Lightning Talk: Report on Software Metrics for Research Software

Gabrielle Allen*, Emily Chen*, Ray Idaszak†, Daniel S. Katz*

*University of Illinois Urbana-Champaign, Urbana, Illinois, USA, Email: {gdallen, echen35, dskatz}@illinois.edu
†RENCI, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA, Email: rayi@renci.org

***Abstract*—We report on a survey to investigate common metrics for research software, following a plan of work established at a WSSSPE3 working group.**

I. Lightning Talk

One of the working groups at WSSSPE3 [1] focused on discussing metrics for research software. Metrics for research software were seen as being important for promotion and tenure, quantifying scientific impact, reducing duplication, and prioritizing development, among other motivations. The group planned to investigate common metrics for research software, to be able to publish a white paper that would be of interest to the community. As a first step in this direction, we began an activity to investigate metrics for software that are being used for the awardees of the National Science Foundation (NSF) Software Infrastructure for Sustained Innovation (SI2) program [2]. All lead principle investigators for SI2 awards were contacted with a request to complete a survey to provide the metrics they had originally proposed to use to assess their software components, and any additional metrics they are currently using.

Specifically, the questions asked of the survey respondents were:

1) Are the software components developed through your award correctly listed at this web site: https://sites.google.com/site/softwarecyberinfrastructure/software/software? If this information is not accurate please list all software components generated through your award here.
2) What metrics did you list in your SSI/SSE proposal or agree to before your award for these software components?
3) If you are now collecting additional metrics for the software components beyond those you planned at that time, what are they?
4) Are there any metrics you planned in your proposal or agreed to before your proposal was awarded that you have since realized that you are not able to collect?
5) Are there any metrics you planned in your proposal or agreed to before your proposal was awarded that you have since realized are not useful?
6) Did you find collecting metrics led to improving the software (e.g. quality, usefulness, sustainability, reliability, performance, impact, etc.)? If so, please indicate which metrics were the most useful and why?

The responses to this survey are currently being collected. We will present an initial analysis of the replies and describe next steps in this activity as a lightning talk at WSSSPE4.

Acknowledgment

Work by E. Chen was supported by the NCSA SPIN program.

References

[1] D. S. Katz, S. T. Choi, K. E. Niemeyer, J. Hetherington, F. Löffler, D. Gunter, R. Idaszak, S. R. Brandt, M. A. Miller, S. Gesing, N. D. Jones, N. Weber, S. Marru, G. Allen, B. Penzenstadler, C. C. Venters, E. Davis, L. Hwang, I. Todorov, A. Patra, and M. de Val-Borro. Report on the Third Workshop on Sustainable Software for Science: Practice and Experiences (WSSSPE3). Technical report, arXiv, 2016. arXiv:1602.02296 [cs.SE].

[2] National Science Foundation. Implementation of NSF CIF21 Software Vision (SW-Vision). http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504817