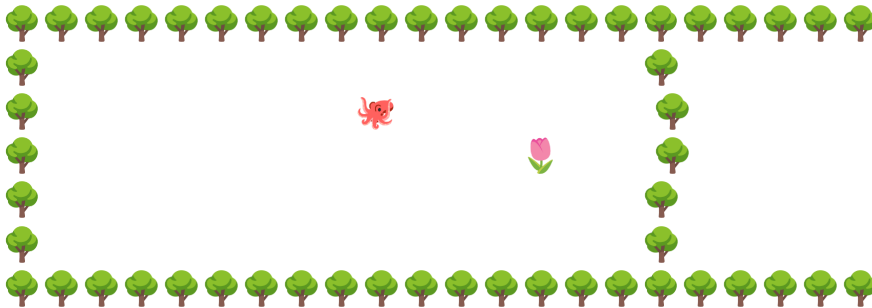


2021-06-24: Community conversations around publishing and scholarly communications
of reproducible research

Relevant links:

- EIG: <https://reproducibility.acm.org/>
- P-WG: <https://reproducibility.acm.org/working-groups/>
- “Taking the Pulse” blog: <https://reproducibility.acm.org/blog/>
- Slides for today:
<https://docs.google.com/presentation/d/1cq3b8QY3qMc-Ad2pcUFqoh9FXBjgePM9Ifj8VNczwI/edit?usp=sharing>
- [Blog post](#) and [notes](#) from previous conversation about reproducibility **solutions**
- [Blog post](#) and [notes](#) from previous conversation about reproducibility **principles**
- [Blog post](#) and [notes](#) from previous conversation about reproducibility **training**

Cursor Garden ↴ rest your cursor here when not writing



Name, Affiliation, Title, ACM member?:

- Vicky Rampin, NYU, Librarian for RDM + Reproducibility, ACM member
- Limor Peer, Yale University, Associate Director for Research at the Institution for Social and Policy Studies, ACM member
- Konrad Hinsén, CNRS, researcher. Not an ACM member. Co-EIC of [ReScience C](#), department editor for [CiSE](#)
- Juan J. Lastra-Díaz. [Universidad Nacional de Educación a Distancia \(UNED\)](#) ad-honorem researcher (Madrid, Spain). Past ACM member. Reproducibility Editor of [Information Systems](#) journal. Senior R&D Engineer at [Lectra](#).
- Hamed Alhoori, Northern Illinois University, ACM member
- Thomas Y. Chen, Academy for Mathematics, Science, and Engineering, Student, ACM Member

- Thu-Mai Christian, Odum Institute, UNC-Chapel Hill, Asst. Dir. for Archives, non-member
- Grigori Fursin, VP of MLOps, OctoML.ai / cTuning foundation, ACM member
- Giulia Guizzardi, eLife Innovation Officer, not ACM member
- Daniel S. Katz, Chief Scientist, NCSA, University of Illinois at Urbana-Champaign (also Associate Editor in Chief, Journal of Open Source Software), ACM member
- Jeremy Leipzig, Content Lead @ Truwl, Drexel PhD candidate, Not ACM member
- Philippe Bonnet, IT University of Copenhagen, ACM member
- Andrea Chiarelli, Senior Consultant (schol comms and open research), Research Consulting (UK), Not ACM member
- Jenny Muilenburg, Research Data Services Librarian, University of Washington, Not ACM member
- Nadica Miljković, Associate Professor at the University of Belgrade - School of Electrical Engineering, not ACM member
- Daniel Nüst, researcher at Institute for Geoinformatics, University of Münster, Reproducibility Chair at AGILE conference, Co-initiator of CODECHECK
- Michael Heroux, Senior Scientist at Sandia National Labs, , Director of Software for the US Exascale Computing Project, ACM Pubs Board member, co-chair of ACM New Publications committee, Chair of NSF ACCI working group on Reproducibility and sustainability

Add ACM efforts on reproducibility to this document:

https://docs.google.com/document/d/10mguVpdYM0OY2bCufT1nWQzQdBlz3_Zzkv2Vys7P2z8/edit

Reproducibility publishing guidelines, statements, and initiatives:

- American Journal of Political Science (AJPS) Verification Policy
<https://ajps.org/ajps-verification-policy/> see “[Guidelines for Preparing Replication Files](#)”
- RDA CURE-FAIR working group will publish (in 2022) standards-based guidelines for how to curate for publishing data and code that are reproducible and FAIR
<https://www.rd-alliance.org/groups/cure-fair-wg>
- Reproducibility guidelines, of Information Systems journal. (2020). Ongoing reproducibility initiative launched in 2016 to publish peer-reviewed “reproducible papers” co-authored by the reviewers. <https://doi.org/10.13140/RG.2.2.34277.22243>
- See also: Introducing three principles and the Reproduction Package, Krafczyk et al., <https://doi.org/10.1098/rsta.2020.0069>

Notes from our conversation:

- Questions from email:
 - What will be the ultimate checklist to publish a model that is deemed reproducible?
 - What semantics needs to be captured to be reproducible?
 - What would we need to describe the model and reproduce it
 - if you can't reproduce the models can we compare models in a meaningful way?
- [Poll everywhere replies / vote: see <https://docs.google.com/spreadsheets/d/1HwVBG41hZfaGwJqjAwTKFXDcnCEiusrVrFjZsDEvVoQ/edit#gid=85790258>]
- Cultural changes are key; people can catch up with technology
- Scaling problem: ability to prove that something can be calculated without fully reproducing (not having to do it all manually)
- Reproducibility editor - "Information Systems" journal
 - Guidelines on software and data and full detail of how to use in order to reproduce results previously published. Reviewers involved in acceptance of paper.
 - *(from chat: I think we should guide by example, than by imposing. At least as adopting technology as publishers; real guidance on how to do this would be helpful. Best practices, etc.)*
- "Reproducible builds" <https://reproducible-builds.org/> - verify the binary is running,
 - *(from chat: I am convinced we can fully automate reproducibility in the long run. The hard part is getting appropriate toolchains accepted. Evidence: Reproducible builds for building software. That's fundamentally the same problem as ours, and it's largely solved.)*
 - *(from chat: <http://james.howison.name/pubs/Howison-2014-RetractBitRottenPublications.pdf> might be the first time I heard about on-going automated reproducibility checking)*
- Proceedings of the VLDB (Very Large Data Bases) - promoting reproducibility; some observations:
 - cannot automate everything;
 - issues w/restrictions when using industry materials;
 - recommend but not enforce;
 - PDF should stay ("PDF is still our currency")

- *(from chat: One pragmatic way to handle PDFs: in e.g. arXiv:1905.04800 we reported on insights into continuous observations of software behaviour/characteristics. The PDF graphs contains links to daily live data, that when attempted to reproduce should match, while the graphs in the PDF itself may be outdated. (We now work on better way to secure that as community consensus backed ground truth.))*
- Comments:
 - Restricted data is not a barrier, can share subset or synthetic data for others to inspect
 - Can we convince industry to play by science rules?
- Reproducibility checks pre-publication
- *(from chat: what is needed is a reproducibility-centered research methodology from the very beginning of any research project)*
- “Less badging and more flagging” -- after peer review, CI to check reproducibility and a flag indicates it is still reproducible
 - Problem of overtime decrease in ability to reproduce; See recent paper on [“Active maintenance”](#)
- Availability rather than reproducibility? Introduce concepts in the original process; validity of claims made in the paper; curation efforts to flag dynamic efforts
- eLife Innovation Initiative - re usage of PDF, trying to close the repro loop [“executable research article”](#) ([Stencila](#)) -- allows execution
- Mike Heroux: It is a lot of work to keep something reproducible going forward. It is possible for active projects that have automated testing process, and can be reasonably done. Beyond that , it is hard to expect active maintenance (IMO). Instead, I think we should recognize that making results reproducible helps assure transparency. Transparency is perhaps the most reasonable outcome from reproducibility efforts. Keeping track of the data, software (exact versions), compute logs, etc., enable the ability to go back later and determine how results were produced and, more importantly, why results might have changed.
- A key question: Standards for transparency vs for reproducibility?
 - Reproducibility sharing and drug testing in sports -- parallels? (idea from <https://everythinghertz.com/> that methods for testing will get better in the future)
 - Today, can try and execute a workflow as best we can and share as much as we can, and have a human assess if working OK, if need then

take a closer look later if it turns out that the research is significant / breakthrough

- *(from chat: in a few years... most of today's software will have collapsed! More on software collapse: <https://hal.archives-ouvertes.fr/hal-02117588>)*
- Computational reproducibility as means to an end: verifiability and reusability.
- *(from chat: I think we need minimal standards for transparency (not the same as a result being reproducible) and for standards to be enforced.)*
- See Library of Congress list of endangered file formats:
<https://www.loc.gov/preservation/resources/rfs/TOC.html>
- Confirmation depth: is it computational repro what we're trying to do or is the main question: do we trust the results
<https://davidsoergel.com/posts/confirmation-depth-as-a-measure-of-reproducible-scientific-research>
 - Reproducibility of one paper vs a field -- diff standards of verification?
 - *(from chat: My hope is that reproducibility enables "tests of robustness", which may involve tools swaps, parameter exploration, or data subsets at the time of peer review. This is roughly analogous to "deep confirmation". Runnability is necessary but not sufficient to allow this.)*
 - *(from chat: Replicability is a basic robustness test. And it really requires reproducibility to be effective.)*
- A challenge, e.g., DB community vs. Earth system sciences and remote sensing: in the latter most data is open (gov. satellites), every community needs to have a convo on what's reproducible for them - what it means for that field
 - Another example: qual research
- Project: 40 interviews re publication of repro research outputs
 - What are the objects that need to be published? Not clear today, disciplines are not evenly mature
 - The word data has different meaning in diff disciplines
 - Data curator / subject libs can support repro efforts on discipline -basis; hard to do centrally
 - How is labor for that supported? resourced?
- How to measure repro efforts in a journal - is it working as it should?
 - Journals are not setting very "sharp" reqs on repro, but repro checks are seen to lead to reputation gain, more credible
 - At AJPS no decline in submissions, impact factor went up

- Main Q is what's the goal -- publishing something that is verified and does not include an error
- *(from chat: for metrics, I lately think first of bringing in the CHAOSS folks <https://chaoss.community> who have expertise in software and related metrics - building them, community consensus around them, using them, etc)*
- Need for training on transparency and repro practices
 - *(from chat: Training the researchers is the first step, for now as eLife we are approaching authors that have already published with us to make their articles executable, and we are using that experience to be able to shape future training and feedback from the authors on the amount of workload)*
- Incentives
 - CASCAD (<https://www.cascad.tech/>), a certification agency, independent agency that issues a badge, for a fee
 - Funders have a role in this: provide funding for review services, possibly third parties like CASCAD
 - Packaging the research in a reproducible way is still difficult but it would be a start
 - Will need to be received well by the community
 - A challenge to researchers is that they don't have time to do everything that they are asked to do
 - Paid peer review -- is this once or over time?
 - Ethically should not expect peer reviewers to work for free; recent poll about paying \$450 for an exhaustive review
 - *(from chat: <https://twitter.com/jermdemo/status/1320130056329568259?s=20>)*
 - Based partly on the premise that reproducible peer review is more time intensive and partly on the underlying ethics
 - Many authors are interested in verification, like copy editing, someone to check the work
 - This discussion relates to the larger cultural exchange question in scholarly communication
 - If people work for free for a commercial entity, this is problematic; people are willing to review voluntarily for their own communities, expecting high quality reviews of their next

submission in return, but not to create value for commercial entities

- What are some success stories of journals who are able to do this?
- *(from chat: The way we are getting more positive feedback at eLife is by connecting personally and implementing technologies that are already used by smaller communities that are obviously already interested in reproducibility but need that final push by trusted stakeholders/influencers)*
- Need to be rewarded, potentially as co-authors (though this would require manuscript acceptance)
- Credit is most important
- *(from chat: Pay them whenever they review, whether we do more reviews seems orthogonal?)*
- *(from chat: IMO, unpaid peer review is an artifact of an academic time when everyone had job security; of course it works out well for journals, too)*
- *(from chat: Do we need to pay people? Not sure. If you publish, you should review. If your community expects 3 reviews per paper, you should do 3X reviews vs submissions. It's more a barter system than a financial transaction, at least at this time.)*
- *(from chat: I wonder about equity issues and unpaid work.)*
- *(from chat: Honestly, I still do not understand how paying the reviewers improve or assure reproducibility. I just see a moral discussion, but not a direct connection to reproducibility.)*
- *(from chat: the amount of effort required to conduct a reproducibility check is far greater than a typical 4 hour review)*
- See <https://dpjedi.org/> The Journal Editors Discussion Interface (JEDI) provides the community of social science journal editors a forum; For those supporting journals, check out the Journal Editors Discussion Interface to discuss these and other issues about scholarly publication issues
- *(from chat: Beyond that, there are many researchers that want to do open and reproducible research, but don't know how to do it and aren't rewarded for doing so. I don't think addressing either of those requires metrics and direct incentives so much as making more training available and having it integrated into other curriculum, and valuing openness in evaluation; Many journals and funders value and evaluate "novelty" and that doesn't have any metrics.)*

- Institutions? How can they be a player in this? Legal compliance is a motivator, OSP office is attentive to this sort of thing (retractions, fraud scandals)
- Study idea: Code Execution in Peer Review. <https://doi.org/10.17605/OSF.IO/X32NC>
 - *(from chat: if anyone is interested to create an overview of reproduction policies at journals, your help would be very welcome - see Code Execution in Peer Review. <https://doi.org/10.17605/OSF.IO/X32NC>)*

Questions on the Taking the Pulse blog post on publishing and scholcomm:

1. How can and should we update scholarly publishing to reflect the richness of methods and underlying materials to support claims?
2. Is there a consensus about what should be published in support of reproducibility?
3. How do Open Scholarship principles interact with reproducibility publishing?
 - a. What about FAIR and CARE Principles?
4. Who is responsible for ensuring that published materials are reproducible?
5. How do journal policies on reproducibility manifest in our communities (e.g.TOP Guidelines)? How well are they followed and enforced?
6. Who reviews reproducibility materials before publication? How is that process integrated with the general publication process? How can or should peer review be updated as a process to align with the goals of promoting reproducible research?
7. How can we align our assessment and metrics to foster an environment that promotes reproducible research?