I contribute my time to service committees, and when I see an opportunity where I could make a substantive change to improve the community, I volunteer for leadership positions. I serve on program committees for top international conferences in Software Engineering (e.g. ICSE, ASE, FSE, ISSTA) and support the organization of these and other conferences by co-chairing committees. I have been recognized as a "distinguished reviewer" on five different highly respected program committees, including most recently at ICSE 2022, with a citation for leading and engaging in the longest discussion on the committee. I review for top journals in my field (e.g. ACM's TOSEM, IEEE's TSE and Springer's EMSE), have served on a DOE funding panel, and have served on 8 NSF panels across several divisions (my home of CCF along with CNS, OAC and SBIR). I have served for three years on the CRA-E undergraduate award committee, and have volunteered time outside of the review period to help improve the review process. At Khoury, I have served on the PhD admissions committee since 2020, and am co-chairing the committee this year. Within Northeastern more broadly, I serve as Khoury's representative on the Cadre of University Marshalls.

Northeastern Leadership and Service

PhD CS Admissions: Between 2020-2022, I served on the PhD CS admissions committee (co-chaired by Daniel and Alina), and as of 2022-2023, am co-chair of the committee with Rob Platt. My primary objective as co-chair of the PhD CS admissions committee is to evolve and improve our college's process for PhD admissions, and to create artifacts that will allow my (eventual) successor to easily build upon our successes. My attention has been on improving the efficiency of our process, while also improving the diversity of our application pool and admitted students.

While serving on the committee, I noticed that the existing processes were struggling to meet the growth in students applying and being admitted – not to mention the growth in the number of faculty recruiting PhD students. Under the prior process, each application received two reviews from the committee, ranking each applicant's "overall merit" on a scale of 1-3. Digesting an entire application, writing a review, and compressing those thoughts into a single score is time consuming — especially when not familiar with the applicant's research area. Moreover, ensuring uniform ranking across reviewers requires extensive training, and even more time on behalf of the committee and chairs. For Fall 2023 admissions, applications were up by 62% compared to Fall 2022 (from 617 to 1,006), and we were extremely pleased to have taken the time to design and implement a new, more scalable process.

Working with the committee, we designed and implemented a new review process that focuses on fast, objective reviews that can be provided by a wider pool of reviewers. Rob and I focused our review process around two central objectives, 1) to help faculty quickly identify applicants who are interested in working with them, and 2) to help faculty quickly search and filter over those applicants based on different criteria. Rather than rank candidates on "overall merit," this year's review process aims to capture standardized information about each applicant such as whether they have relevant prior work experience, their topics of interest, and faculty who may overlap in interests. Given that the new review form is more objective (and hence, reviews should be more uniform across reviewers), we reduced the number of reviewers per-applicant (to one), and assigned faculty on the committee as "discussion leads" for applicants, responsible for spot-checking the quality of reviews. This process is largely implemented within HotCRP, and I have been maintaining a timeline of the various activities that I've performed so that future co-chairs can quickly understand the role. Feedback thus far from the committee and faculty has been generally positive, and at the end of the admissions cycle we will distribute a survey to help determine how to further refine the process.

Towards improving the diversity of our PhD application pool and accepted students, in the 2021 and 2022 application cycles I proposed and implemented a pre-application mentorship program. Applicants were matched with current PhD students, who provided feedback on personal statements and met 1:1 on

Zoom to answer questions and provide additional feedback. In each year, we provided mentorship to 30-40 applicants. For this year's 2023 application cycle, I evolved this program to reach more students, creating virtual info and Q&A sessions. In total, we organized eight info sessions and 225 applicants attended. Each info session was led by faculty from different areas: one focusing on URGs, and the others organized by research area. I worked with the marketing staff to publicize these events, and with the admissions staff to respond to procedural questions from applicants. I created a slide deck that provided an overview of the program and answers to frequently asked questions, and encouraged the faculty leading each info session to use these slides as-is or personalize them. At the end of the application season, I will analyze the impact of this program by matching info session attendees against PhD applications, using this feedback to improve the program next year.

CAREER Club: In 2017, Abhi Shelat created an informal mentorship program to support faculty applying for NSF CAREER awards. I joined Khoury in 2020, and having previously received a CAREER award, volunteered to participate in the CAREER club as a mentor every year since. In Fall 2022, I took over the faculty lead role of the CAREER club, working with Jane and Andrea to refine the organization of the mentoring program. This year's CAREER club features much more structure to help applicants develop and refine their project ideas, and I am very grateful to Jane and Andrea for their organizational support. I have discussed the structure of our CAREER club with colleagues at other institutions, generally receiving the response of "wow, I wish that we had that!" Over the coming years, I look forward to sharing artifacts from our mentoring program (e.g. syllabus and notes) publicly, creating resources that can be adopted more broadly within Northeastern and elsewhere.

National and International Professional Leadership

Student Mentoring Workshops: I have co-organized NSF-supported undergraduate mentoring workshops at ICSE (2022) and at SPLASH (2017, 2018, 2019, 2020). These workshops focus on broadening the participation of underrepresented groups in PL and SE research. The workshop generally has gender parity, with a significant representation of underrepresented minorities. As a white, cis male benefiting from many privileges in my academic career, my goal has been to stand behind my (diverse) co-organizers in order to execute their vision. Generally, my activities have focused on marketing, scheduling, and communicating details with participants, as opposed to setting the goals or format of the workshop. For example: at the ICSE 2022 mentoring workshop, my co-organizers proposed that we restructure the program from workshop presentations with Q&A to be exclusively centered around 1:1 and small group mentoring, with a social outing for mentors and mentees. With my co-organizers, we brainstormed the list of mentors to invite. Then, I did most of the remaining work to make this event happen: to take that list of names and dig up the email addresses for those (roughly 100) community members, to reach out to each target mentor, schedule the meetings and publicize the events.

Artifact Evaluation: An increasing trend in CS conferences is to organize artifact evaluation processes to recognize authors' efforts towards ensuring that their tools and datasets are available and reusable, and moreover, to integrate these artifacts into our publication process. SIGSOFT's artifact evaluation process started at ESEC/FSE in 2011, and has now spread to become commonplace at most of our conferences. In the intervening time, we have learned a lot about and refined the process. However, we are still not done: recent studies have shown that there is confusion on the part of both reviewers and authors [3], [4]. Authors are uncertain of the effort needed to prepare an artifact that passes the muster of the artifact review committee. Through these studies, reviewers also report that without clear criteria for acceptance, it is difficult to calibrate reviews and conduct the process fairly. I personally witnessed this mismatch with my students' recent ICSE 2022 artifact for our fuzzer CONFETTI [5]: The experimental evaluation requires 300 4-core VMs for 24 hours. We provided the reviewers with several options for evaluating this artifact (to run a shorter evaluation, to inspect and plot the raw data generated by our experiments, or to run the complete experiment using our hardware). One reviewer advocated for rejecting this artifact on the basis

that the README was too long, and that this was too much information/too complicated.

To encourage a co-mingling of ideas about artifact evaluation between the PL and SE communities, I served as the PLDI 2020 Artifact Evaluation co-chair. Working with an ad-hoc group of colleagues in SE on this topic, I also co-chaired a SIGSOFT Executive Committee Working Group on Open Science and Artifact Evaluation. One recent output of this working group was a retrospective evaluation of artifacts published in SE venues over the past 10 years [1]. We found that the current reward mechanism for artifacts (one or several badges displayed on the top of an article) does not significantly impact the visibility of articles through citations. We also found that, particularly for research that requires large, cloud-based evaluations (like our CONFETTI work), authors struggle to create easily reproducible artifacts. My current research project on automating reproducibility for software experiments aims to create new processes and tools for constructing large-scale software experiments [2]. The core idea of this project is to ensure that software experiments are reproducible and portable from their inception by creating a reusable ecosystem of open source components.

Conference Organization: Aside from mentoring workshops and artifact evaluation, I have served in a variety of leadership roles on organizing committees of software engineering conferences. My overall philosophy for these roles is to 1) find a role where I can bring a unique skillset, 2) execute and improve the existing process, and 3) create reusable artifacts that future organizers can benefit from. For example: in the process of organizing the virtualization for ICSE and ISSTA in 2020, I noticed that much of the steps could be automated by software, and co-founded an open source project that was subsequently used to support multiple iterations of conferences across different communities [1]. Most recently, I am co-chairing the ISSTA tool demonstrations track, and organizing a workshop on software engineering education at ICSE.

Referenced Publications

[5] J. Kukucka, L. G. Ganchinho de Pina, P. Ammann, and J. Bell, *CONFETTI: Amplifying concolic guidance for fuzzers (github reusable artifact)*, https://github.com/neu-se/confetti-artifact, Jan. 2022.

[1] S. Winter, C. S. Timperley, B. Hermann, J. Cito, J. Bell, M. Hilton, and D. Beyer, "A retrospective study of one decade of artifact evaluations," in *Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, 2022.

[2] E. Barr, J. Bell, M. Hilton, C. Timperley, and S. Mechtaev, "Continuously accelerating research," in *Proceedings of the 2023 ACM/IEEE International Conference on Software Engineering, New Ideas and Emerging Results*, 2023.

External References

[3] C. S. Timperley, L. Herckis, C. L. Goues, and M. Hilton, "Understanding and improving artifact sharing in software engineering research," *Empir. Softw. Eng.*, vol. 26, no. 4, p. 67, 2021.

[4] B. Hermann, S. Winter, and J. Siegmund, "Community expectations for research artifacts and evaluation processes," in *Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, 2020, pp. 469–480.

---

[1]https://github.com/clowdr-app/clowdr