



# Advancing Reproducibility in Parallel and Distributed Systems Research

**Manish Parashar**, University of Utah

*This installment of Computer's series highlighting the work published in IEEE Computer Society journals comes from IEEE Transactions on Parallel and Distributed Systems.*

**W**hile reproducibility has always been fundamental to scientific and technical research, it has taken on an increased urgency, impacting society's trust in research results.<sup>1</sup> Recognizing both this importance and urgency, *IEEE Transactions on Parallel and Distributed Systems (TPDS)* remains committed to enabling reproducible research through transparency and the availability and

potential reuse of the code and data associated with its publications.

Specifically, for a few years, the TPDS Reproducibility Initiative<sup>2,3</sup> has been exploring postpublication peer review of the code associated with published articles. Authors who have published in TPDS can make their published work more reproducible and earn a reproducibility badge by submitting their associated code for postpublication peer review. To date, this pilot has largely focused

on two badges, Code Available and Code Reviewed, and has successfully badged more than a dozen articles.

Ensuring reproducibility has also taken on additional complexity due to the increasingly central roles played by computation and data.<sup>4</sup> While TPDS's goal has always been to include badges for reproducing research results using the code and/or data provided, the nature of research in the parallel and distributed systems covered by TPDS makes it challenging to evaluate the code and data for reproducibility. This is because such an evaluation may require access to specific hardware, system architectures and scales, operating system configurations, and so



on, which may not be feasible or practical. While technologies such as containerization can help address some of these challenges, providing reviewers access to an execution environment in

for badging.<sup>6</sup> TPDS is exploring and piloting solutions to some of these issues.

For example, TPDS has been analyzing the artifact submission and review processes and improving them. It has

most importantly, authors for submitting artifacts along with their articles. **C**

Authors who have published in TPDS can make their published work more reproducible and earn a reproducibility badge by submitting their associated code for postpublication peer review.

which they can effectively reproduce the full range of research published in TPDS—including, for example, research involving very-large-scale parallelism, large data volumes, and low-level system/middleware services—is typically infeasible.

Consequently, TPDS piloted an alternate approach where members of the community can submit short, supplemental “critique” articles that present their experiences in reproducing published results using the artifacts and/or evaluations or experiences with published artifacts. These supplemental article submissions are reviewed; if accepted, they are linked to the original publication and citable, serving to help validate the reproducibility of the original work. This approach has been implemented in two special sections consisting of a primary article and multiple critique articles that reproduce the results of the primary one.<sup>5</sup> TPDS is also partnering with national providers of computing resources to provide resources to support the evaluation of artifacts.

Achieving reproducibility also requires deploying the necessary support infrastructure, including enabling the submission and evaluation of data, code, and other artifacts across the breadth of topics covered; the creation of integrated workflows for processing and publishing submissions; and complying with emerging standards, for example,

been updating the guidelines provided to authors, which help authors with packaging, documenting, and submitting their artifacts and assist reviewers in assessing them. Additionally, TPDS recently named an associate editor-in-chief for reproducibility and is appointing a Reproducibility Editorial Board and Reproducibility Review Board to oversee the badging process, support the evaluation of artifacts, and continuously improve this process as well as share best practices. TPDS will also leverage a conference management system to support the artifact evaluation and badging workflows. Finally, it will be expanding the set of repositories that authors can use to submit their artifacts to include Dryad, Figshare, Harvard Dataverse, and Zenodo (including GitHub via Zenodo), in addition to CodeOcean and DataPort.

**W**e plan to continue to push forward on these important issues and will share any lessons learned. You can always find the latest information about the TPDS reproducibility initiative online at <https://www.computer.org/digital-library/journals/td/tpds-reproducibility-initiative>.<sup>2</sup> I would like to thank the TPDS Editorial Board and reviewers for their efforts in moving this initiative forward; IEEE Computer Society for its support; and,

## REFERENCES

1. M. Parashar, “Leveraging the National Academies’ reproducibility and replication in science report to advance reproducibility in publishing,” *Harvard Data Sci. Rev.*, vol. 2, no. 4, 2020, doi: 10.1162/99608f92.b69d3134.
2. “IEEE Transactions on Parallel and Distributed Systems (TPDS) Reproducibility Initiative,” IEEE Computer Society, 2020. <https://www.computer.org/digital-library/journals/td/tpds-reproducibility-initiative>
3. M. Parashar, “The reproducibility initiative,” *Computer*, vol. 52, no. 11, pp. 7–8, 2019, doi: 10.1109/MC.2019.2935265.
4. “Report from the Ad Hoc Committee on Open Science and Reproducibility,” IEEE Computer Society, 2020. <https://www.computer.org/volunteering/boards-and-committees/Open-Science-Reproducibility>
5. M. Parashar, “EiC editorial: Advancing reproducibility in parallel and distributed systems research,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 33, no. 9, pp. 2010–2010, 2022, doi: 10.1109/TPDS.2021.3137871.
6. “Taxonomy, definitions, and recognition badging scheme working group,” NISO, Baltimore, MD, USA, 2021. [Online]. Available: <https://www.niso.org/standards-committees/reproducibility-badging>

**MANISH PARASHAR** is the director of the Scientific Computing and Imaging Institute, the chair in Computational Science and Engineering, a professor in the School of Computing at the University of Utah, Salt Lake City, Utah, USA, and the editor in chief of *IEEE Transactions on Parallel and Distributed Systems*. Contact him at [parashar@ieee.org](mailto:parashar@ieee.org).