Reproducibility Challenge Webinar

SC21 Student Cluster Competition

Le Mai Weakley, PhD
SC21 Reproducibility Challenge Chair

Announcements

- Checkout the Google FAQ/Announcement doc! (Teams should have access)
- Accept the invite to the VSCC/SCC Slack channel (SC21 Students@SC workspace).
  - To get an invitation email student-cluster-competition@info.supercomputing.org

Today's agenda:

- Introductions
- Speakers
- Reproducibility Challenge Recommendations
- Q&A

Speakers

Michela Taufer

Founder of Reproducibility Effort at SC

Stephen Lien Harrell

SC21 Reproducibility Journal Issue Chair

Carlos Maltzahn

SC21 Reproducibility Chair

Ankit Srivastava

Author of Reproducibility Challenge paper

What is the Reproducibility Challenge?

- One of the applications in the SCC/VSCC (in addition to Cardiod, Quantum Espresso, mystery app and the benchmarks)
- Teams will attempt to reproduce part of the results of an accepted paper from prior SC
- Scores from the challenge are based on written reports
- Submission of Digital Artifact
- Highest scoring reports along with application author invited to do special journal issue

Reproducibility in HPC and Our Everyday Lives

Illustration by Kallum Best

"Do we trust the experimental results published in, e.g., the Nature journal? "
Yes, because …

- **The experiment was reproduced multiple times converging to the same scientific conclusions**

Image sources:
https://www.alfadispenser.com/2018/05/30/repeatability-vs-reproducibility-a-little-talk/
https://undsci.berkeley.edu/article/0_0_0/howscienceworks_17
http://www.lateralmag.com/articles/issue-28/editorial/replication-crisis

Reproducibility in HPC and Our Everyday Lives

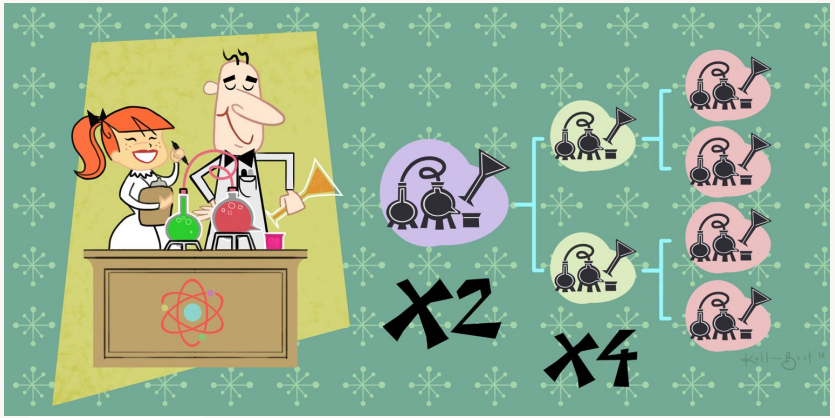
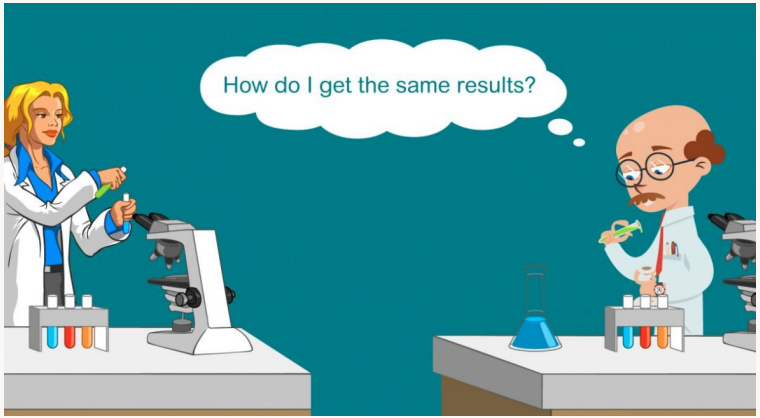


Illustration by Kallum Best



"Do we trust the experimental results published in, e.g., the Nature journal? "
Yes, because …

- The experiment was reproduced multiple times converging to the same scientific conclusions
- **The experimental process was documented step by step**

Image sources:
https://www.alfadispenser.com/2018/05/30/repeatability-vs-reproducibility-a-little-talk/
https://undsci.berkeley.edu/article/0_0_0/howscienceworks_17
http://www.lateralmag.com/articles/issue-28/editorial/replication-crisis

Reproducibility in HPC and Our Everyday Lives

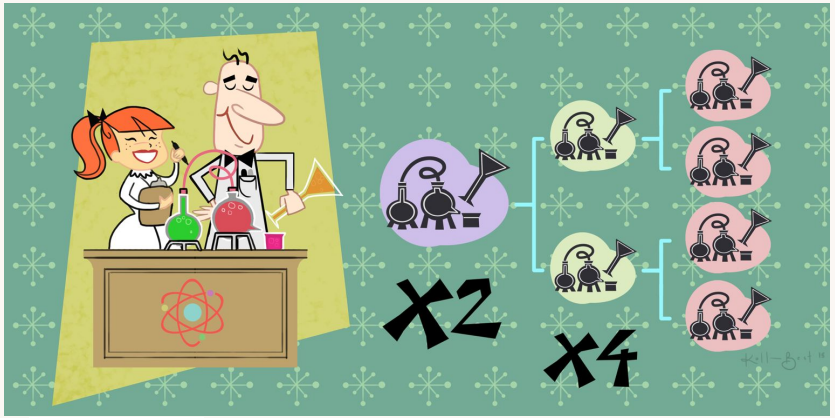
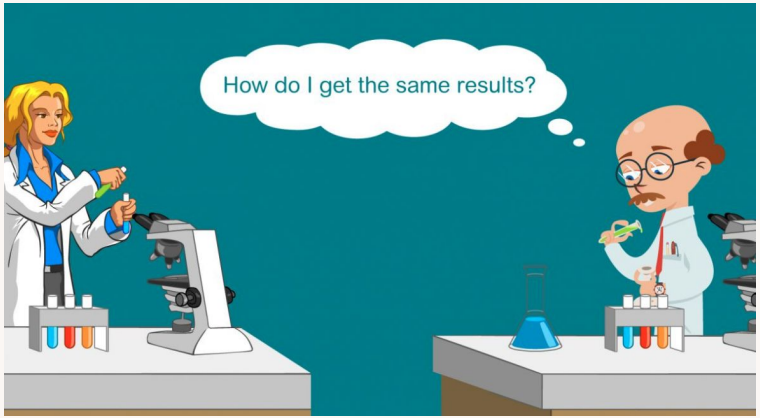


Illustration by Kallum Best



"Do we trust the experimental results published in, e.g., the Nature journal? "
Yes, because …

- The experiment was reproduced multiple times converging to the same scientific conclusions
- The process was documented step by step

"What is one of the advantages of a reproducible and well documented experiment?"

- **Everyone can reproduce the experiment**
- **Anyone can build new science by leveraging the reproduced results**

Image sources:
https://www.alfadispenser.com/2018/05/30/repeatability-vs-reproducibility-a-little-talk/
https://undsci.berkeley.edu/article/0_0_0/howscienceworks_17
http://www.lateralmag.com/articles/issue-28/editorial/replication-crisis

Reproducibility in HPC and Our Everyday Lives

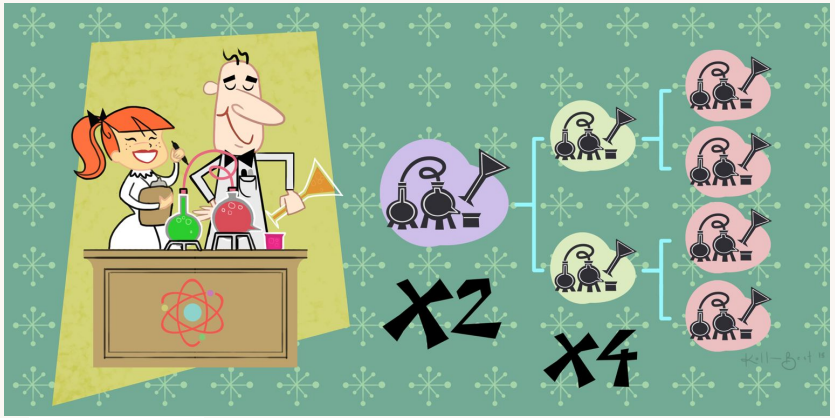
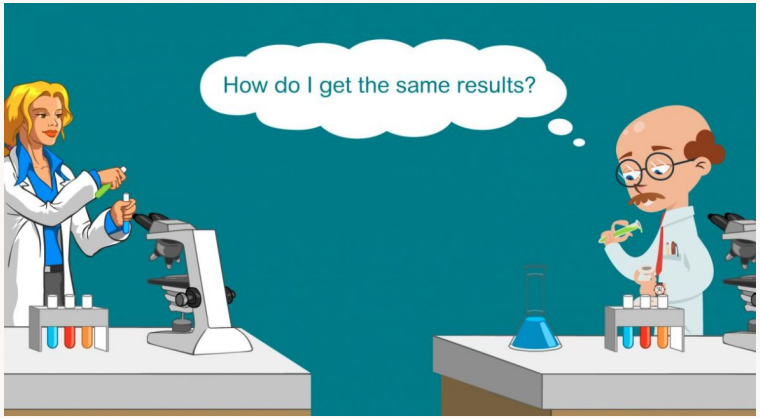


Illustration by Kallum Best



"Do we trust the experimental results published in, e.g., the Nature journal? "
Yes, because …

- The experiment was reproduced multiple times converging to the same scientific conclusions
- The process was documented step by step

"What is one of the advantages of a reproducible and well documented experiment?"

- Everyone can reproduce the experiment
- Anyone can build new science by leveraging the reproduced results

What if we replace "*experimental results*" with "computational results"?

Image sources:
https://www.alfadispenser.com/2018/05/30/repeatability-vs-reproducibility-a-little-talk/
https://undsci.berkeley.edu/article/0_0_0/howscienceworks_17
http://www.lateralmag.com/articles/issue-28/editorial/replication-crisis

Lifecycle of Reproducibility at SC

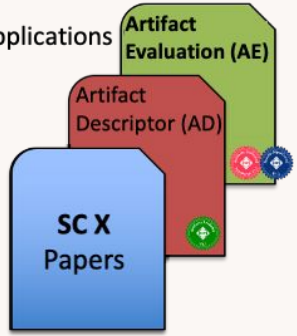
Technical Program @ SC X

Select BP/BSP/BRA candidates

Select badge applications

Review AD/AE

Review papers



Lifecycle of Reproducibility at SC

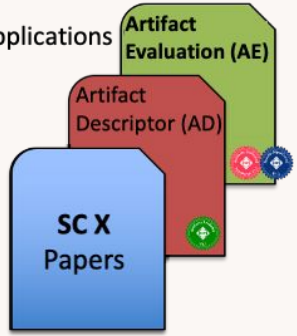
Technical Program @ SC X

Select BP/BSP/BRA candidates

Select badge applications

Review AD/AE

Review papers



Select one (1) **SC X** paper for **SC X+1** SCC

Lifecycle of Reproducibility at SC

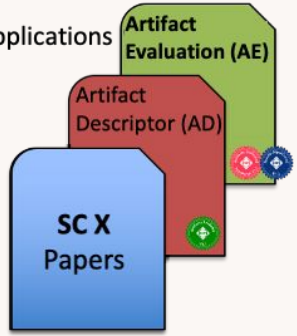
Technical Program @ SC X

Select BP/BSP/BRA candidates

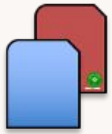
Select badge applications

Review AD/AE

Review papers



Select one (1) **SC X** paper for **SC X+1** SCC



Generate replication benchmark for diverse set of HPC platforms

Lifecycle of Reproducibility at SC

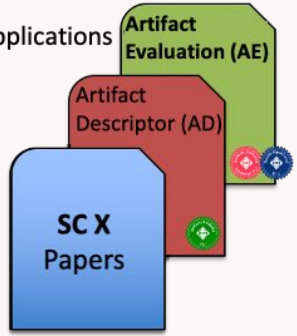
Technical Program @ SC X

Select BP/BSP/BRA candidates

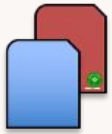
Select badge applications

Review AD/AE

Review papers



Select one (1) **SC X** paper for **SC X+1** SCC



Generate replication benchmark for diverse set of HPC platforms

Student Cluster Competition @ SC X+1

Partner with vendors

Build a cluster

Test performance benchmarks

Replicate **SC X** Paper

Generate replication reports

Lifecycle of Reproducibility at SC

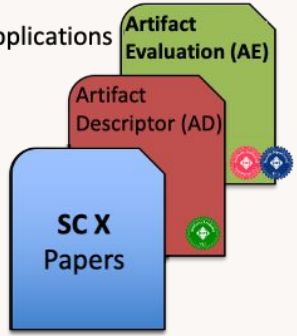
Technical Program @ SC X

Select BP/BSP/BRA candidates

Select badge applications

Review AD/AE

Review papers



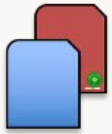
Select one (1) **SC X** paper for **SC X+1** SCC

Technical Program @ SC X+1

Assign badge to SC X paper

Give SIGHPC certificate to SC X paper authors

Present Journal SI with SCC reports from SCC @ SC X-1



Generate replication benchmark for diverse set of HPC platforms

Student Cluster Competition @ SC X+1

Partner with vendors

Build a cluster

Test performance benchmarks

Replicate **SC X** Paper

Generate replication reports

Lifecycle of Reproducibility at SC

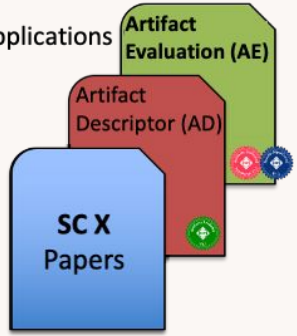
Technical Program @ SC X

Select BP/BSP/BRA candidates

Select badge applications

Review AD/AE

Review papers



Technical Program @ SC X+2

Review IEEE TPDS paper with SCC reports from SCC @ SC X+1

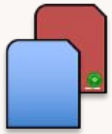
Select one (1) **SC X** paper for **SC X+1** SCC

Technical Program @ SC X+1

Assign badge to SC X paper

Give SIGHPC certificate to SC X paper authors

Present Journal SI with SCC reports from SCC @ SC X-1



Generate replication benchmark for diverse set of HPC platforms

Student Cluster Competition @ SC X+1

Partner with vendors

Build a cluster

Test performance benchmarks

Replicate **SC X** Paper

Generate replication reports

Strengthening Reproducibility for SC 21 and Beyond

August 26, 2021

SC's leadership in Reproducibility

The **SC steering committee approved the reproducibility initiative**. Authors of SC15 papers were **invited to submit an AD Appendix after the conference**: one paper did so, became the source for the **SC16 Student Cluster Competition Reproducibility Challenge** and the first SC paper to display an **ACM badge**.

Authors submitting to the SC16 conference could **optionally submit an AD Appendix**: nine authors submitted one, three were finalists, and one was selected to become the source for the SC17 Student Cluster Competition Reproducibility Challenge.

SC17 made the **AD Appendix a requirement to be considered for the Best Paper or Best Student Paper awards**. SC17 also **introduced the Computational Results Analysis (CRA) Appendix**. 40% of submitted and 50% of accepted papers included an AD appendix; nine submitted papers (six accepted) included a CRA Appendix.

SC18 extended the **option of submitting AD Appendices to Workshops and Posters**. **The CRA Appendix was renamed Artifact Evaluation (AE) Appendix**, and limited to four pages. AD Appendices were limited to 2 pages and remained optional (but required for consideration as Best Paper/Best Student Paper, and also Best Poster/Best Student Poster).

SC19: **AD Appendices were mandatory for all submissions**. AE Appendices were still optional, and both were **submitted via a standard form** in the conference submission system. **Three new subcommittees, with their chairs, were introduced** in support of the SC Reproducibility Initiative.

SC20 **expanded the Initiative to transparency and reproducibility to reflect scientific rigor through disclosure particularly in research involving AI**. The **AD Appendix was streamlined for reduced researcher burden and to align with open science principles**. An additional track undertook a **formal survey of community sentiment about SC reproducibility** with the objective of publishable results. The majority of survey participants who went through the AD/AE Appendices process expressed that they now think differently about theirs and others' research after having gone through the process.

SC21 Reproducibility Team Moves SC Substantially Forward

- Bring SC in line with best practices used by other conferences
  - AD/AE committee led by
    - Tanu Malik, Co-Chair (Assist. Prof., DePaul U), 2019 NSF CAREER awardee for her work on computational reproducibility and 2019 Fellow for Better Scientific Software (BSSw)
    - Anjo Vahldiek-Oberwagner, Co-Chair (Rsrch. Sci., Intel Labs), AE co-chair at OSDI'20 , AE committee member at SOSP'19 & USENIX Security'20
  - Clear separation of responsibilities of AD/AE committee and TPCs
    - AD review is a signal for TPCs
    - AE review of badge applications of accepted papers only

  - AE Infrastructure for authors *and* reviewers provided by Chameleon Cloud, CloudLab, XSEDE's JetStream 2, SDSC
    - Webinar for authors in June. Advertised through TPCs
    - Webinar for AE reviewers in July
    - Working on long-term hosting opportunities for publicly available artifacts

Reproducibility Initiative Chair
Carlos Maltzahn, University of California, Santa Cruz

Reproducibility Initiative Vice Chair
Ivo Jimenez, University of California, Santa Cruz

Reproducibility Challenge Chair
Le Mai Weakley, Indiana University

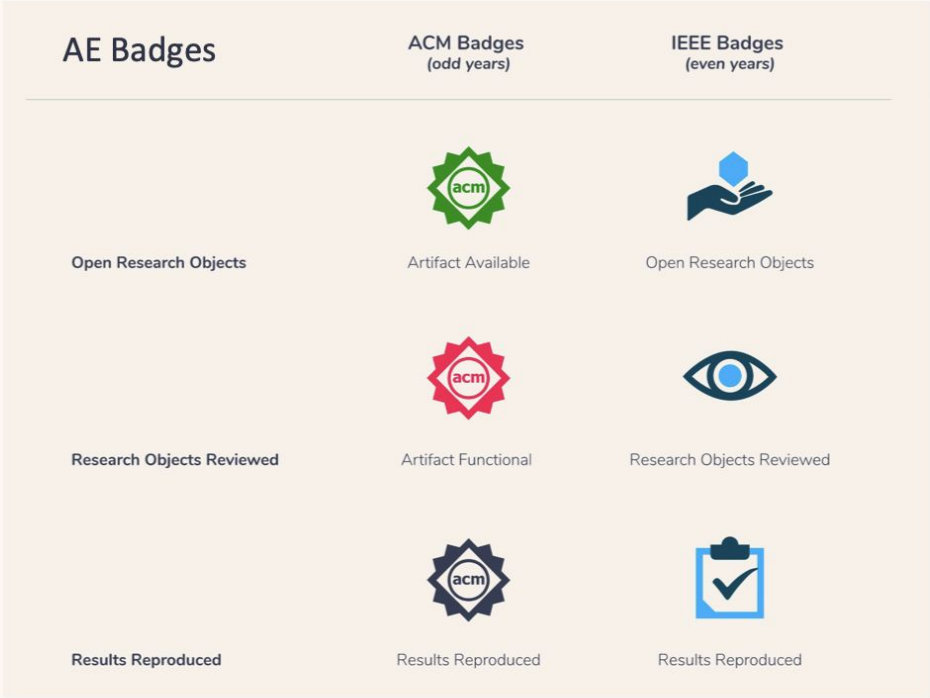
AD/AE Appendices Co-Chairs
Tanu Malik, DePaul University
Anjo Vahldiek-Oberwanger, Intel

Journal Special Issue Chair
Stephen Harrell, Texas Advanced Computing Center

Journal Special Issue Vice Chair
Scott A. Michael, Indiana University

SC21 Reproducibility Team Moves SC Substantially Forward

- Increase the types of badges and number of papers receiving them
  - Artifact Available, Artifact Functional, Artifact Evaluated
    - ACM and IEEE agreed on equivalent badges
  - Over 50% of submitted papers have applied for badges, almost all of these for all three
- Develop new processes that are "reproducible" for future SC conferences
  - Started briefing next year's Chair of Reproducibility Initative, Bilel Hadri
  - Recruit from AE leadership of best practices used by other conferences
  - Extensive documentation, monitoring of time required
  - BONUS: Streamlines Reproducibility selection for Student Cluster Competition



SC Best Reproducibility Advancement Award

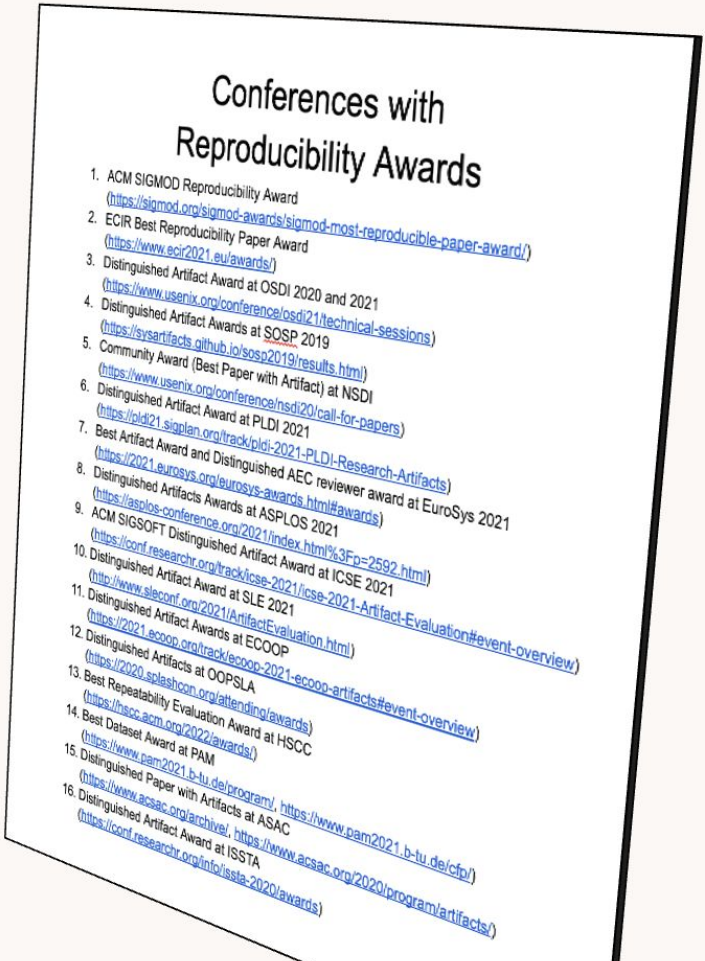
Approved by Steering Committee on 8/12/21

Why it is important to recognize contributions to reproducibility

- Highlight exemplars of best practices in reproducibility to move the community forward
- Evolve the criteria as the community moves forward

Why this should be an SC Conference Award

- Leverage this year's effort to advance best practices for future SC conferences
- Committee structure and AD/AE submission process makes this feasible and realistic



Conferences with Reproducibility Awards

1. ACM SIGMOD Reproducibility Award (https://sigmod.org/sigmod-awards/sigmod-most-reproducible-paper-award/)
2. ECIR Best Reproducibility Paper Award (https://www.ecir2021.eu/awards/)
3. Distinguished Artifact Award at OSDI 2020 and 2021 (https://www.usenix.org/conference/osdi21/technical-sessions)
4. Distinguished Artifact Awards at SOSP 2019 (https://sysartifacts.github.io/sosp2019/results.html)
5. Community Award (Best Paper with Artifact) at NSDI (https://www.usenix.org/conference/nsdi20/call-for-papers)
6. Distinguished Artifact Award at PLDI 2021 (https://pldi21.sigplan.org/track/pldi-2021-PLDI-Research-Artifacts)
7. Best Artifact Award and Distinguished AEC reviewer award at EuroSys 2021 (https://2021.eurosys.org/eurosys-awards.html#awards)
8. Distinguished Artifacts Awards at ASPLOS 2021 (https://asplos-conference.org/2021/index.html%3Fp=2592.html)
9. ACM SIGSOFT Distinguished Artifact Award at ICSE 2021 (https://conf.researchr.org/track/icse-2021/icse-2021-Artifact-Evaluation#event-overview)
10. Distinguished Artifact Award at SLE 2021 (http://www.sleconf.org/2021/ArtifactEvaluation.html)
11. Distinguished Artifact Awards at ECOOP (https://2021.ecoop.org/track/ecoop-2021-ecoop-artifacts#event-overview)
12. Distinguished Artifacts at OOPSLA (https://2020.splashcon.org/attending/awards)
13. Best Repeatability Evaluation Award at HSCC (https://hscc.acm.org/2022/awards/)
14. Best Dataset Award at PAM (https://www.pam2021.b-tu.de/program/, https://www.pam2021.b-tu.de/cfp/)
15. Distinguished Paper with Artifacts at ASAC (https://www.acsac.org/archive/, https://www.acsac.org/2020/program/artifacts/)
16. Distinguished Artifact Award at ISSTA (https://conf.researchr.org/info/issta-2020/awards)

Selecting the Best Reproducibility Advancement Winner

1. Self-Nomination by Authors (via submission form)
2. Nomination by (TPC+AD/AE) reviewers (subset of accepted papers)

3. List goes to Reproducibility Award Committee:
   - TPC Chairs + TPC members who reviewed papers that were nominated
   - AD/AE chairs + 2 AD/AE reviewers
   - Reproducibility Initiative Members (Reproducibility Challenge, Special Issue journal chairs)

4. Prune submissions down to 3-4 finalists

5. Each Finalist is assigned an external champion, i.e. an AD/AE reviewer who already reviewed the paper's/artifact's badge application and will present the artifact to the Reproducibility Award Committee

6. Select Winner

Speaker: Ankit Srivastava

Reproducibility Challenge Reports

What to put in it?

- Describe your architecture, hardware and software
- Describe your experimental setup
- Describe your experimental design (amount of runs, implementation details)
- Describe and plot outcomes of your experiments
- Compare your results to the paper's results
- Frame your report in terms of the challenge

Reproducibility Challenge Reports

Preparing for the challenge ahead of time

- Start early!
- Write what you can
  - Things that are getting firmed up (descriptions of infrastructure, experimental setup)
- Check out old reports
  - I will post some resources in the Google group soon
- Try out weak and strong scaling
  - Design the experiment
  - Plan for contingencies
  - Write up what you can when you get this settled
- Think and ask about differences that could contribute to differences in results
  - Be curious!

Questions?

Thank you for attending and good luck!

Title credits from Meet Me In St. Louis, an MGM movie 1944