DOE Software Stewardship Challenges in Diversity, Professional Development, and Retention of Research Software Engineers

Authors: Miranda Mundt (Sandia National Laboratories), Vanessa Sochat (Lawrence Livermore National Laboratory), Daniel S. Katz (University of Illinois at Urbana-Champaign), Sandra Gesing (University of Illinois Discovery Partners Institute), Verónica G. Melesse Vergara (Oak Ridge National Laboratory)

Introduction

We are members of the United States Research Software Engineer (US-RSE) Association[1], a group consisting of Research Software Engineers (RSEs) and allies with various educational backgrounds, areas of expertise, and professional experiences. Our association focuses on supporting the increasingly important role of RSEs in the computational science and engineering (CSE) community.

The introduction of the research software engineer as a professional designation reflects a shift in scientific computing towards prioritizing the development and application of research software engineering practices, while recognizing that they are different than traditional software engineering methods as used in industry. The CSE community can greatly benefit from the ingrained expertise of dedicated software engineering experts, similar to how research in many disciplines (e.g., climate science) requires a multi-disciplinary approach, with experts in different fields of science and technology collaborating towards a common goal. As we look to the future, having software engineering experts as integral members of scientific computing teams and serving as the contact point for researchers applying computational methods for their research must be on the forefront of our minds, along with the cultivation of a diverse community which can fuel innovation for next-generation efforts.

To move in this direction, we have identified three key challenge areas where progress could enable more effective scientific software innovation, thus supporting software stewardship efforts: (1) expanding the diversity of the overall workforce, (2) training and mentorship of both new candidates and the existing workforce, and (3) retention of the workforce through established career growth and development opportunities.

Expanding Workforce Diversity

Diversity is one of the leading contributors to innovation in the workplace. In a recent study on the effect of diversity on industry, firms that prioritize diversity, particularly at a leadership level, are "45% likelier to report that their firm's market share grew over the previous year and 70%

[1] https://us-rse.org

likelier to report that the firm captured a new market" (Hewlett et al. 2013). Conversely, companies that lack diversity are far more common and result in women and people of color frequently being silenced or ignored. In their 2020 analysis, Hofstra et al. find that overall, underrepresented groups produce higher rates of scientific novelty, but these innovations are regularly devalued and discounted (Hofstra 2020).

While some institutions actively recruit RSEs, for others (who also rely on them), current recruiting for research software engineers is largely non-existent. It is done to some degree via professional organizations or targeted institutions (e.g., through career fairs at large, ivy league schools). In a study on recruiting women and minorities in radiation oncology, "the ultimate pursuit [of the field] begins with a student's initial exposure to the specialty … Unfortunately, the … probability of any student being exposed [to the field] is relatively low" (Suneja et al. 2019). Frequently, new researchers are recruited through employee referrals (e.g., the "old boys' club") or through existing strategic partnerships with a select set of universities. Indeed, according to their 2003 study, Breaugh et al. find that employee referrals received a higher percentage of job offers than those recruited from colleges or job fairs (Breaugh 2003). Employees will tend to refer candidates who act and think like themselves, or, "the 'autocatalytic character of elite production' as one generation of star scientist is recruiting the next generation of star scientists and providing them with guidance so making it more likely that those scientists will be part of the elite themselves one day" (Böhm & Phillips 2015). In other words, existing members of the workforce will gravitate towards those similar to themselves. These recruitment practices contribute to the lack of diversity in the workplace.

Of course, while diversity concerns can only partially be addressed by changes to the recruitment process, getting diverse candidates in the door helps present new opportunities for larger cultural change. To better encourage a diverse set of people to work on research software, we need to allocate resources towards recruiting underrepresented groups (e.g., diversity in race, gender, ethnicity, abilities, neurodiversity, etc.) and conduct outreach at a wider swath of universities, and not just where we have existing strategic partnerships. Several of the authors of this response have witnessed this lack of diversity firsthand. For example we have teams composed primarily of graduates of a single university, and often even from the same advisor, or entire working groups of primarily one demographic. While there is merit in using a trusted pipeline for finding quality candidates, this siloed approach excludes innovative, creative researchers who have not had the privilege of being a part of these same recruiting pipelines. To this end, we recommend that the DOE change its recruitment practices to include: implementing outreach programs geared towards candidates from underrepresented universities, auditing existing job postings to promote more inclusive language and a broader candidate pool on future postings, and offering internships to members of underrepresented groups, particularly through proactive engagement with community colleges. In addition, DOE needs to address cultural issues related to equity and inclusion, in part so that the diverse gains made in recruiting are retained, but these require institution-specific cultural studies, which are outside the scope of this response.

Training and Mentorship

Research software engineers have unique needs for training as compared to domain science experts (Milewicz & Mundt 2021). These needs include consistent access to training on new and existing technologies, such as best practices for coding, testing, performance, portability, etc., as well as access to interdisciplinary mentorship networks, longer term mentor/mentee relationships, and a focus on soft skill development, needs which are currently met for domain scientists but often are not for RSEs.

By the very nature of the position, RSEs are generally more fragmented in their workload than their scientific research counterparts (Katz et al. 2019). Frequently RSEs are members on several projects and are seen only as valuable as the skills they bring, rather than as a staff member to be grown and encouraged. As a result, time is infrequently allocated to RSEs for professional development.

As of now, there is no overarching Department of Energy (DOE) supported program geared towards dedicated training and mentorship efforts for research software engineering within the scientific domain. An individual institution may have an internal mentorship program or workshops, but these programs lack standardization of content and networks in order to broaden access to information and training opportunities. Particularly for members of underrepresented groups, access to mentorship, formalized training, and funding for development opportunities have been shown to be barriers to entry into the broader research community (Kumwenda et al. 2017), and consequently would also be barriers to entry into the RSE community.

From a software stewardship perspective, the lack of cross-discipline mentoring networks affects not only RSEs, but the larger CSE community. Interdisciplinary work is of high importance as collaborative software development between those of different backgrounds to achieve a common goal has become mainstream (Burnell et al. 2002). To address this challenge, DOE could standardize and fund mentorships and targeted training programs across the DOE. The 2021 NSF CSSI call for proposals (https://www.nsf.gov/pubs/2021/nsf21617/nsf21617.htm) requires respondents to include a "CI Professional Mentoring and/or Professional Development Plan" to address this. DOE could add a similar requirement to any solicitations that focus on software development and maintenance.

Retention through Career Growth and Development

The next challenge is the lack of dedicated time and funding for individual development and career growth opportunities. Particularly in the case of RSEs, their workload is already split between multiple projects, and they are relying on the direct funds of those projects. If they are fortunate enough to have access to overhead funds, as is the case at some national labs, universities, and private institutions, most of these funds are normally allocated for activities to further develop tools and utilities, as opposed to personal development.

The software engineering profession, as an example, has a relatively low retention rate. A 2010 study found that one-third of new hires across all industry software engineers were likely to leave an organization within 2 years (Stein & Christiansen 2010), and in 2016, the tenure of software developers in large tech companies ranged from 1.5 to 2.3 years (Dice Insights 2016). Given the longevity of many scientific research initiatives, the steep learning curve for new developers entering a scientific software project, and time from science to usable software, we cannot afford a turnover rate that is even close to that of industry. We recognize, though, that some turnover, particularly between institutions that develop and maintain research software, is positive and is one of the means by which best practices are developed and transferred. To retain our workforce, we must provide ample growth and development opportunities to compete with industry. Research is a long-term process, and thus the importance of retaining RSEs is essential to scientific software success. This can be achieved via competitive salaries, clearly defined career growth to include all roles (not just those who publish regularly) (Katz et al. 2019, Katz et al. 2021), and prolonged stability. Additionally, there must be dedicated time committed to everyone to pursue their own personal development as well as opportunities to explore new work paths and styles. With respect to software engineering, the only way to stay abreast of the newest technology and changes in standards and best practices is to engage in constant skill development activities.

Conclusion

In this response, we have argued that research software engineering is one of the disciplines that is essential for modern science, which depends on modern software, and that this discipline requires dedicated practitioners (research software engineers or RSEs). We've identified three challenges related to RSEs that need to be addressed in order to increase the quality of newly developed software and properly maintained existing software, and the quality of the science and engineering that depends on it. These challenges are the diversity of the RSE community, training of RSEs, and established career paths for RSEs. To address diversity concerns, we suggest changes in recruitment policies such as outreach to underrepresented universities, adjustment to language in future job postings, and internships targeted at underrepresented groups. To address training concerns, we recommend standardized mentorship and training programs, as well as focus on development activities in future DOE funding calls. Finally, to address retention concerns, we recommend mirroring industry practices in terms of competitive salaries, clearly defined career growth, and prolonged stability for RSEs. If we wish to continue to develop innovative scientific software that serves the DOE's mission, we must ensure the robustness, stability, and diversity of those who support our portion of the software ecosystem.

Please direct any questions or comments about this RFI response to Miranda Mundt (email: mmundt@sandia.gov).

Acknowledgements

- Contributions to this response were made by a member of Sandia National Laboratories, which is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DENA0003525.
  (Sandia Review ID: SAND2021-15647 O)
- Contributions to this response were made by a member of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.

References

- Böhm, L., & Phillips, R. (2015). Recruiting and Retaining the Best Scientists: A Researcher's Perspective. *IUP Journal of Organizational Behavior*, *14*(2), 7-27.
- Suneja, G., Mattes, M. D., Vega, R. B. M., Escorcia, F. E., Lawton, C., Greenberger, J., ... & Siker, M. (2020). Pathways for recruiting and retaining women and underrepresented minority clinicians and physician scientists into the radiation oncology workforce: a summary of the 2019 ASTRO/NCI diversity symposium session at the ASTRO Annual Meeting. *Advances in Radiation Oncology*, *5*(5), 798-803.
- Hewlett, S. A., Marshall, M., & Sherbin, L. (2013). How diversity can drive innovation. *Harvard business review*, *91*(12), 30-30.
- Breaugh, J. A., Greising, L. A., Taggart, J. W., & Chen, H. (2003). The relationship of recruiting sources and pre‐hire outcomes: Examination of yield ratios and applicant quality. *Journal of Applied Social Psychology*, *33*(11), 2267-2287.
- Hofstra, B., Kulkarni, V. V., Galvez, S. M. N., He, B., Jurafsky, D., & McFarland, D. A. (2020). The diversity–innovation paradox in science. *Proceedings of the National Academy of Sciences*, *117*(17), 9284-9291.
- Milewicz, R., & Mundt, M. (2021). An Exploration of the Mentorship Needs of Research Software Engineers. *arXiv preprint arXiv:2110.02251*.
- Katz, D. S., McHenry, K., Lee, J. S. (2021, September). Senior Level RSE Career Paths (with an S), presentation at SeptembRSE, the Fifth Conference of Research Software Engineers, 10.5281/zenodo.5531839.
- Katz, D. S., McHenry, K., Reinking, C., & Haines, R. (2019, May). Research software development & management in universities: case studies from Manchester's RSDS group, Illinois' NCSA, and Notre Dame's CRC. In *2019 IEEE/ACM 14th International Workshop on Software Engineering for Science (SE4Science)* (pp. 17-24). IEEE.
- Kumwenda, S., El Hadji, A. N., Orondo, P. W., William, P., Oyinlola, L., Bongo, G. N., & Chiwona, B. (2017). Challenges facing young African scientists in their research careers: A qualitative exploratory study. *Malawi Medical Journal*, *29*(1), 1-4.
- Burnell, L. J., Priest, J. W., & Durrett, J. B. (2002). Teaching distributed multidisciplinary software development. *IEEE software*, *19*(5), 86-93.
- Stein, M., & Christiansen, L. (2010). *Successful onboarding*. McGraw-Hill Professional Publishing.
- Dice Insights. (2016) *How long do tech pros stay in their jobs?* https://insights.dice.com/2016/07/08/how-long-do-tech-pros-stay-in-their-jobs/