Reproducibility and ACM
Transactions on Mathematical Software

Michael Heroux

Sandia National Laboratories

Reproducibility & Independent Verification Requirement

- In order to publish a paper: *Someone other than the authors must be able to reproduce the computational results.*

- Latitude in "reproduce":
  - Exactly the same numerical results?
  - Exactly the same runtime?
  - Close, in the opinion of an expert reviewer?

- What about:
  - Access to the same computing environment?
  - High end systems?

- Lots of challenges.

- But just the *expectation [threat]* can drive efforts…

Sandia National Laboratories

Fruits of the Threat

- **Source management tools:** In order to guarantee that results can be reproduced, the software must be preserved so that the exact version used to produce results is available at a later date.

- **Use of other standard tools and platforms:** In order to reduce the complexity of an environment, standard software libraries and computing environments will be helpful.

- **Documentation:** Independent verification requires that someone else understand how to use your software.

- **Source code standards:** Improves the ability of others to read your source code.

- **Testing:** Investment in greater testing makes sense because the software will be used by others.

- **High-quality software engineering environment:** If a research team is serious about producing high-quality, reproducible and verifiable results, it will want to invest in a high-quality SE environment to improve team efficiency.

Thank you for taking the time to consider our paper for your journal.

XXX has agreed to undergo the RCR process should the paper proceed far enough in the review process to qualify. ***To make this easier we have preserved the exact copy of the code used for the results (including additional code for generating detailed statistics that is not in the library version of the code).***

Reproducibility Terminology

V. Stodden, D. H. Bailey, J. Borwein, R. J. LeVeque, W. Rider, and W. Stein. 2013. Setting the Default to Reproducible: Reproducibility in Computational and Experimental Mathematics. (2013). http://icerm.brown.edu/html/programs/topical/tw12 5 rcem/icerm report.pdf

- **Reviewable Research**. The descriptions of the research methods can be independently assessed and the results judged credible. (This includes both traditional peer review and community review, and does not necessarily imply reproducibility.)

- **Replicable Research**. Tools are made available that would allow one to duplicate the results of the research, for example by running the authors' code to produce the plots shown in the publication. (Here tools might be limited in scope, e.g., only essential data or executables, and might only be made available to referees or only upon request.)

- **Confirmable Research**. The main conclusions of the research can be attained independently without the use of software provided by the author. (But using the complete description of algorithms and methodology provided in the publication and any supplementary materials.)

- **Auditable Research**. Sufficient records (including data and software) have been archived so that the research can be defended later if necessary or differences between independent confirmations resolved. The archive might be private, as with traditional laboratory notebooks.

- **Open or Reproducible Research.** Auditable research made openly available. This comprised well-documented and fully open code and data that are publicly available that would allow one to (a) fully audit the computational procedure, (b) replicate and also independently reproduce the results of the research, and (c) extend the results or apply the method to new problems.

- TOMS RCR Initiative: Referee Data.
- Why TOMS? Tradition of real software that others use.
- Two categories: Algorithms, Research.
- TOMS Algorithms Category:
  - Software Submitted with manuscript.
  - Both are thoroughly reviewe.
- TOMS Research Category:
  - Stronger: Previous implicit "real software" requirement is explicit.
  - New: Special designation for replicated results.

ACM TOMS Reproducible Computational Results (RCR) Process

- Submission: Optional (for now) RCR option.
- Standard reviewer assignment: Nothing changes.
- RCR reviewer assignment:
  – Concurrent with the first round of standard reviews
  – Known to and works with the authors during the RCR process.
- RCR process:
  – Multi-faceted approach.
- Publication:
  – Replicated Computational Results Designation.
  – The RCR referee acknowledged.
  – Review report appears with published manuscript.

Sandia National Laboratories

RCR Process

- Independent replication:
  - Transfer of or pointer to software given to RCR reviewer.
  - Guest account, access to software on author's system.
  - Detailed observation of the authors replicating the results.
- Review of computational results artifacts:
  - Results may be from a system that is no longer available.
  - Leadership class computing system.
  - In this situation:
    - Careful documentation of the process.
    - Software should have its own substantial verification process.

Sandia National Laboratories

Discussion

- How important is reproducibility to our communities? Why?

- Can IDEAS utilize reproducibility expectations to improve quality? How?