

Revisions for CISE paper

Abstract:

1. Clarify the LIGO project

This has been clarified in the text.

Introduction

2. It might help readers here to clarify explicitly what you mean by ‘unsustainable’.

This has been clarified in the text.

3. Can you make it clearer what the precise link between unsustainable software and the reproducibility crises is?

We clarified this in the introduction.

4. Page2, Lines 29-31. I wonder if here it might be worth mentioning that scientists might often design software for a single purpose or application, with little thought to the longer term or wider usage of that software. Another problem is that software created in academia is not subject to the same level of scrutiny as that created by software communities. For example, it is quite strange that little or no consideration is ever given to software or data shared as part of scholarly publications, and it is rarely checked as part of the peer review process. Indeed, getting recognition for software is a kind of weird process at the moment, as typically it has to be packaged up as part of a traditional journal and manuscript-based process. Which is a bit silly.

We believe this is the message throughout the paper. That software is developed ad hoc without regard for the high standards seen in industry, with a major reason being lack of incentives (which we are addressing in a small way through accelerated review of associated publications). We don’t mention the single purpose aspect given that research software covers a pretty wide range (cross domain infrastructure tools to the kinds of single purpose software that you suggest). We review both types of software.

5. Just a question for the authors that might be worth considering, do you think perceptions about software sustainability are changing? In line with changing government mandates around open science, and increasing concern about reproducibility?

This discussion is well beyond the scope of this paper.

6. Page 2, Line 44: I think these are just collectively known as ‘The Carpentries’ now? Also, the Open Science MOOC (which I run) is working on

training students/researchers more on these things too (see eg here), which might be worth mentioning. I'm obviously biased..

The reviewer is correct in that the umbrella organization is known as The Carpentries - fixed. While we appreciate additional suggestions for training programs to mention, the two distinct types of courses we mention (university and non-university) serve to make the point.

7. Page 2, Line 54: Who is 'us'?

Clarified in the text.

8. Page 3, line 19: Not clear what these 'two major things' are, based on the next paragraphs. Just highlight this a little more to make it clearer.

This has been clarified in the text.

9. For the R packages, do you perhaps have any idea of usage statistics here? That would be really cool to emphasise the transformative impact that rOpenSci has had across many research fields.

This is also beyond the scope of the paper and would require a longer discussion around software metrics. We do have partial download counts from 1 of 120 mirrors that serve R packages. However these counts are not a good indication of use, especially since various services (including our own) are frequently downloading packages for testing and containerization.

10. For the peer review paragraph, I wonder if it might be worth highlighting that what you've kind of done is combine some of the best aspects of open source (e.g., open collaboration) with the best aspects of science (e.g., rigorous and reproducible research) to create a really strong community around rOpenSci. That's sort of the impression I get. And it might be worth noting the results of Prechelt et al. (2017) here on how the community considers the current status of peer review in software engineering (<https://arxiv.org/pdf/1706.07196.pdf>).

We say this in the introduction. Our text is: This approach brings together best practices for publication peer-review along with new practices that are unique to reviewing software. This system deliberately combines elements of traditional academic peer review (external peers), with practices from open-source software review.

Challenges with research software

11. Maybe for each section here, describe briefly what they actually mean? People not familiar with software might not know, for example, what documentation actually is. Would help this paper appeal to a wider audience.

We updated terms that were not clear in the section.

Software review as a service

12. What do you mean by ‘the review process is fully open’ in this regard? There are a lot of potential interpretations for this: <https://f1000research.com/articles/6-588/v2> (this issue again on page 6, line 13).

Thanks for pointing us to this great read. We have amended the text to reflect precise OPR traits covered by our process.

Advantages and limitations

13. Page 6, lines 46-49, I have pretty much no idea what any of these things are. I think given that you are trying to make the case that software sustainability, and its advantages, are useful for the wider research community, being careful to explain any specialist terms like this is important for accessibility.

We updated terms that were not clear in the section.

Incentives

14. Page 7, line 51: Publications, which are sort of like a boxed up package that act to provide a narrative/advert for research, and the actual communication of the process of the research itself is often an afterthought.
15. Is it not a bit weird again that we are still submitting software to the same incentive systems that are based around journal systems? Does the increasing importance of software trigger the need for a greater diversification of research evaluation processes, rather than just a narrow focus on citations? For example, the research community digests and re-uses software much more differently than it does for research papers, and this means that we need to look at different ways of providing credit for these things (see here for some discussion on this matter: <https://academic.oup.com/femsle/advance-article/doi/10.1093/femsle/fny204/5078345>).

Sure this is a bit weird, but for now the academic promotion and tenure systems are strongly tied to publication records. So anything we can do to give academics traditional credit will help them advance in academia until we can solve the larger issue of credit for activities beyond publications.

16. For example, we don't want to just provide incentives for engagement, but for high quality engagements, which is possible through software-based process which are more collaborative, and a sort of 'reviewing the reviewers' system. So it might be possible to do things like create virtual rewards such as points, badges, or abilities that capture the altruistic nature of software development and couple it to the development of researcher profiles. Having incentives based on social processes rather than journal brands seems like a nice future to me.

We updated one of the bullet points under benefits of ropensci peer review to cover this point.

General comments

17. I know this paper focuses on rOpenSci, but I wonder if it might just be worth adding a nod to some of the other initiatives out there too? For example, ReScience (rescience.github.io), created to publish replication efforts in computational science. And the ACM task force on data, software and reproducibility in publication: <https://www.acm.org/data-software-reproducibility> (some other examples given here: <https://f1000research.com/articles/6-1151/v3>).

We would be happy to mention other related efforts. However, ReScience (where the lead author is an editor) does not actually review or publish software. The focus of the journal is on replication, not the software. There is almost no overlap in the areas covered by the two efforts. Looking over examples in the linked paper, we have already listed JOSS which occupies a very similar space and has strong connections with this effort. It is also not clear what the task force aims to do (However members of the task force are also senior personnel on the URSSI project which is an ongoing effort aimed at tackling this issue of software sustainability).