], Dana Robinson[5]
[1]*UIUC,* [2]*ORNL,* [3]*SNL,* [4]*Kitware Inc.,* [5]*The HDF Group*

Center for Open-Source Research Software Sustainability and Advancement

CORSA

- Vision: Become a community of practice that creates pathways for open-source scientific software projects to avail themselves of resources for long-term growth, stewardship, advancement, and innovation.

- Scope: research and scientific computing software projects

Greg Watson, Oak Ridge National Laboratory

Addi Thakur Malviya, Oak Ridge National Laboratory

Elaine Raybourn, Sandia National Laboratories

Daniel S. Katz, University of Illinois Urbana Champaign

Dana Robinson, The HDF Group

Bill Hoffman, Kitware, Inc.

Consortium for the Advancement of Scientific Software

- A federation of member organizations supporting specific communities or providing crosscutting capabilities related to scientific software stewardship

- **Inward-facing activities** to support and enhance software products
  - Improving software development practices, sustainability, quality, trustworthiness
  - Improving user experience, integration into the larger ecosystem

- **Outward-facing activities** to
  - Curate the collection of software products,
  - Help teams find and engage with users
  - Help the community find and use software that might benefit them

https://cass.community

Origins and Founding Members

Launched by 8 software stewardship organizations (SSOs) supported by the U.S. DOE, Office of Advanced Scientific Computing Research (ASCR)

- Targeted ecosystem initially includes the ECP Software Technologies software products and other ASCR-supported software

COLABS	**CORSA**	**FASTMATH**	**PESO**
Training, workforce development, and building the RSE community	Partnering with foundations to provide sustainable pathways for scientific software	Stewardship, advancement, and integration for math and ML/AI packages	Stewarding, evolving and integrating a cohesive ecosystem for DOE software

RAPIDS	**S4PST**	**STEP**	**SWAS**
Stewardship, advancement, and integration for data, visualization and ML/AI packages	Stewardship, advancement and engagement for programming systems	Stewardship, advancement of software tools for understanding performance and behavior	Stewardship and project support for scientific workflow software and its community

CORSA goals

Create pathways to open source software foundations

Empower communities with software sustainability metrics

Facilitate cross-cutting activities to address the needs of unique and diverse communities

With varied formats and methods

Provide objective guidance to software communities

Foundation pathways

CORSA's foundation pathways goals

- **Communicate the benefits** of joining a foundation to projects.

- **Streamline the process of joining a foundation** in collaboration with foundations.

- **Work with projects to understand foundations' lifecycle processes** and with foundations to understand project lifecycles.

- **Identify various support models for software projects within foundations.**

- Tailor support models (pathways) to different contexts and project types.

CORSA has set up a public **open source foundations forum** to support the growth of the ecosystem toward foundation policy stewardship and resource and information exchange.

Why open-source foundations matter for sustainability?

- Foundations foster a connected ecosystem where projects can collaborate with others.

- Foundations provide access to a broader pool of resources, expertise, and support networks that individual projects may struggle to maintain independently.

- Provides a neutral and independent home for projects, ensuring they aren't tied to a single developer or institution, which can improve long-term viability.

- Foundations streamline the process of joining by offering support models and governance frameworks.

- Foundations showcase projects to potential funders, users, and collaborators, enhancing visibility and long-term support.

Foundations CORSA works with

- Initially focusing on
    - NumFOCUS
    - High Performance Software Foundation (HPSF) (a Linux Foundation subfoundation)

  as foundations that CASS projects have already joined

- Also includes
    - Apache Software Foundation
    - Eclipse Foundation
    - LF AI & Data Foundation (a Linux Foundation subfoundation)
    - Open Molecular Software Foundation

    - Open to more

CORSA's general foundations work

- Work with foundations (and their documents) to provide
  - Instructions for project on how to join
  - What projects need to have/do before they can join

- Have collected and provide guidance, documents, examples, templates for open-source software projects, such as on governance, roadmaps, contributing, code of conduct, licenses, etc. (public on GitHub)
  - Generally needed for projects to join foundations

- Providing general benefit of foundations, as well as benefits of each

- Running a monthly foundations forum: zoom call to talk about general issues and project/foundation experiences (sign up for info via CORSA mailing list)

CORSA's specific projects & foundations work

- Talking to projects about why they have / haven't joined a foundation

- How can we help projects overcome perceived obstacles

  - Can be related to local environment or foundations

- How can we help foundations better reach out to projects

- What can we learn to help other projects

Software sustainability metrics

Defining software sustainability

- Software sustainability refers to the capacity of software to endure and remain operational, useful, and relevant over time.

- It involves ensuring that the software can be maintained, adapted, and improved as needs evolve, even as original developers move on or organizational priorities shift.

- It's not just about keeping software running but ensuring it remains effective, valuable, and easy to maintain or improve over the long term.

Measuring software sustainability

Purpose of software sustainability metrics: measuring the sustainability of scientific and research software.

Questions: Which metrics would be most effective? Which are feasible to collect?

Motivation: Projects need to know what to improve and how well they are doing; Funders need to know what they could invest in

Software community has developed various approaches to this

- Metrics collected and described by CHAOSS
- Software Sustainability Matrix from Kitware
- OpenSSF Scorecard and Criticality Score

Interest from NumFOCUS, Research Software Alliance (ReSA), and other orgs

Potential software sustainability considerations

Lots of people have thought about this, and come up with:

- Maintainability
- Reusability
- Portability
- Scalability
- Community Engagement and Support
- Reproducibility
- Interoperability
- Funding and Institutional Support
- Testing and Quality
- Impact and Recognition

CORSA has set up a **metrics working group** to support the understand metrics and how to collect them.

References:
https://chaoss.community/kb-metrics-and-metrics-models/
https://www.software.ac.uk/publication/better-software-better-research
https://bssw.io/blog_posts/use-of-software-metrics-in-research-software

Insights from projects at CORSA workshops and meetings

Key Takeaways from these Discussions on Sustainability

Sustainability as an ecosystem: It's not just about individual software but the broader software ecosystem, where interdependencies are key.

Competition and funding pressures: Sustainability is not just technical but also influenced by the competitive landscape (e.g., dominant players in the field) and the availability of funding.

Adaptability perspective: Sustainability, especially for HPC software, also depends on how easy the software is to install, maintain, and adapt across different systems.

CORSA Metrics Goals and Activities

1. Work with the CASS (Consortium for the Advancement of Scientific Software) and broader scientific software community (e.g., CHAOSS) to identify and develop software sustainability metrics

2. Develop and/or coordinate infrastructure to collect and present metrics

3. Provide training sessions or workshops to help use and interpret metrics

4. Facilitate the integrating of software sustainability metrics into project management processes

5. Assist CASS projects in effectively communicating their sustainability efforts and achievements

CASS Metrics Working Group

- CORSA organizes the CASS Metrics Working group.

- WG goals:

  - To identify metrics important/ relevant to the CASS community

  - To work collaboratively with CASS Impact Framework Working group towards implementing/operationalizing these

  - To work with and help individual CASS projects towards optimizing, improving these metrics per their project goals

- Metrics WG is held bi-weekly via zoom on Wednesdays 4 pm EST. Contact Addi Malviya-Thakur (malviyaa@ornl.gov) for joining info

CASS Sustainability Training

- Is more help (training, hand-on support) needed for projects to become more sustainable?
- Currently being considered by CORSA
  - Starting discussions with CASS projects/SSOs about providing training
  - Can we collaborate with CHAOSS folks who are developing similar training for CZI EOSS awardees? Or just use material?
  - Also investigating also existing material from other sources

Wrap-up

- Please sign up to the CORSA mailing list at https://corsa.center
    - We will post information on upcoming events and meetings as well as distribute the monthly CORSA newsletter
- Access our resources on foundation membership and open source project best practice guides
- Join the monthly Foundations Forum meeting
- Join the bi-weekly Metrics Working Group meeting
- Please reach out to us at info@corsa.center if you would like to get more involved!

Thank you!

We encourage you to get involved in the software sustainability community

https://corsa.center

Subscribe to our group!