AGILE Conference Reproducibility Review 2020

Daniel Nüst, AGILE 2020 Reproducibility Chair
daniel.nuest@uni-muenster.de
https://bit.ly/agile2020-repro-review-slides

AGILE Association of Geographic Information Laboratories in Europe

REPRODUCIBLE AGILE

My name is Daniel Nüst and I have the honor this year to serve as the AGILE conference's Reproducibility Chair.
In the next 5 minutes, I will report about the current status of the reproductions of accepted AGILE full papers.

But first, let's quickly recap how we got here.

AGILE Reproducible Paper Guidelines

https://doi.org/10.17605/OSF.IO/CB7Z8

Created by AGILE Initiative in 2019, see report at https://osf.io/hupxr/

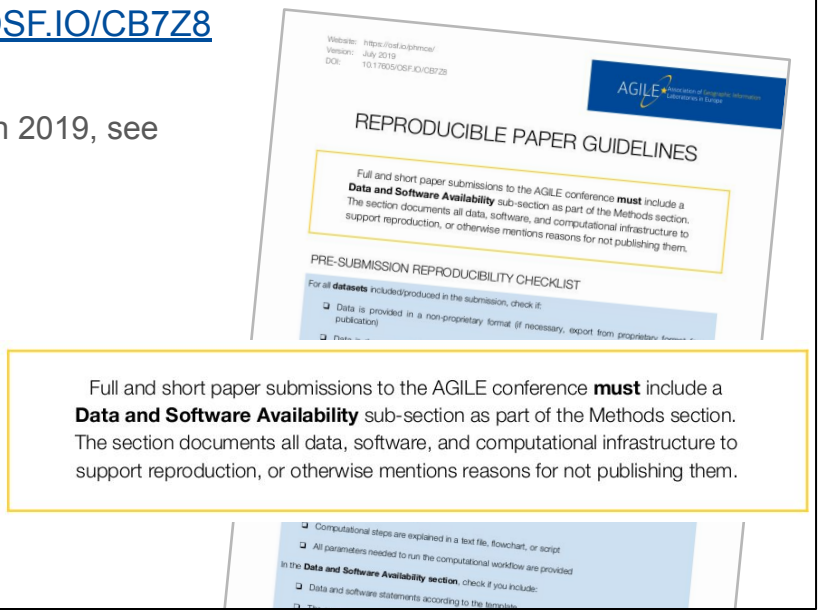
Transparency

Promotion

Acknowledge spectrum

GIScience

https://osf.io/phmce/wiki/home/



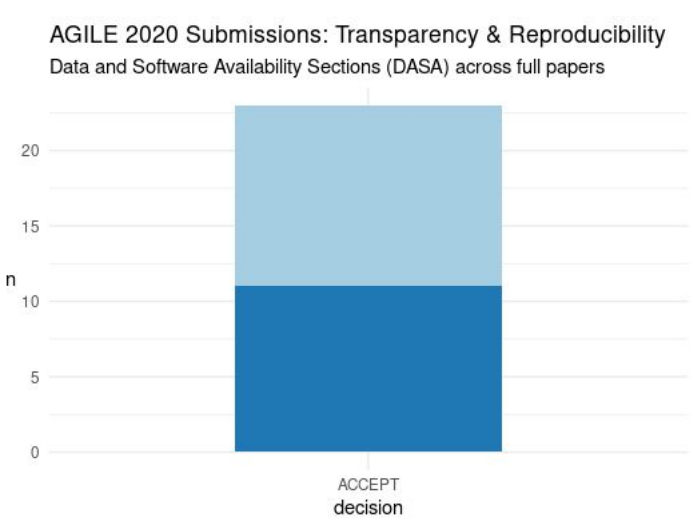
Full and short paper submissions to the AGILE conference **must** include a **Data and Software Availability** sub-section as part of the Methods section. The section documents all data, software, and computational infrastructure to support reproduction, or otherwise mentions reasons for not publishing them.

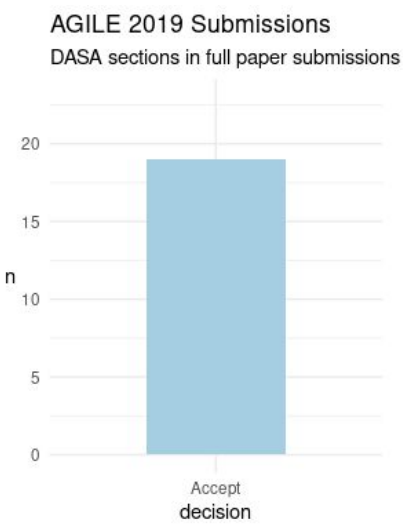
In 2019, AGILE supported an initiative for developing guidelines for reproducible papers.
These guidelines are the first compilation of tipps and tools how to increase transparency in GIScience papers, including a wiki with examples for typical GIScience datasets and workflows.
Their intention is to promote positive examples of reproducible computational research.

Primarily, the guidelines include AUTHOR GUIDELINES, that is how to incorporate DATA and COMPUTATIONAL WORKFLOWS IN RESEARCH PAPERS and a section on WRITING THE DATA AND SOFTWARE AVAILABILITY SECTION, or “DASA” section.

Having a DASA section is the only mandatory requirement! It is absolutely possible not to provide any links to data and code and comply with the guidelines at the same time.

These guidelines were put into the call for papers as a recommendation.





0% **of rejected papers have a DASA section (correlation, not cause?)**
48% **of accepted full papers have DASA section**

Reproducible research and GIScience: an evaluation using AGILE conference papers
https://doi.org/10.7717/peerj.5072

Let's take a look at some numbers:
Out of 23 accepted full papers, 11 have a DASA section, which is about 48%.
None of the rejected papers had a DASA section, but that is likely a correlation and not a causation of course, though one main arguments for higher transparency and reproducibility is indeed that the method can be presented in a more convincing way to reviewers.

I think this is a really great achievement for the whole community and I want to thank all authors of these 11 manuscripts for their open-mindedness and their engagement for open science.

Let's not forget that in 2019, **none [next]** of the papers had a DASA section, and a study from 2018 showed that none of the 32 best paper nominees (full and short papers) from 2010 to 2017 would have provided enough information to even attempt a reproduction.

So how many reproductions did we attempt this year?

Reproducibility Reports

https://doi.org/10.17605/OSF.IO/6K5FH

2 published (both partially successful) **2019: 0**

6 on-going

14 not possible/not attempted (5 of which after communication with authors):

- authors not contacted if there was **no starting point** in the paper (no reference to data or code despite clearly computational workflow)
- **documentation insufficient** for third party (link to "X on GitHub", generic dataset citation, not specific file/subset/scene), including **manual** steps
- **sensitive data** on human subjects or **confidential/commercial data** (EULA prohibits data redistribution but access not scripted)
- **proprietary** software

Up to today, we have published **2 [next]** partially successful reproductions.
"Partially", because not all parts of the articles are reproducible.
Both of these little successes follow suggestions form the guidelines: 1 provides a computing environment with the click of a single link based on Binder, and the other used an anonymous Figshare record for data and code already in the first anonymous submission.

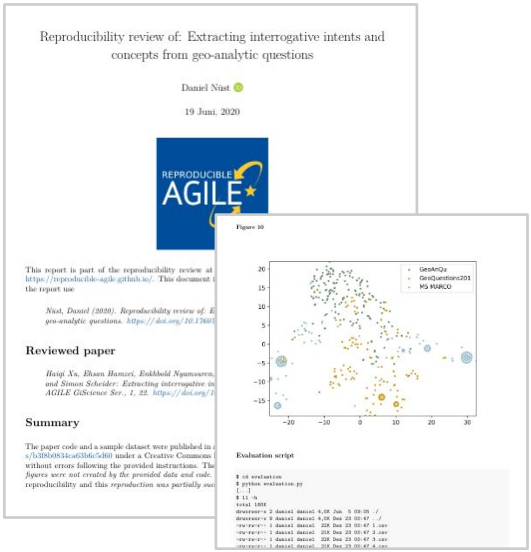
6 reproductions are still on-going.

14 papers were assessed as not possible to reproduce, for different reasons. Here are a few examples of barriers the reproducibility reviewers faced:

- **[next]** If authors provided no starting point, that is no link to data or code at all, we did not attempt a reproduction
- **[next]** Often *documentation* was not detailed enough for a third party to reproduce a workflow
- **[next]** Some papers use sensitive/confidential/ orcommercial data, or
- **[next]** rely proprietary software

I think this is a good first step, because remember: just last year, the number of even partial successful reproductions that we know of is zero.

Two partially successful reproductions published


Reproducibility review of: Extracting interrogative intents and concepts from geo-analytic questions

Daniel Nüst

19 Juni, 2020

REPRODUCIBLE AGILE

Reviewed paper

Summary

Figure 10

GeoAnQu
GeoQuestions201
MS MARCO

Evaluation script


Reproducibility review of: Integrating cellular automata and discrete global grid systems: a case study into wildfire modelling

Daniel Nüst

2020-06-20

REPRODUCIBLE AGILE

Reviewed paper

Summary

wind_0.3_50.png

model-output-example.png

https://doi.org/10.17605/OSF.IO/7XRQG

https://doi.org/10.17605/OSF.IO/ZTC7M

Here are the links to the two reproduction reports published on the Open Science Framework.

Please take a look and let me know if you think these kind of reports are useful for authors, students, readers, or reviewers, and how we can improve them.

General observations and lessons learned

- **Real test run** (cf. original plan)
  > need guidance on *time* spent by repro reviewers + **"easier reproducible" papers**
- **Reproduction not attempted != bad science**, reproducibility is not binary but a ***spectrum***
  > continue education on reproducibility, increase requirements while practices spread in community
- Additional **reproducibility questions for scientific reviewer** not without errors (missed DASA), reviewers commented on reproducibility, but no attempts to reproduce > evolve reviewer roles and tasks
- Repro reviews **were less strict than original ideal** (*partially* reproducible, not only *with* DASA section)
  > promote positive examples and don't expect perfection
- **Cancellation** of conference wrecked **schedule** > goal for the future is still to be done by time of conference, but need better alignment with review and publication milestones
- **Non-blindness** is required (e.g., GitHub links), higher use of anonymisation only works for some analyses
- Decided to skip **short papers** early on > probably better, community/repro reviewer still learning
- Possibility for authors to **object report publication** is good for interim period, but bad for acknowledging the reproducibility reviewer's work > do not continue to ask for permission
- "Must" in guidelines and "should" in CfP was **confusing (CfP trumped GL)**

This is the first reproducibility review at an AGILE conference, and as you can expect from a first time, we deviated quite a lot from the original plan, which is fine: most importantly, we wanted to take home lessons to improve future reviews. For the sake of time, let me focus on a few items on this list.

First, a not attempted reproduction does not mean there is bad science, but we (= the reproducibility reviewers) really experienced the spectrum of reproducible research first hand. We need to continue the education within the community on how to write more reproducible papers. This will allow the community to increase requirements over time.

Second, the cancellation of the physical conference and subsequent changes really messed up our schedule. We need to rethink how we integrate reproducibility reviews into the review and publication process, and how much importance we want to ascribe.

Third, you **cannot** assess the reproducibility of an anonymised manuscript.

Next steps

- Complete remaining **reproducibility reviews**
- Coordinate with Copernicus Publications on **badges**
- **Update AGILE Reproducible Paper Guidelines**
  - Add “Reproducibility Reviewer Guidelines” section
  - More languages.. 🇮🇹 ?
- **Discuss and coordinate** with Council/Org Committee/Community
  - **Mandatory** DASA @ AGILE 2021? 2022? Roadmap needed!
  - Reproducibility review @ **AGILE 2021**?
    - **Opportunity: involve ECR in peer review as reproducibility reviewers**
    - process (change order to parallel/before > both review for decision)
    - no option to object report publication
    - more radical measures?

I’d like finish with a short outlook on the next steps.

Most importantly, we want to finish the remaining reviews.

Next, the “AGILE reproducible” badges will be added to the conference proceedings website.

Later, we will update the guidelines so we have a smoother and even more successful reproducibility review at the conference next year.
At least we hope these will be one.
In my personal opinion, there should be one, and the conference organisers and AGILE council ideally provide a roadmap until when the DASA section is mandatory for all papers.
Also, the reproducibility review is a great opportunity to involve early career researchers in peer review, and it would be great to grow the reproducibility committee in that direction.

If you are wondering what other steps the community can take to evolve peer review at AGILE conferences, I encourage you to take a look at the AGILE initiative’s report - linked here.

(Will we build a share infrastructure for online computations? Should we reject irreproducible papers in 2023?)

Thank you!

What are your questions?

https://discourse.agile-online.org/c/reproducible

https://bit.ly/agile2020-repro-review-slides

Reproducibility Committee 2020 + Initiative
Daniel Nüst (University of Münster, GER)
Frank Ostermann (University of Twente, NEL)
Carlos Granell (Universitat of Jaume I, ESP)
Alexander Kmoch (University of Tartu, EST)
Barbara Hofer (University of Salzburg, AUT)
Rusne Sileryte (TU Delft)

REPRODUCIBLE AGILE

https://reproducible-agile.github.io/

Word-stem cloud of all AGILE 2020 submissions (full/short/poster)

We might not have time for questions and comments today, so I invite you to reach out to me directly, join the AGILE online community forum, find out more about the Reproducible AGILE initiative, or take another look at these slides - you see the links to all these resources here and I posted a link to the slide in the online meetings chat.

Finally, *I am not doing all of this alone*!
I want to thank my fellow members of the reproducibility committee for their hard work:
THANK YOU Frank, Carlos, and Alex for contributing to the reviews this year.

Thank you all for your attention.