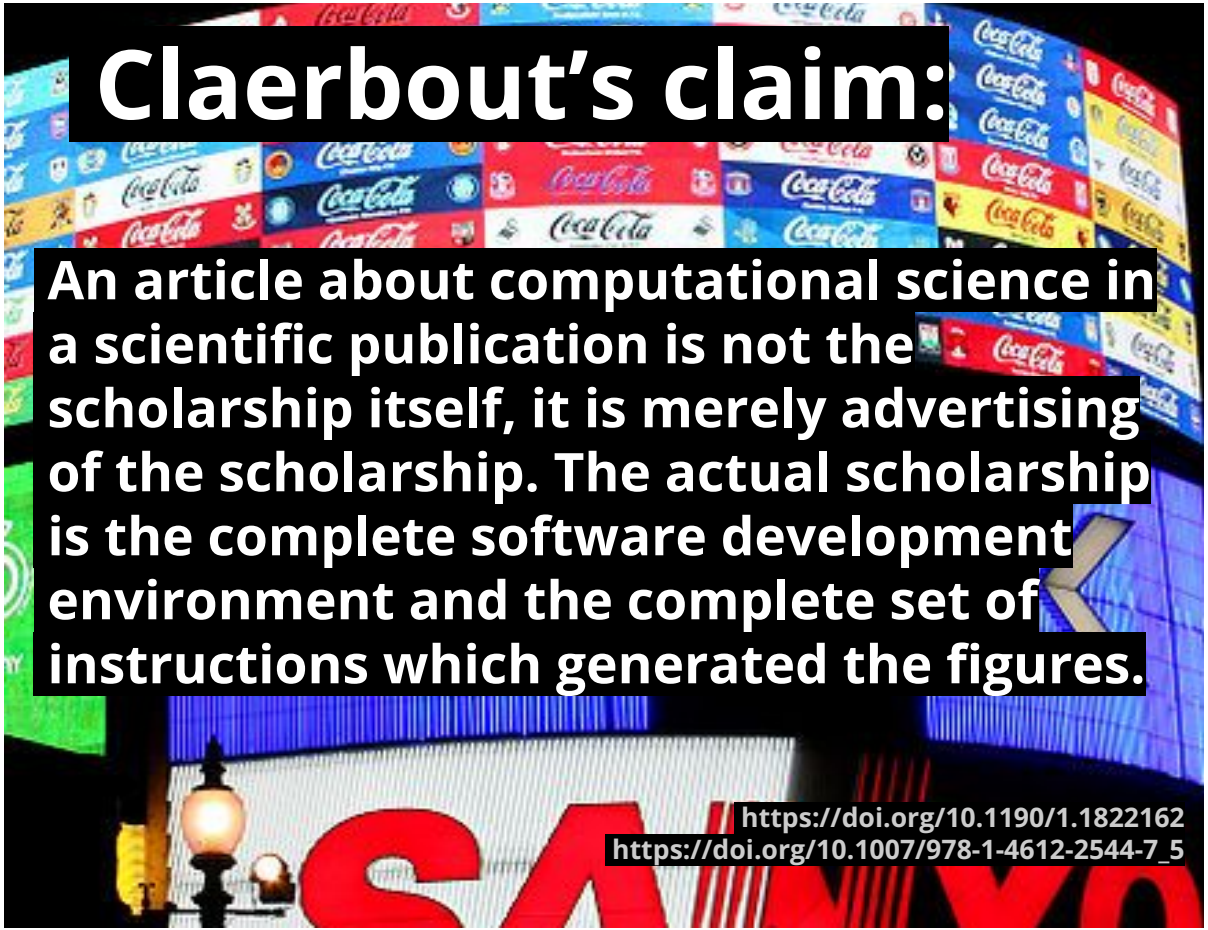
A peer review process for higher reproducibility of publications in GIScience can also work for Earth System Sciences

Daniel Nüst[1]
Frank O. Ostermann[2]
Carlos Granell[3]

[1] Technische Universität Dresden, Chair of Geoinformatics, Geosciences, Dresden, Germany (daniel.nuest@tu-dresden.de)
[2] Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands (f.o.ostermann@utwente.nl)
[3] Institute of New Imaging Technologies, Universitat Jaume I de Castellón, Castellón, Spain (carlos.granell@uji.es)

Please take pictures of the slides!

EGU 2023, Vienna
10.5194/egusphere-egu23-15384
https://bit.ly/peer-review-process-ess



AGILE Association of Geographic Information Laboratories in Europe

HOME | AGILE COMMUNITY | CONFERENCE 2023 | PAST CONFERENCES & PROCEEDINGS | AGILE ACTIVITIES | FUNDING INITIATIVES | FAQ

Members

Search for: e.g. City, Organisation, Person in AGILE memberbase GO

Name	City	Country
52 North	Muenster	Germany
Bell Edwards Geographic Data Institute, School of Geography & Sustainable Development, University of St Andrews	St Andrews, Scotland	United Kingdom
Bochum University of Applied Sciences, Department of Geodesy	Bochum	Germany
Carinthian Technical Institute	Villach	Austria
Catholic University of Leuven	Leuven-Heverlee	Belgium
Centro ALGORITMI, Escola de Engenharia, Universidade do Minho	Guimarães	Portugal

Who is AGILE?
https://agile-online.org/

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

2017, ‘18 & ‘19: Workshops on reproducibility
2019: Reproducible publications at AGILE conferences (AGILE Initiative supported by AGILE Council)
2020: First AGILE reproducibility review
2021: Second AGILE reproducibility review
2022: Third AGILE reproducibility review - guidelines mandatory
2023: Still going!

AGILE Reproducible Paper Guidelines

(v1)

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

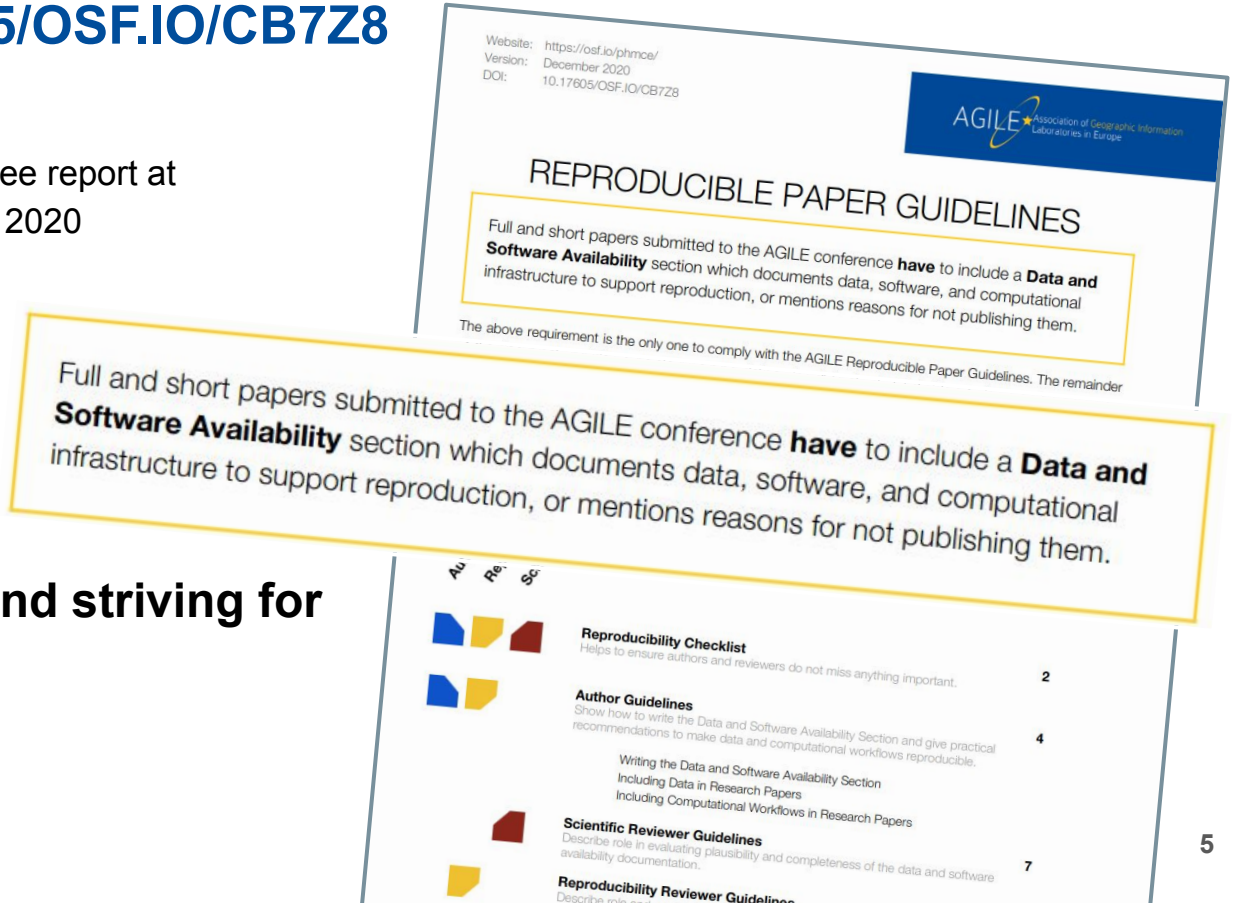
Created by AGILE Initiative in 2019 (see report at **https://osf.io/hupxr/**) and updated in 2020

Transparency over enforcement

Promote, don't exclude

Acknowledge spectrum and striving for ideal

Tailored to GIScience



The guidelines

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

Reproducibility checklist

Author guidelines

Writing DASA section
Data in Research Papers
Computational workflows in Research Papers

Reviewer guidelines

Reproducibility reviewer guidelines

Background

Website: https://osf.io/phmce/
Version: December 2020
DOI: 10.17605/OSF.IO/CB7Z8

AGILE Association of Geographic Information Laboratories in Europe

REPRODUCIBLE PAPER GUIDELINES

Full and short papers submitted to the AGILE conference **have** to include a **Data and Software Availability** section which documents data, software, and computational infrastructure to support reproduction, or mentions reasons for not publishing them.

The above requirement is the only one to comply with the AGILE Reproducible Paper Guidelines. The remainder of the document provides concrete recommendations for all involved stakeholders to increase transparency, reproducibility, and openness of computational GIScience research. The following table of contents shows the recommended parts for different readers. Familiarity with all sections is, of course, beneficial.

Further resources

These guidelines can not cover all details of the reproducibility review at AGILE conferences. For more information for authors, translations, and practical examples see the guidelines wiki. For more information about the review process and deadlines, see the process description. For any questions, please visit the AGILE Discourse server's forum for the Reproducible Paper Guidelines.

Thanks!

To the contributors to the AGILE Reproducible Paper Guidelines

Daniel Nüst

Frank Ostermann

Rusne Sileryte

Barbara Hofer

Carlos Granell

Marta Teperek

Anita Graser

Karl Broman

Kristina Hettne

Connie Clare

Frederique Belliard

Yan Wang

https://giphy.com/gifs/HBOMax-Pnh0Lou03fv92J4puZ

Review process

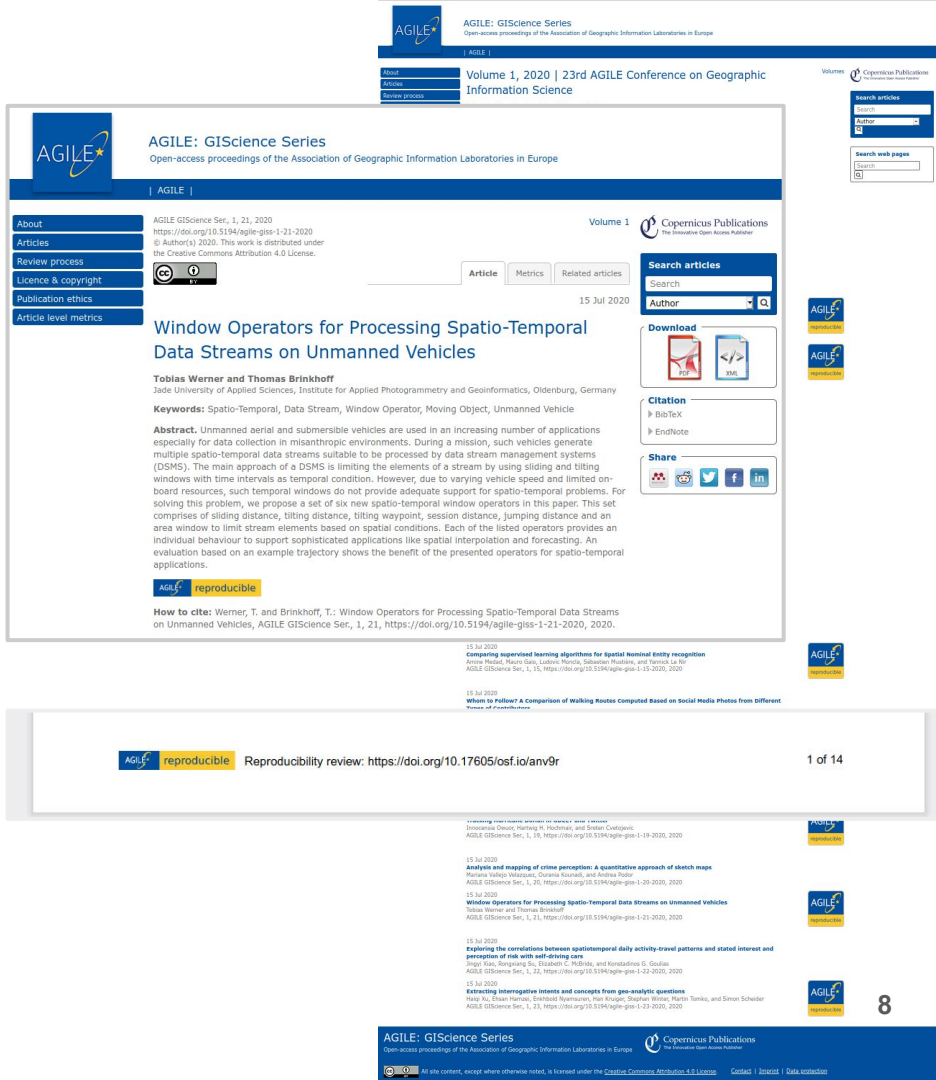
Proceedings: **agile-giscince-series.net/review_process.html**
Process documentation: **osf.io/7rjpe/**

Reproducibility review of **full papers** *after* accept/reject decisions by scientific reviewers

Reproducibility review & communication

Community conference & volunteers

Badges on proceedings website and article title page link to **reproducibility reports (Thanks Copernicus!)**



Example reproducibility reviews from AGILE 2022

Reproducibility review of: Understanding the Imperfection of 3D point Cloud and Semantic Segmentation algorithms for 3D Models of Indoor Environment

Rémy Decoupes

2022-06-10

This report is part of the reproducibility review at the AGILE conference. For more information see https://reproducible-agile.github.io/. This document is published on OSF at https://doi.org/10.17605/OSF.IO/Z7P8K. To cite the report use

Decoupes, R. (2022, May 25). Reproducibility review of: Understanding the Imperfection of 3D point Cloud and Semantic Segmentation algorithms for 3D Models of Indoor Environment. https://doi.org/10.17605/OSF.IO/Z7P8K

Reviewed paper

Cai, G., and Pan Y.: Understanding the Imperfection of 3D point Cloud and Semantic Segmentation algorithms for 3D Models of Indoor Environment, AGILE GIScience Ser., 3, 2, https://doi.org/10.5194/agile-giss-3-2-2022

Summary

The code and data provided by the authors allow to **partially reproduce** the computational work presented in the Section 4.4, *Interactive Exploration of Data Imperfection for Model Tuning*, of the paper. The model training (PointNet++) and the input data for Section 4.4 are not reproducible by the provided code but the authors added a note, in the GitHub repository of their project, explaining how the data was generated. Three figures in Section 4.4 are fully reproducible (5, 6, and 7), 2 partially (4 and 8) and 4 are not (3, 9, 10, and 11).

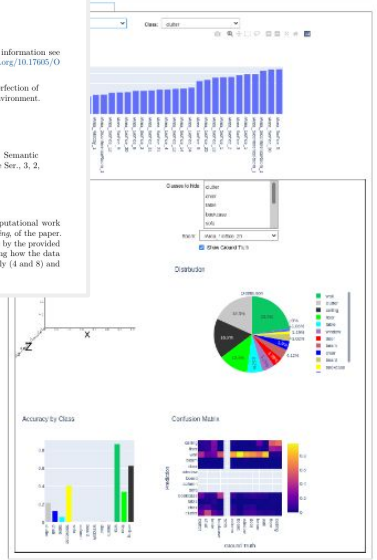


Figure 2: visual-exploration-dashboard-pt2.ipynb: Tab2. Partially reproduction of Figure 4.

https://doi.org/10.17605/OSF.IO/Z7P8K

Reproducibility review of: Unlocking social network analysis methods for studying human mobility

Jakub Krukar

2022-06-09

This report is part of the reproducibility review at the A
https://reproducible-agile.github.io/. This document is pu
the report use

Krukar, J. (2022, May 23). Reproducibility review of: Un
studying human mobility. https://doi.org/10.17605/OSF.I

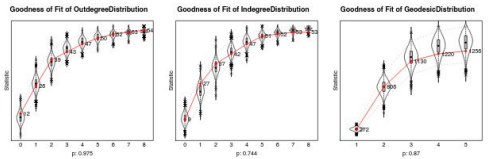
Reviewed paper

Wiedemann, N., Martin, H., and Raubal, M.: Unlocking
human mobility, AGILE GIScience Ser., 3, 19, https://doi.

Summary

The paper provides a link to a GitHub repository that w
improved by the authors after an email exchange. The rep
presented in the paper but most results based on this dat
minor disparities due to automated scaling of graphs. In s
duced. The repository is well-documented, it includes the
and the authors' response to questions and bugs has been

files seem identical with those displayed in the figure. Figure 4 has been successfully reproduced.

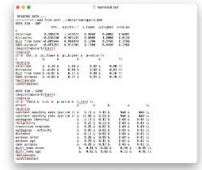


Other figures

In addition, the 'results_foursquare' folder contains two files that are not clearly described but do not directly correspond to any figure shown in the manuscript: 'hist_jac1.pdf' and 'hist_jac2.pdf'.

Table 1

The 'results_foursquare' directory contains a file 'terminal.txt' with some numerical results. The first set of results correspond to the numerical results presented in Table 1, although are not described as such. Numerical results of Table 1 have been successfully reproduced.



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

CODE CHECK ✓

https://codecheck.org.uk/

Independent execution of computations underlying research articles.

Follow us on YouTube: https://www.youtube.com/@cdchck

Principles

1. Codecheckers **record** but don't investigate or fix.
2. **Communication** between humans is key.
3. **Credit** is given to codecheckers.
4. Workflows must be **auditable**.
5. **Open** by default and **transitional** by disposition.

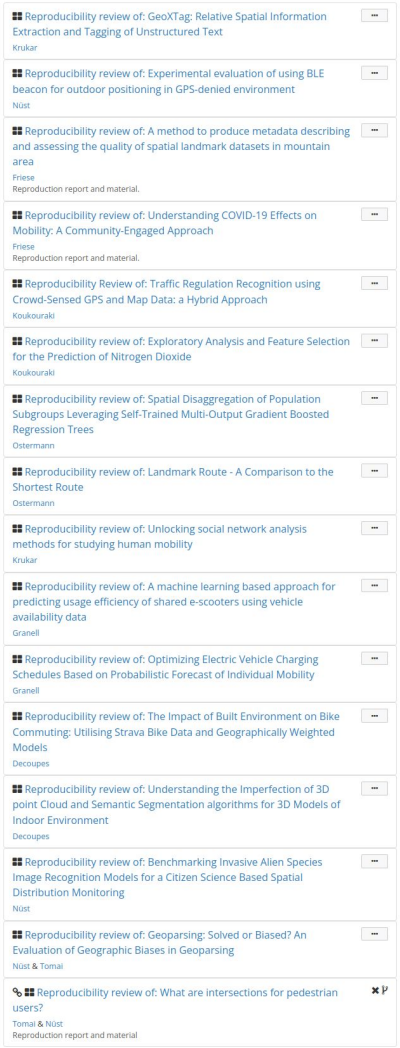
Reproducibility review results

16 reproducibility reports published
= 73% of accepted full papers

2021: 9 2020: 6
2018-19:? ..-2017: 0?! (see 10.7717/peerj.5072)

6 not reproducible:

- authors say too difficult / too busy
- no data nor code (tutorial, conceptual)
- big data + prop. tool & code not working



AGILE reproducible

https://osf.io/r5w79/

Thanks!

To all reproducibility reviewers

Alexander Kmoch
Anita Graser
Arun Sharma
Carlos Granell
Daniel Nüst
Eftychia Koukouraki
Eleni Tomai
Frank Ostermann
Jakub Krukar
Philipp Friese
Rémy Decoupes
Nina Wiedemann
Mehtab Alam Syed

https://codecheck.org.uk/register/

https://giphy.com/gifs/HBOMax-Pnh0Lou03fv92J4puZ

From AGILE to ESS

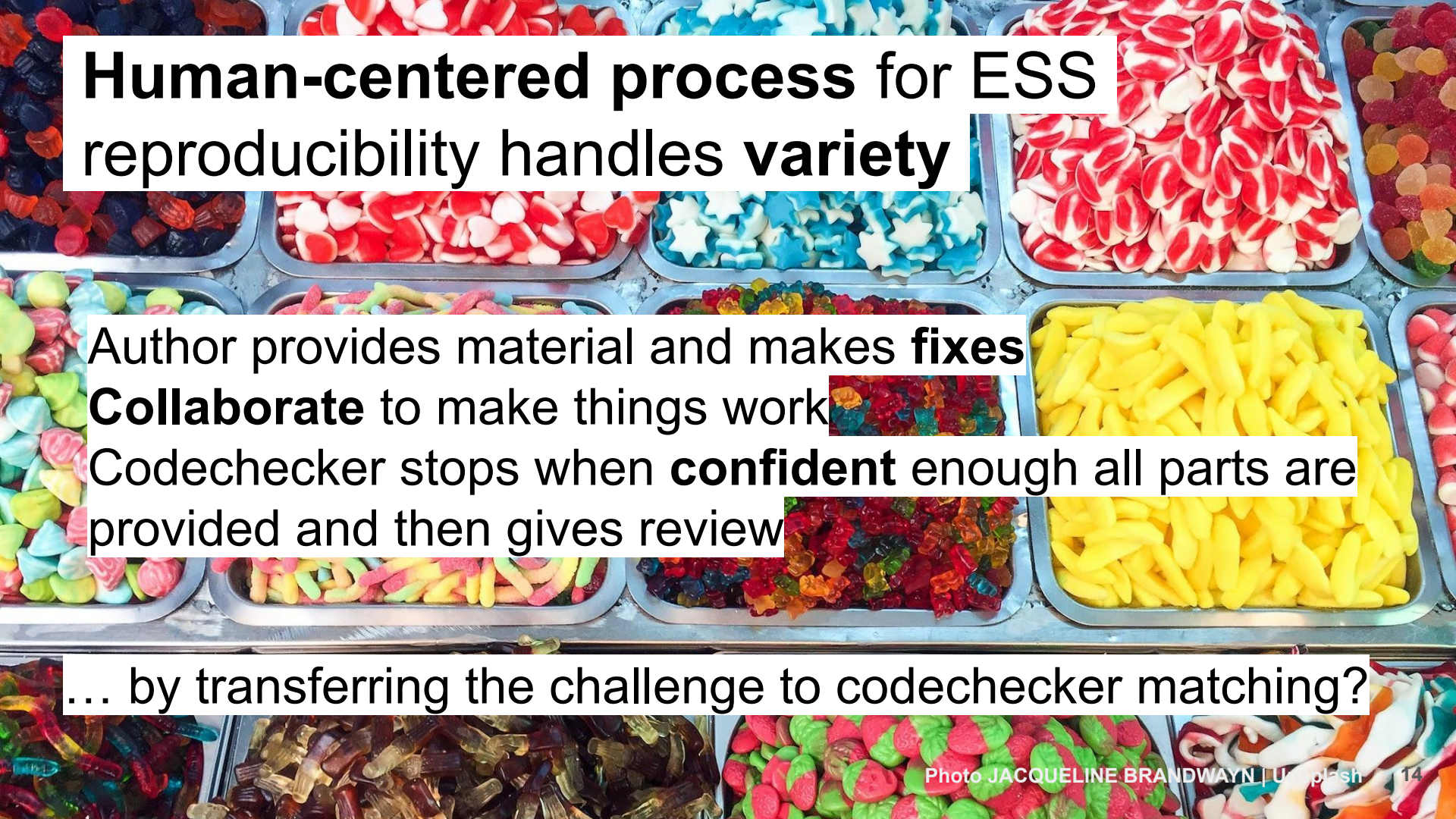
Four ideas that are **transferable** to Earth System Sciences: human-centered process, communicative focus, supportive framework of tools and infrastructure, and candidates.

https://giphy.com/gifs/acorn-tv-wine-cooking-recipe-BhsSZR0k6EMxjCQDwP

Human-centered process for ESS reproducibility handles **variety**

Author provides material and makes **fixes**
Collaborate to make things work
Codechecker stops when **confident** enough all parts are provided and then gives review

… by transferring the challenge to codechecker matching?



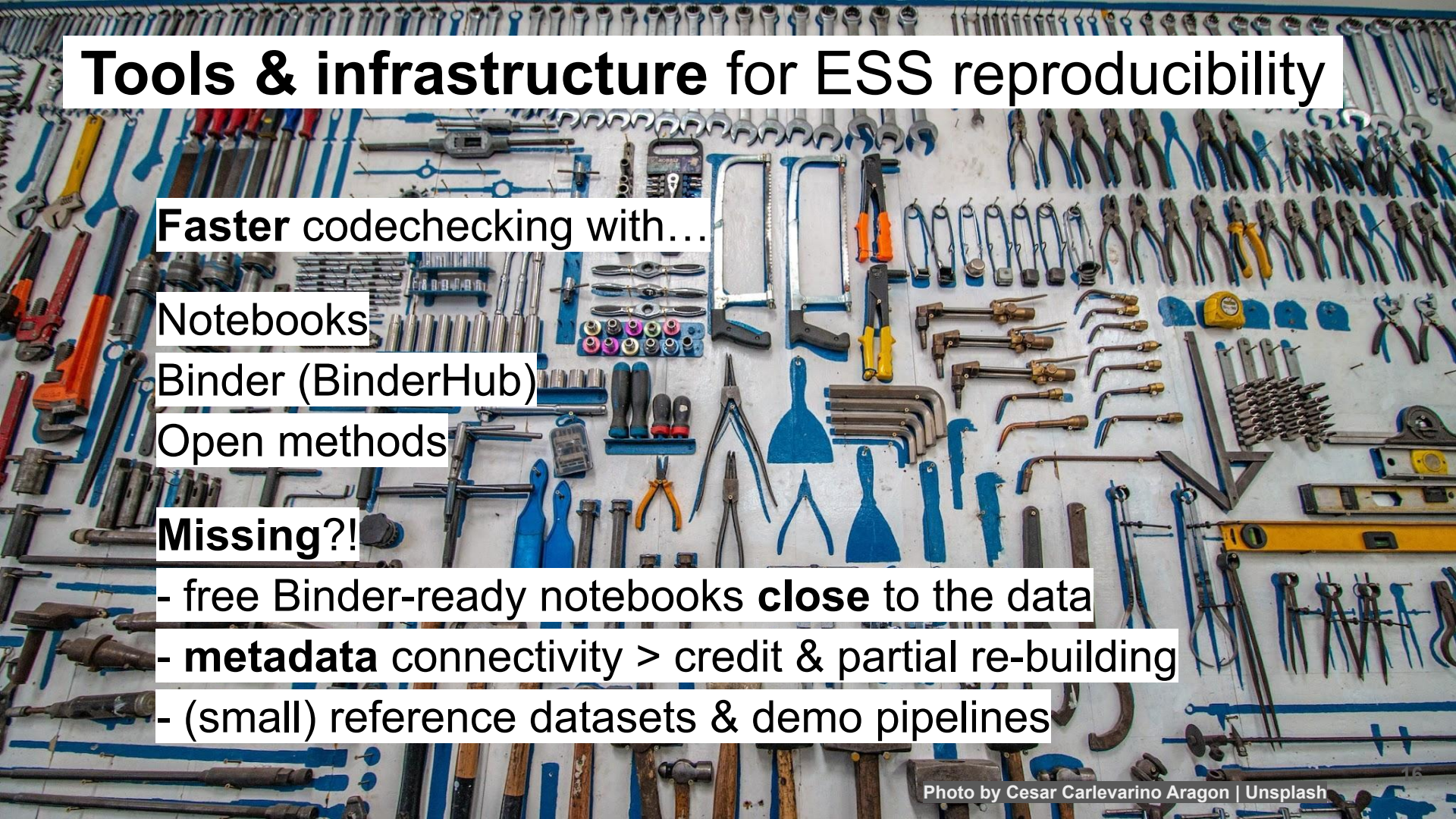
Communicative focus for ESS reproducibility

Avoids rules & automation playing **catch** with innovation & technology
Avoids unification or **limitation** of researcher freedom

Continuous development of shared **practice** over time
& **definition** of “reproducibility” and “how reproducible is enough"

Targets **attitudes** towards sharing, transparency, and openness
Positive learning **experience** for all involved roles

Tools & infrastructure for ESS reproducibility



Faster codechecking with…

Notebooks

Binder (BinderHub)

Open methods

Missing?!

- free Binder-ready notebooks **close** to the data
- **metadata** connectivity > credit & partial re-building
- (small) reference datasets & demo pipelines

Photo by Cesar Carlevarino Aragon | Unsplash

Candidates for ESS reproducibility

Earth System Science Data (**ESSD**)

All EGU/**Copernicus.org**/AGU journals?

ES³

???

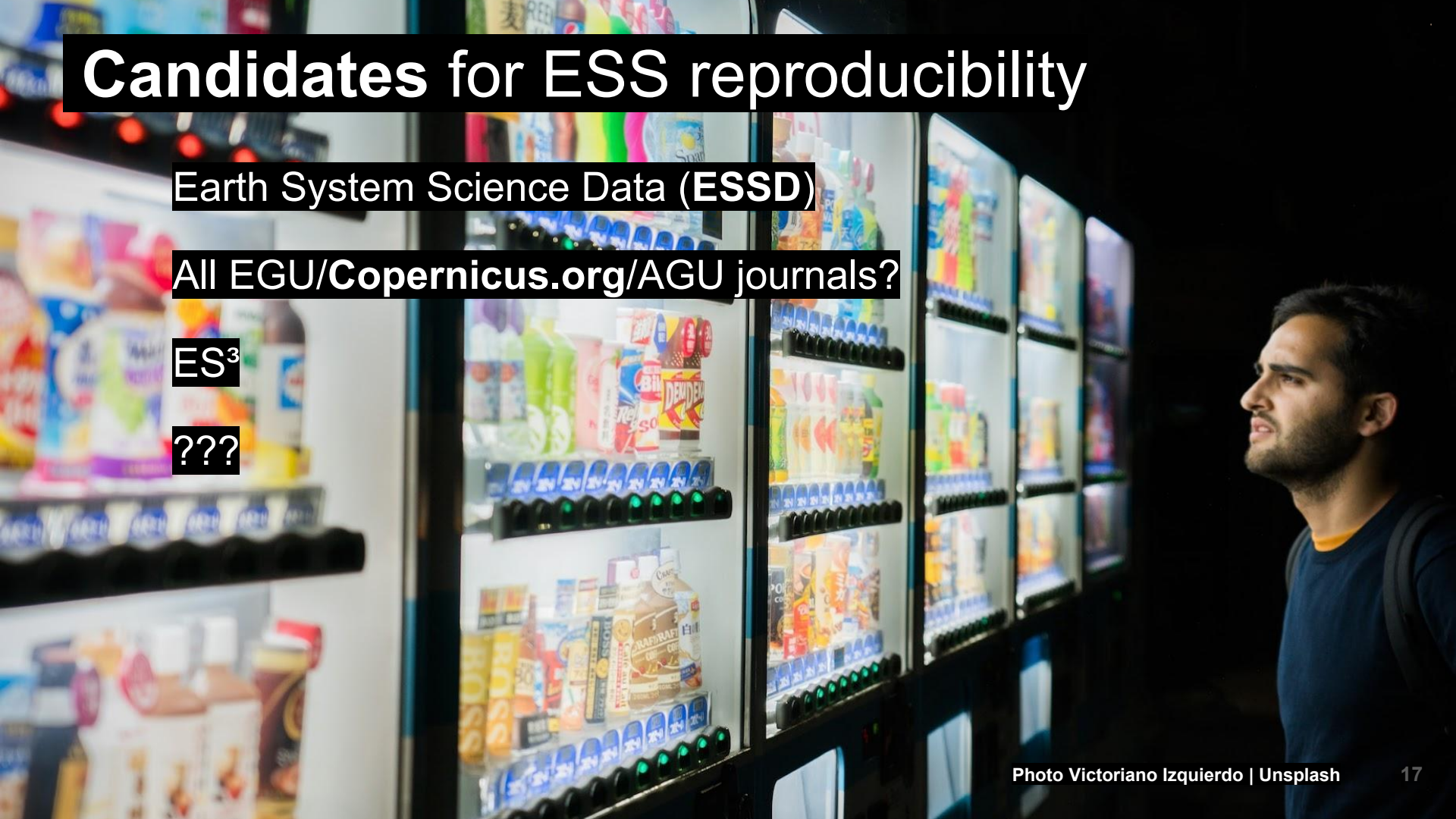


Photo Victoriano Izquierdo | Unsplash

https://media.giphy.com/media/mbhseRYedIG5W/giphy.gif

What are the steps that you could take? Five year plan:

1. Initiate **discourse** with a core team in a (small, sub) **community** (e.g., small event, journal)
2. Document **state** of reproducibility & find **leadership** support
3. Establish guidelines (do-ocracy) & **apply** to current work (preprints?)
4. **Continue** development and **positive** discourse
5. Give talk about experiences and **share**

Thank you!

Questions?
@nordhomen | daniel.nuest@tu-dresden.de

Nüst, D., Ostermann, F. O., and Granell, C.: **A peer review process for higher reproducibility of publications in GIScience can also work for Earth System Sciences**, EGU General Assembly 2023, Vienna, Austria, 24–28 Apr 2023, EGU23-15384, https://doi.org/10.5194/egusphere-egu23-15384, 2023.

Abstract

The Reproducible AGILE initiative (https://reproducible-agile.github.io/) successfully established a code execution procedure following the CODECHECK principles (https://doi.org/10.12688/f1000research.51738.2) at the AGILE conference series (https://agile-online.org/conference). The AGILE conference is a medium-sized community-led conference in the domains of Geographic Information Science (GIScience), geoinformatics, and related fields. The conference is organised under the umbrella of the Association of Geographic Information Laboratories in Europe (AGILE).

Starting with a series of workshops on reproducibility from 2017 to 2019, a group of Open Science enthusiasts with the support of the AGILE Council (https://agile-online.org/agile-actions/current-initiatives/reproducible-publications-at-agile-conferences) was able to introduce guidelines for sharing reproducible workflows (https://doi.org/10.17605/OSF.IO/CB7Z8) and establish a reproducibility committee that conducts code executions for all accepted full papers.

In this presentation, we provide details of the taken steps and the encountered obstacles towards the current state. We revisit the process and abstract a series of actions that similar events or even journals may take to introduce a shift towards higher reproducibility of research publications in a specific community of practice.

We discuss the taken approach in the light of the challenges for reproducibility in Earth System Sciences (ESS) around four main ideas.
First, Reproducible AGILE's human-centered process is able to handle the increasingly complex, large and varying data-based workflows in ESS because of the clear guidance on responsibilities (What should the author provide? How far does the reproducibility reviewer need to go?).
Second, the communicative focus of the process is very well suited to, over time, help to establish a shared practice based on current technical developments, such as FAIR Digital Objects, and to reform attitudes towards openness, transparency and sharing. A code execution following the CODECHECK principles is a learning experience that may sustainably change researcher behaviours and practice. At the same time, Reproducible AGILE's approach avoids playing catch-up with technology and does not limit researcher freedom or includes a need to unitise researcher workflows beyond providing instructions suitable for a human evaluator, similar to academic peer review.
Third, while being agnostic of technology and infrastructures, a supportive framework of tools and infrastructure can of course increase the efficiency of conducting a code execution. We outline how existing infrastructures may serve this need and what is still missing.
Fourth, we list potential candidates of event series or journals that could introduce a code checking procedure because of their organisational setup or steps towards more open scholarship that were already taken.

How to cite: Nüst, D., Ostermann, F. O., and Granell, C.: A peer review process for higher reproducibility of publications in GIScience can also work for Earth System Sciences, EGU General Assembly 2023, Vienna, Austria, 24–28 Apr 2023, EGU23-15384, https://doi.org/10.5194/egusphere-egu23-15384, 2023.

Next steps

Do it again in 2023 🎉

… and for real?
Reject irreproducible papers?

Fix reproducibility review vs. schedule

🛠️ **Revise guidelines?** 🇮🇹 🇫🇷 🇨🇳

Grow reproducibility reviewer team (= YOU!)

Longitudinal meta-research study

Towards opening scholarship

- Format-free first submission
- Review/Publish computational notebooks
- Require CRediT
- Reviewer activity @ ORCID
- Open review (if tenured?)
- Discourse on peer review (read this)

Example reproducibility review reports from AGILE 2022

Reproducibility review of: Exploratory Analysis and Feature Selection for the Prediction of Nitrogen Dioxide

Eftychia Koukouraki

2022-06-15

This report is part of the reproducibility review at the AGILE https://reproducible-agile.github.io/. This document is publishe cite the report use

Koukouraki, E. (2022, June). Reproducibility review of: Explora the Prediction of Nitrogen Dioxide. https://doi.org/10.17605/O

Reviewed paper

Iskandaryan, D., Di Sabatino, S., Ramos, F., and Trilles, S.: Exp for the Prediction of Nitrogen Dioxide, AGILE GIScience Ser., 3 6-2022

Summary

The paper evaluates the competence of selected features in t Machine Learning. For this reproducibility review, the Figures an *Results* were considered, while the Figures of *Section 3 - Explora* corresponding analysis was provided as a Github repository an code were provided through a Zenodo repository. The reproduc ones reported in the paper, so the **reproduction** of the paper i

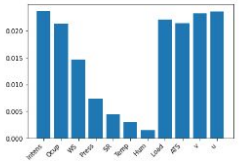


Figure 2: Corresponds to Figure 14 of the reproduced paper

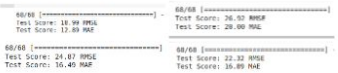


Figure 3: Corresponds to Table 2 of the reproduced paper

```
68/68 [==============================]
Test Score: 6.81 RMSE
Test Score: 5.97 MAE

68/68 [==============================]
Test Score: 5.61 RMSE
Test Score: 5.18 MAE

68/68 [==============================]
Test Score: 3.55 RMSE
Test Score: 3.07 MAE

68/68 [==============================]
Test Score: 4.90 RMSE
Test Score: 4.37 MAE

68/68 [==============================]
Test Score: 3.44 RMSE
Test Score: 2.87 MAE

68/68 [==============================]
Test Score: 19.91 RMSE
Test Score: 15.51 MAE
```

Figure 4: Corresponds to Table 3 of the reproduced paper

Reproducibility review of: Geoparsing: Solved or Biased? An Evaluation of Geographic Biases in Geoparsing

Daniel Nüst, Eleni Tomai

2022-06-10

This report is part of the reproducibility review at the AG https://reproducible-agile.github.io/. This document is publ cite the report use

Nüst, Daniel, and Eleni Tomai. 2022. Reproducibili Biased? An Evaluation of Geographic Biases in Geopar OSF.IO/3DSMV

Reviewed paper

Liu, Z., Janowicz, K., Cai, L., Zhu, R., Mai, G., Biased? An Evaluation of Geographic Biases in Geo https://doi.org/10.5194/agile-giss-3-9-2022

Summary

The article presents an evaluation of geoparsing performanc methods from various sources. Though preprocessing steps a software could not be evaluated, one of two toponym resolu The provided notebooks for exploratory analysis, calculat evaluation could be run and the outputs match the data an this reproducibility report can confirm a partially successf which authors provide reasonable but improvable documen their computational workflow.

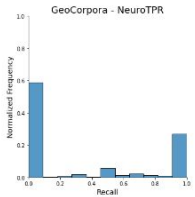


Figure 2: Figure 1 (b)

```
jupyter notebook standard-deviation-toponym-resolution-ambiguity.ipynb
```

After fixing paths, the calculated statistics match the values in **Table 1**:

```
# Edinburgh
Washington County      median_error_distance
std                         807.704385
Clinton          median_error_distance
std                              NaN
Springfield      median_error_distance
std                              NaN
Greenville       median_error_distance
std                         407.517261
Kingston         median_error_distance
std                         680.481725
Georgetown       median_error_distance
std                              NaN
Hamilton         median_error_distance
std                        4105.574962
Jefferson County       median_error_distance
std                              NaN
Newport          median_error_distance
std                        5705.194849
Franklin County        median_error_distance
std                              NaN

# CamCorder
Washington County      median_error_distance
std                         692.177184
Clinton          median_error_distance
std                         539.43373
Springfield      median_error_distance
std                        2770.634287
Greenville       median_error_distance
std                         514.681684
Kingston         median_error_distance
std                         226.447057
Georgetown       median_error_distance
std                        2742.785136
Hamilton         median_error_distance
std                         622.237286
Jefferson County       median_error_distance
std                         384.380587
Newport          median_error_distance
std                         145.050226
Franklin County        median_error_distance
std                         414.985655
```

Spatial Autocorrelation Analysis

This part of the workflow requires the proprietary software ArcGIS pro, for which I do not have access.

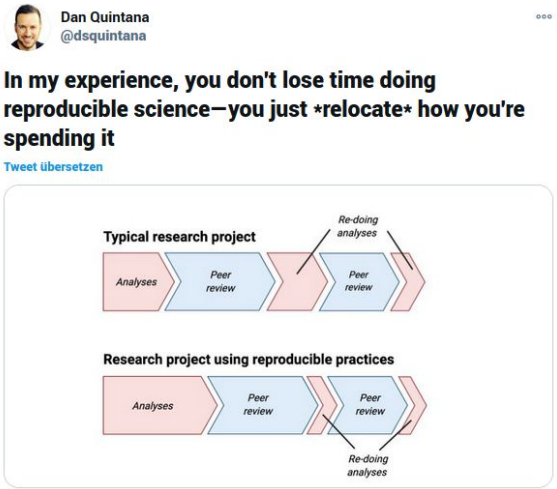
https://doi.org/10.17605/OSF.IO/W7VPH

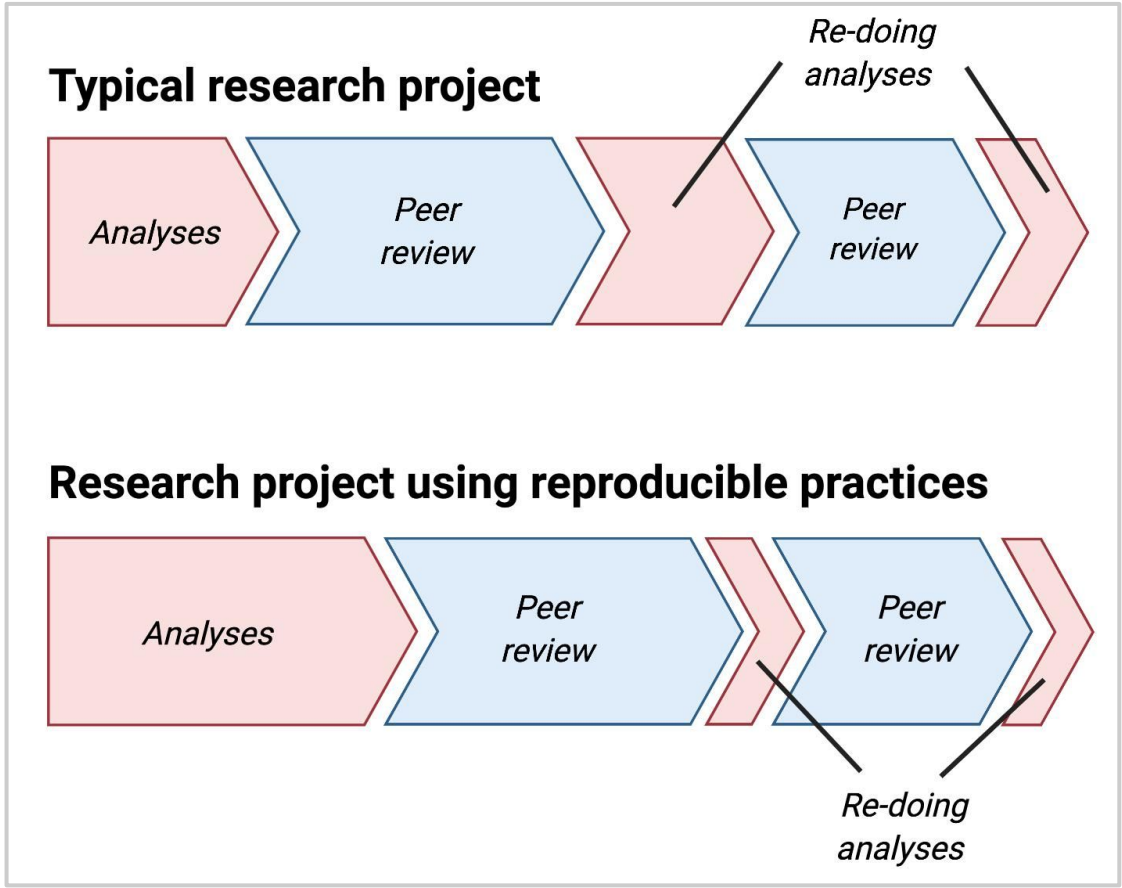
https://doi.org/10.17605/OSF.IO/3DSMV

Collaboration with

2022-002	Z7P8K	conference (AGILEGIS)	41	https://doi.org/10.17605/osf.io/z7p8k	2022-07-09
2022-003	JDTN3	conference (AGILEGIS)	41	https://doi.org/10.17605/OSF.IO/JDTN3	2022-07-09
2022-004	XPG6Y	conference (AGILEGIS)	41	https://doi.org/10.17605/osf.io/XPG6Y	2021-07-09
2022-005	CDFAH	conference (AGILEGIS)	41	https://doi.org/10.17605/osf.io/cdfah	2022-07-09
2022-006	W7VPH	conference (AGILEGIS)	41	https://doi.org/10.17605/osf.io/W7VPH	2022-07-09
2022-007	3DSMV	conference (AGILEGIS)	41	https://doi.org/10.17605/OSF.IO/3DSMV	2022-07-09
2022-008	K78EB	conference (AGILEGIS)	41	https://doi.org/10.17605/OSF.IO/K78EB	2022-07-09
2022-009	94VNX	conference (AGILEGIS)	41	https://doi.org/10.17605/osf.io/94vnx	2021-07-09
2022-010	8B7MR	conference (AGILEGIS)	41	https://doi.org/10.17605/osf.io/8b7mr	2022-07-09
2022-011	KF8SR	conference (AGILEGIS)	41	https://doi.org/10.17605/osf.io/kf8sr	2022-07-09
2022-012	R6PSQ	conference (AGILEGIS)	41	https://doi.org/10.17605/osf.io/r6psq	2022-07-09
2022-013	3G9S8	conference (AGILEGIS)	41	https://doi.org/10.17605/osf.io/3g9s8	2022-07-09
2022-014	6S2GP	conference (AGILEGIS)	41	https://doi.org/10.17605/osf.io/6s2gp	2022-07-09
2022-015	MVQCW	conference (AGILEGIS)	41	https://doi.org/10.17605/osf.io/mvqcw	2021-07-09
2022-016	DJFC2	conference (AGILEGIS)	41	https://doi.org/10.17605/OSF.IO/DJFC2	2021-07-09
2022-017	WNCSM	conference (AGILEGIS)	41	https://doi.org/10.17605/osf.io/wncsm	2022-07-09

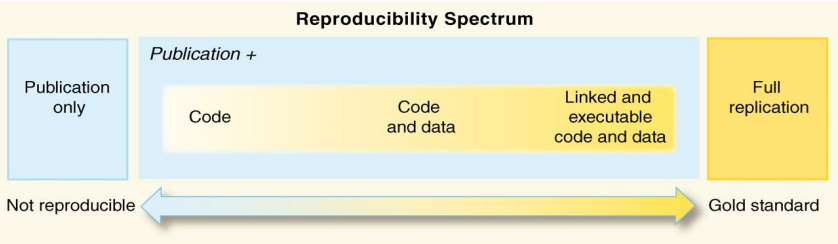
https://codecheck.org.uk/register/



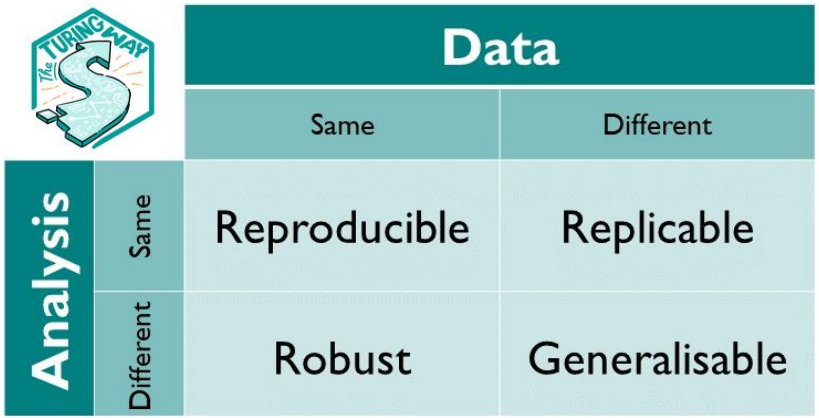


Quintana, D. S. (2020, November 28). Five things about open and reproducible science that every early career researcher should know. https://doi.org/10.17605/OSF.IO/DZTVQ

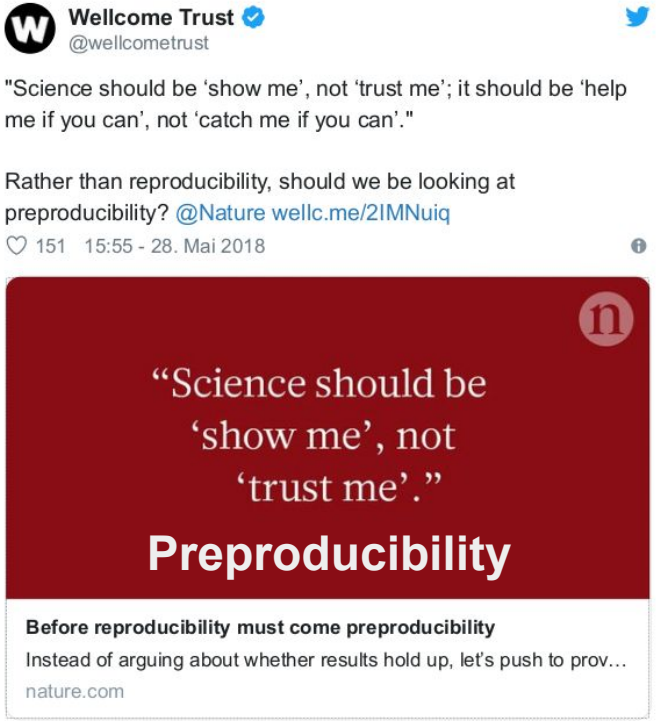
Reproducible Research & Open Science



https://doi.org/10.1126/science.1213847



https://the-turing-way.netlify.app/reproducible-research/overview/overview-definitions.html



https://doi.org/10.1038/d41586-018-05256

Checklist and writing the DASA section

REPRODUCIBILITY CHECKLIST

For all **datasets** included/produced in the paper, check if data:

- ❑ Is provided in a non-proprietary format
- ❑ Is documented for third parties to reuse
- ❑ Is accessible in a public repository and has an open data licence

For all **software tools/libraries/packages** and **computational workflows** included/produced, check if:

- ❑ Reproduction steps are explained in a README (plain text file), flowchart, or script
- ❑ Computational environments (including hardware) are documented or provided
- ❑ Versions of relevant software components (libraries, packages) are provided
- ❑ All parameters and expected execution times for the computational workflow are provided
- ❑ Software developed by the authors is available in a public repository and has an open licence
- ❑ There is a clear connection between **tables, figures, maps, and statistical values** and the data and code that they are based on, e.g., using file names or documentation in the README

In the **Data and Software Availability section**, check if you include:

- ❑ Data and software statements (see examples below)
- ❑ The reasons, if any, for not being able to share (parts of) data or code

For all **data and software** check that:

- ❑ All datasets and code (used or mentioned) are assigned DOIs
- ❑ Datasets and code are cited throughout the paper

After acceptance in the **camera-ready paper** check that:

- ❑ If data has been shared privately or anonymously for peer review, they are updated with all metadata and accessible via a DOI and referenced from the paper
- ❑ If a reproducibility review report will be published for your paper, a DOI URL in the Data and Software Availability section is included using the following template:
  A reproducibility report for this paper is available confirming that [considerable parts of the computational workflow / all results / Figures 1 and 4] could be independently reproduced, see https://doi.org/link_to_report.

WRITING THE DATA AND SOFTWARE AVAILABILITY SECTION

The DASA section provides references to where data, software and documentation is available (e.g., paper section or README file) and under what conditions (e.g., copyright, licenses or access procedures for protected data). It should be concise and contain persistent links to repositories using Digital Object Identifiers[7] (DOI). You may remove links for anonymity during peer review ("xxx"), or share anonymized links[8] if your repository supports them. Data, software and (third-party) tools should be cited following recommended citation or standard citation guidelines. Possible statements for the DASA section are provided below. You may include one of these statements or draft your own.

Statements for non-computational or conceptual work

No data or code was collected, developed, or used in this work.

The full list of reviewed literature is available at [link to attachment or citable deposit of bibliography].

The full concept maps are available at [link] *and the ideas were first sketched in a blog post at* [link].

Research data/code supporting this publication ...

… is available in [name of the repository(-ies)] *and is accessible via the following DOI* [DOI link(s)]

… was accessed on [date of dataset access/download] *with the following* [query parameters, if applicable] *under the license* [dataset license].

… was downloaded manually using the services at [name of organisation] *(using a departmental subscription for costs) and* [name of organisation]. *The compiled dataset cannot be redistributed due to licensing restrictions.*

...is not available due to [indicate reasons, e.g., licenses, sensitive data on human subjects, privacy statements; if there are processes to obtain the data, describe them].

The computational workflow supporting this publication …

… is executed via [choose, e.g., a single command/file, a workflow management software, a set of numbered scripts] *published under license* [the license] *at* [DOI of repository].

… is published in a [language] *module/package at* [link of software project]. *The used version is archived at* [DOI of repository].

… is provided as a [container/VM] *published at* [DOI of repository] *with instructions included in the file README.md in the repository.*

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

The guidelines for data

REPRODUCIBLE AGILE

"What if…" and
Examples not shown

INCLUDING DATA IN RESEARCH PAPERS

	Minimum requirements	Recommended practices
What?	• All input data and configuration • Data description/documentation, including provenance, field or column types, etc. • If data is retrieved from an external source, documentation on collection queries and download steps	• Standardised, discipline-specific metadata[8] and ontologies to describe your data • Data download scripts
Where?	• Publish data in a public repository providing a DOI • Cite data (including date and version) in the paper	• Discipline- or data type-specific repository[9] • Include recommended citation in dataset description (unless already provided by repository) • Create a registration for OSF projects[10] and use the DOI to cite it
How?	• Use open data formats; export from proprietary format for publication • Specify the license	• Use plain text-based file formats

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

The guidelines for computational workflows

REPRODUCIBLE AGILE

INCLUDING COMPUTATIONAL WORKFLOWS IN RESEARCH PAPERS

	Minimum requirements	Recommended practices
What? Computational environment	• Describe the used environment and computational infrastructure, e.g., hardware specs, operating system • List software versions • Cite used software[14]	• Provide the actual environment, e.g., a Dockerfile + container[15] or a Virtual Machine (e.g., using OSGeo-Live) • Provide a pinned freeze of your dependencies (structured configuration files with dependency information) • Add a colophon or "reproducibility receipt"[16] to your notebooks • Installation and execution instructions for different operating systems
Computation steps	• Document the detailed steps in a text file and/or flowchart (every action/click) • Document expected execution times given computing power unless negligible • Ask a colleague to try out the instructions	• Scripts/models and a README file that explains their use • All figures are fully scripted and a peer has read your README's instructions (incl. interactive visualisations and interactive adjustments • Multi-panel plots are composited with scripts[17] • Software package with structured metadata[18], tests/CI[19], and a pipeline framework[20] or workflow language[21] • Live documents for analyses, e.g., Binder[22] • Live demo of APIs/online applications (e.g., anonymous cloud resources, such as Google Cloud Run or AWS) • Subset or a synthetic dataset for quick evaluation
Where?	• Repository providing a persistent identifier, e.g., a DOI or SWHID[23]	• Versioned code repository, such as GitHub or GitLab, and ongoing open development
How? Tools used	• Use generally available tools (avoid proprietary tools that are not available to reviewers and other researchers)	• Use and create Open Source tools • Cite core modules/tools/language used
Development practices	• Use clear licenses[24] that fit your environment • Follow one of "Good enough practices in scientific computing"[25]	• Follow all "Good enough practices.." Use development guidelines for your environment / language of choice (e.g., for R[26])

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

Scientific reviewer guidelines… concerning the reproducibility review only!

SCIENTIFIC REVIEWER GUIDELINES

This section clarifies the expectations and role of the scientific reviewer with respect to the reproducible paper guidelines. For information for the Reproducibility Reviewer, please see the following section.

Reproducibility is considered good scientific practice that provides input for the quality assessment of a paper. Therefore, reviewers of AGILE papers should be aware of the **author guidelines on reproducibility** and be familiar with the **reproducibility checklist**, as well as the expected content of the **mandatory data and software availability section.** Using this information, reviewers should evaluate the plausibility and completeness of the data and software availability documentation, and whenever possible and readily available **include feedback on reproducibility aspects** in their comments. Scientific reviewers are free to but **are not expected to attempt reproductions of computations**.

Data and software availability documentation provide an additional set of information for assessing the quality of research presented in a manuscript. Reviewers are asked to know about the AGILE reproducible paper guidelines and to consider the level of reproducibility reached in a manuscript. To do so, they shall assume the position of someone who would like to reproduce the submitted work to assess whether the provided material is likely to allow reproduction of the submitted work. Based on this impression, reviewers may challenge authors regarding the level of reproducibility reached, if any statements are made regarding reproducibility in a manuscript.

Scientific reviewers are not required to actually reproduce a manuscript, but, if the data and code are provided in an anonymous format, and if a reviewer attempts to reproduce all or parts of the submitted work, then they are asked to document the process and outcomes (see Reproducibility Reviewer Guidelines below). Please reach out to the reproducibility chair if you are keen on conducting a reproducibility review for a paper you are reviewing.

The peer review of AGILE papers is a fully anonymous peer review, i.e. authors and reviewers do not know each other's identity. Reviewers should be supportive to authors and consider potential limitations in access to resources due to anonymisation. Since the provision of information to help reproduction of a paper can accidentally lead to disclosure of an author's identity, the reviewers should not use any such additional information to the disadvantage of the authors. The reviewers' comments provided to the authors are expected to be neutral[28] and contribute to improved reproducibility of the reported findings.

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

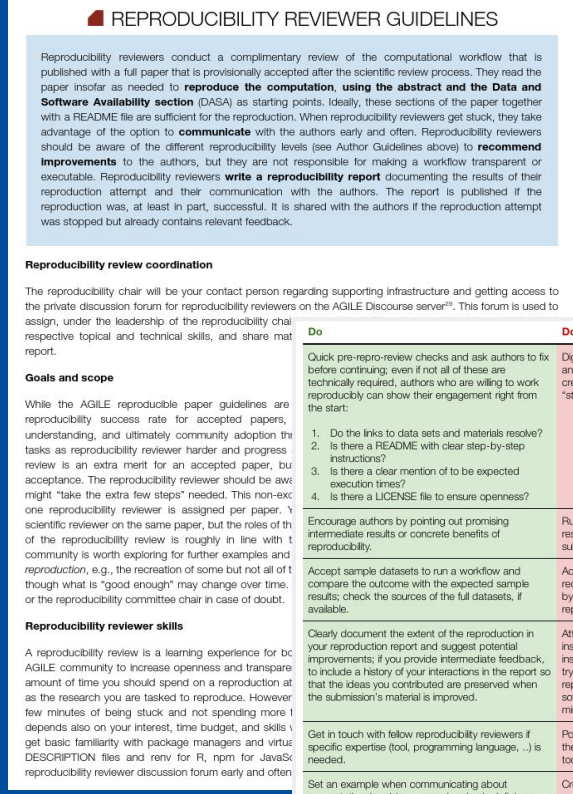
The guidelines for reproducibility reviewers

Ideal vs. realistic

Role & skills

Examples for “Do’s and Don’ts”:

- Do shift burden to author
- Do encourage and set examples
- Do *not* accept private data sharing
- Document your work in report (impact)
- Be kind (career stage, knowledge, privileges)
- No rummaging



REPRODUCIBILITY REVIEWER GUIDELINES

Reproducibility reviewers conduct a complimentary review of the computational workflow that is published with a full paper that is provisionally accepted after the scientific review process. They read the paper insofar as needed to **reproduce the computation**, **using the abstract and the Data and Software Availability section** (DASA) as starting points. Ideally, these sections of the paper together with a README file are sufficient for the reproduction. When reproducibility reviewers get stuck, they take advantage of the option to **communicate** with the authors early and often. Reproducibility reviewers should be aware of the different reproducibility levels (see Author Guidelines above) to **recommend improvements** to the authors, but they are not responsible for making a workflow transparent or executable. Reproducibility reviewers **write a reproducibility report** documenting the results of their reproduction attempt and their communication with the authors. The report is published if the reproduction was, at least in part, successful. It is shared with the authors if the reproduction attempt was stopped but already contains relevant feedback.

Reproducibility review coordination

The reproducibility chair will be your contact person regarding supporting infrastructure and getting access to the private discussion forum for reproducibility reviewers on the AGILE Discourse server[29]. This forum is used to assign, under the leadership of the reproducibility chai
respective topical and technical skills, and share mat
report.

Goals and scope

While the AGILE reproducible paper guidelines are
reproducibility success rate for accepted papers,
understanding, and ultimately community adoption th
tasks as reproducibility reviewer harder and progress
review is an extra merit for an accepted paper, bu
acceptance. The reproducibility reviewer should be awa
might “take the extra few steps” needed. This non-exc
one reproducibility reviewer is assigned per paper.
scientific reviewer on the same paper, but the roles of th
of the reproducibility review is roughly in line with
community is worth exploring for further examples and
reproduction, e.g., the recreation of some but not all of
though what is “good enough” may change over time.
or the reproducibility committee chair in case of doubt.

Reproducibility reviewer skills

A reproducibility review is a learning experience for bo
AGILE community to increase openness and transpare
amount of time you should spend on a reproduction at
as the research you are tasked to reproduce. However
few minutes of being stuck and not spending more
depends also on your interest, time budget, and skills
get basic familiarity with package managers and virtu
DESCRIPTION files and renv for R, npm for JavaS
reproducibility reviewer discussion forum early and often

Do	Don’t
Quick pre-repro-review checks and ask authors to fix before continuing; even if not all of these are technically required, authors who are willing to work reproducibly can show their engagement right from the start: 1. Do the links to data sets and materials resolve? 2. Is there a README with clear step-by-step instructions? 3. Is there a clear mention of to be expected execution times? 4. Is there a LICENSE file to ensure openness?	Dig across badly or un-documented collections of files and functions to identify which part of the code/data creates which figure/table/output; find or build the “start button” yourself.
Encourage authors by pointing out promising intermediate results or concrete benefits of reproducibility.	Run workflows requiring considerable computational resources (unless interesting for you) but ask for data subsets for demonstration purposes.
Accept sample datasets to run a workflow and compare the outcome with the expected sample results; check the sources of the full datasets, if available.	Accept private sharing of data or code, unless strictly required for protection of sensitive data. All changes by the author should update to the public reproduction material.
Clearly document the extent of the reproduction in your reproduction report and suggest potential improvements; if you provide intermediate feedback, to include a history of your interactions in the report so that the ideas you contributed are preserved when the submission’s material is improved.	Attempt to install software without any instructions, install binary software of unknown origin, or try to fix installation problems you encounter on your machine; try to install without (a) asking for help from a fellow reproducibility reviewer who is familiar with the software, or (b) asking the author to help, providing a minimal reproducible example of your problem.
Get in touch with fellow reproducibility reviewers if specific expertise (tool, programming language, ..) is needed.	Point out or even fix problems that are not specific to the submission, e.g., general problems in a software tool.
Set an example when communicating about computational problems, e.g., by clearly defining your system (OS version, language version, etc.)	Create accounts on any service or platform to access code, data, or other resources.
Ask specific questions or point out concrete problems that may lead authors to improve their material, including referencing these guidelines or concrete tools/methods that you already (!) know about, especially if you suspect that the author might now be familiar with them (e.g., version pinning/dependency management, absolute paths).	Fix anything (unless you really enjoy doing so), e.g., • compiler problems, • outdated libraries, • broken paths, or • Incomplete computing environment specifications, especially if the author can fix them even quicker.
Make sure that you are aware of any templates or specific resources provided for reproducibility reviewers from the reproducibility committee chair before starting your review.	
Consider the author’s background, career stage, and position to be aware of (a lack of) privileges or institutional power to decide how much support you provide and how you communicate; your reproducibility review can be a contribution to improve equity and inclusion in academia.	Be a bro.

General observations and lessons learned (2021!)

- **Further improvement over last years submissions - better prepared workflows!** Biggest hurdles remain: insufficient documentation, no “quick” variant or lack of expected data size/runtime, links Figures < > Scripts
- **Community understanding better, but needs time**: Had to remind authors to add DASA section - how can we be clearer in the communication? Camera-ready papers by authors possible, but exhausting.
- Additional **reproducibility questions for scientific reviewers worked better**, but triggering only by regular reviewers doesn’t work well - fortunately not too many submission to check for repro chair
- Repro reviews **were less strict than original ideal but on par with last year**
  > promote positive examples and don’t expect perfection
- **Non-blindness** served its purpose, but unblinding also delayed procedures
- **Schedule** still very much a challenge, partly because infrastructure (EasyChair) does not enable reviewer roles and communication > working around that with scripts and scraping
- **Improvements to process were good**: clarity in communication for authors that **DASA section is mandatory**, not attempting short papers, do not offer authors to object to report publications (no problems!)
- **Reproduction not attempted != bad science**, reproducibility is not binary but a ***spectrum***
  > continue education on reproducibility, increase requirements while practices spread in community

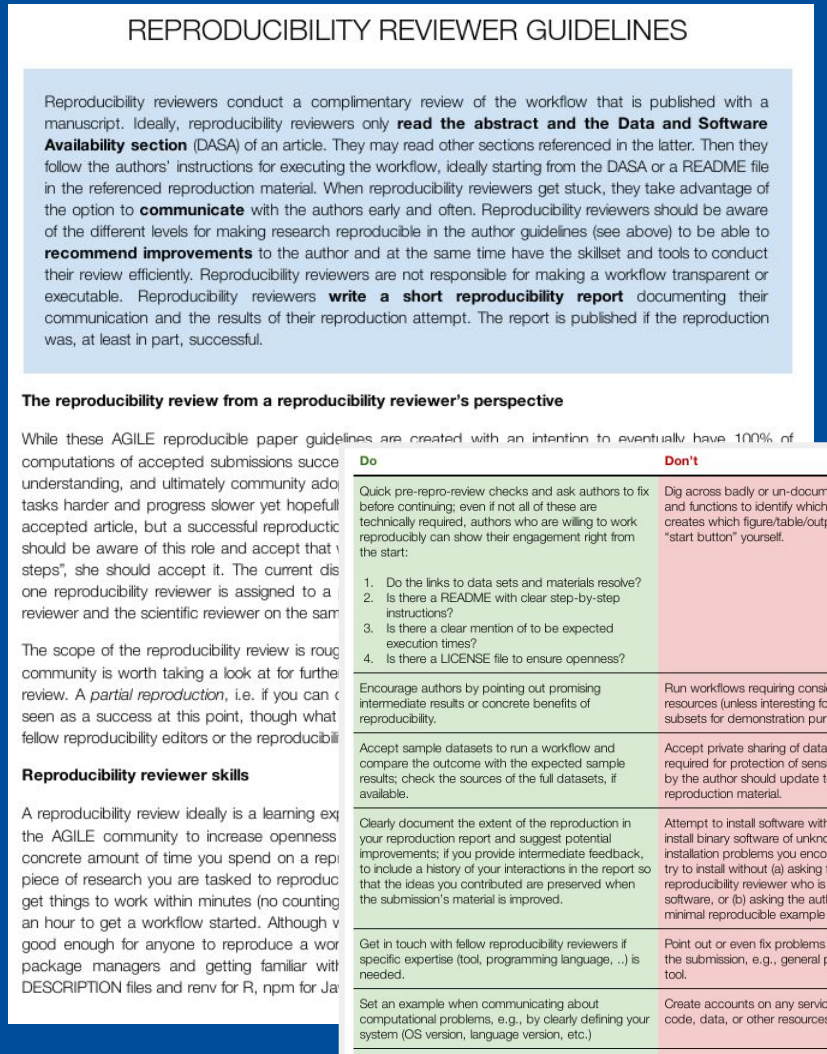
The guidelines for reproducibility reviewers

Ideal vs. realistic

Role

Skills

Do's & dont's



REPRODUCIBILITY REVIEWER GUIDELINES

Reproducibility reviewers conduct a complimentary review of the workflow that is published with a manuscript. Ideally, reproducibility reviewers only **read the abstract and the Data and Software Availability section** (DASA) of an article. They may read other sections referenced in the latter. Then they follow the authors' instructions for executing the workflow, ideally starting from the DASA or a README file in the referenced reproduction material. When reproducibility reviewers get stuck, they take advantage of the option to **communicate** with the authors early and often. Reproducibility reviewers should be aware of the different levels for making research reproducible in the author guidelines (see above) to be able to **recommend improvements** to the author and at the same time have the skillset and tools to conduct their review efficiently. Reproducibility reviewers are not responsible for making a workflow transparent or executable. Reproducibility reviewers **write a short reproducibility report** documenting their communication and the results of their reproduction attempt. The report is published if the reproduction was, at least in part, successful.

The reproducibility review from a reproducibility reviewer's perspective

While these AGILE reproducible paper guidelines are created with an intention to eventually have 100% of computations of accepted submissions succe
understanding, and ultimately community ado
tasks harder and progress slower yet hopefull
accepted article, but a successful reproductic
should be aware of this role and accept that
steps", she should accept it. The current dis
one reproducibility reviewer is assigned to a
reviewer and the scientific reviewer on the sam

The scope of the reproducibility review is roug
community is worth taking a look at for furthe
review. A *partial reproduction*, i.e. if you can
seen as a success at this point, though what
fellow reproducibility editors or the reproducibi

Reproducibility reviewer skills

A reproducibility review ideally is a learning ex
the AGILE community to increase openness
concrete amount of time you spend on a rep
piece of research you are tasked to reproduc
get things to work within minutes (no counting
an hour to get a workflow started. Although v
good enough for anyone to reproduce a wor
package managers and getting familiar wit
DESCRIPTION files and renv for R, npm for Ja

Do	Don't
Quick pre-repro-review checks and ask authors to fix before continuing; even if not all of these are technically required, authors who are willing to work reproducibly can show their engagement right from the start: 1. Do the links to data sets and materials resolve? 2. Is there a README with clear step-by-step instructions? 3. Is there a clear mention of to be expected execution times? 4. Is there a LICENSE file to ensure openness?	Dig across badly or un-documented collections of files and functions to identify which part of the code/data creates which figure/table/output; find or build the "start button" yourself.
Encourage authors by pointing out promising intermediate results or concrete benefits of reproducibility.	Run workflows requiring considerable computational resources (unless interesting for you) but ask for data subsets for demonstration purposes.
Accept sample datasets to run a workflow and compare the outcome with the expected sample results; check the sources of the full datasets, if available.	Accept private sharing of data or code, unless strictly required for protection of sensitive data. All changes by the author should update to the public reproduction material.
Clearly document the extent of the reproduction in your reproduction report and suggest potential improvements; if you provide intermediate feedback, to include a history of your interactions in the report so that the ideas you contributed are preserved when the submission's material is improved.	Attempt to install software without any instructions, install binary software of unknown origin, or try to fix installation problems you encounter on your machine; try to install without (a) asking for help from a fellow reproducibility reviewer who is familiar with the software, or (b) asking the author to help, providing a minimal reproducible example of your problem.
Get in touch with fellow reproducibility reviewers if specific expertise (tool, programming language, ..) is needed.	Point out or even fix problems that are not specific to the submission, e.g., general problems in a software tool.
Set an example when communicating about computational problems, e.g., by clearly defining your system (OS version, language version, etc.)	Create accounts on any service or platform to access code, data, or other resources.

How to put your community on a path towards more reproducibility in 5 ~~easy~~ hard steps

1. Build a team of enthusiasts (workshop, social events)
2. Assess the current state and raise awareness (workshop, paper)
3. Institutional support (🙏 AGILE Council 🙏 + committee chairs)
4. Positive encouragement (no reproduction != bad science)
5. Keep at it!