Toward a Framework for Evaluating Software Success: A Proposed First Step

Stan Ahalt¹, Bruce Berriman², Maxine Brown³, Jeffrey Carver⁴, Neil Chue Hong⁵, Allison Fish⁶, Ray Idaszak¹, Greg Newman⁷, Dhabaleswar Panda⁸, Abani Patra⁹, Elbridge Gerry Puckett⁶, Chris Roland¹⁰, Douglas Thain¹¹, Selcuk Uluagac¹², Bo Zhang¹³

**Abstract**

Software is a particularly critical technology in many computational science and engineering (CSE) sectors. Consequently, software is increasingly becoming an important component in the evaluation of competitive grants and the execution of research projects. As a result, software can be viewed as a scholarly contribution and has been proposed as a new factor to consider in tenure and promotion processes. However, existing metrics for evaluating the capability, use, reusability, or success of software are sorely lacking. This lack of software metrics permits the development of software based on poor development practices, which in turn allows poorly written software to “fly under the radar” in the scientific community and persist undetected. The absence of evaluation by knowledgeable peers often leads to the establishment and adoption of tools based on aggressive promotion by developers, ease-of-use, and other peripheral factors, hindering the sustainability, usefulness, and uptake of software and even leading to unreliable scientific findings. All of these factors mean that addressing the current lack of software evaluation metrics and methods is not just a question of increasing scientific productivity, but also a matter of preventing poor science.

As a first step toward creating a methodology and framework for developing and evaluating software success metrics for the CSE community, we propose the creation of a software “peer-review group.” This group, comprised of grant recipients funded to develop sustainable software, would meet periodically to evaluate their own and each others’ software, developing and refining success metrics along the way. We envision the group as a pilot test for a potential larger-scale effort to establish a more formal framework for software success metrics and evaluation.

**White paper corresponding to this poster:** [http://dx.doi.org/10.6084/m9.figshare.1561451](http://dx.doi.org/10.6084/m9.figshare.1561451)

---

**Further Questions**

- How will peer-review of code and corresponding peer-review of software success metrics be performed and tied together across different research domains?
- What type of infrastructure will best facilitate this?
- How will ever-increasing open participation by the community be nurtured and measured?
- What governance structure will be put in place to ensure shared community governance of the software and metric peer-review process?
- How will factors of software success be evaluated: e.g. criticality, value, sustainability, usability, reusability, performance, functionality, capability, availability, scientific impact, usefulness, reliability?

---

**Next Steps**

- Use CSESSP, WSSSPE, other future events to build this community and framework & author more white papers
- Continually take advantage of these venues
- Collaborate with other metric & framework projects
- Consider proposal around this idea to give enough runway to build community & metrics
- Encouraging the community to take up this idea and run with it

---

¹RENCI, University of North Carolina at Chapel Hill; ²California Institute of Technology; ³University of Illinois at Chicago; ⁴University of Alabama; ⁵Software Sustainability Institute, University of Edinburgh; ⁶University of California, Davis; ⁷Colorado State University; ⁸The Ohio State University; ⁹The State University of New York, University at Buffalo; ¹⁰North Carolina State University; ¹¹University of Notre Dame; ¹²Florida International University; ¹³Indiana University