

# Code execution during peer review: CODECHECK and RSE reviewers

**16th HiRSE Seminar**

**Daniel Nüst**

Technische Universität Dresden (NFDI4Earth) | Reproducible AGILE

**<https://bit.ly/hirse-codecheck> | [10.5281/zenodo.7919697](https://zenodo.org/record/7919697)**



License: Create Commons  
Attribution 4.0 International (CC BY 4.0)  
<https://creativecommons.org/licenses/by/4.0/>

# Closed and irreproducible research



# Claerbout's claim:

An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures.

<https://doi.org/10.1190/1.1822162>

[https://doi.org/10.1007/978-1-4612-2544-7\\_5](https://doi.org/10.1007/978-1-4612-2544-7_5)

## How to draw an owl

1.



1. Draw some circles

2.



2. Draw the rest of the fucking owl

# HOW TO: DRAW A HORSE

BY VAN OKTOP

---



① DRAW 2 CIRCLES



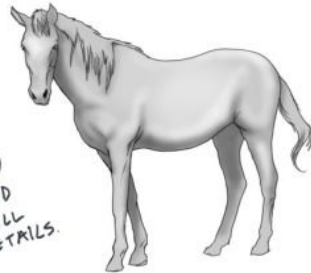
② DRAW THE LEGS



③ DRAW THE FACE



④ DRAW THE HAIR



⑤  
ADD  
SMALL  
DETAILS.

# One thing

**Have a README: all else is details.**

Show willingness to help, but don't stop publishing because lacking docs.

Hard to document for someone else > document for future you, add more on demand.

Inspired by Greg Wilson's Teaching Tech Together (<http://teachtogether.tech/en/index.html>) Rule 1.

# Four things on reproducible research

Have a README: all else is details.

Have a colleague run your workflow before submission.

Reproduce papers (demand material as reviewer or return the favour 🙌).

Publish code and data, cite it.

Rule 1 inspired by Greg Wilson's Teaching Tech Together (<http://teachtogether.tech/en/index.html>) Rule 1.

# The Turing Way >>> <https://the-turing-way.netlify.app/>



## The Turing Way

Search this book...

Welcome

Guide for Reproducible Research

Guide for Project Design

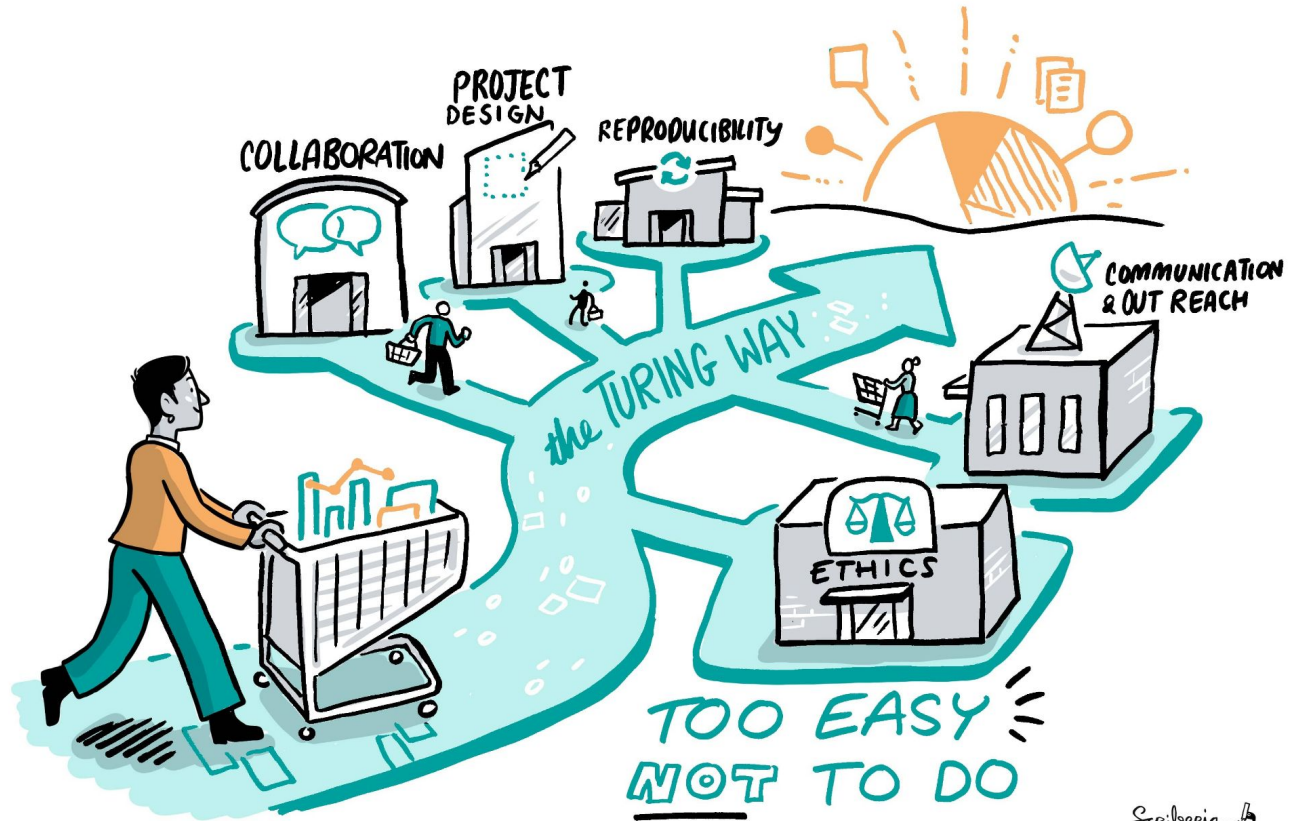
Guide for Communication

Guide for Collaboration

Guide for Ethical Research

Community Handbook

Afterword



The Turing Way project illustration by Scriberia. Zenodo. <http://doi.org/10.5281/zenodo.3332807>

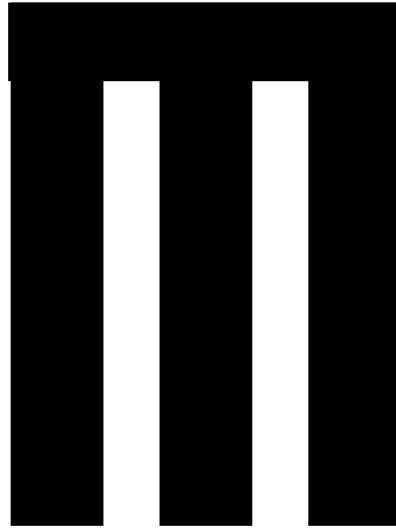


**<Excursion>**

# Traditional and modern scientists

Broad knowledge: across disciplines  
collaborate with other experts, apply outside of own field

Deep knowledge: expertise and  
skills within a single field



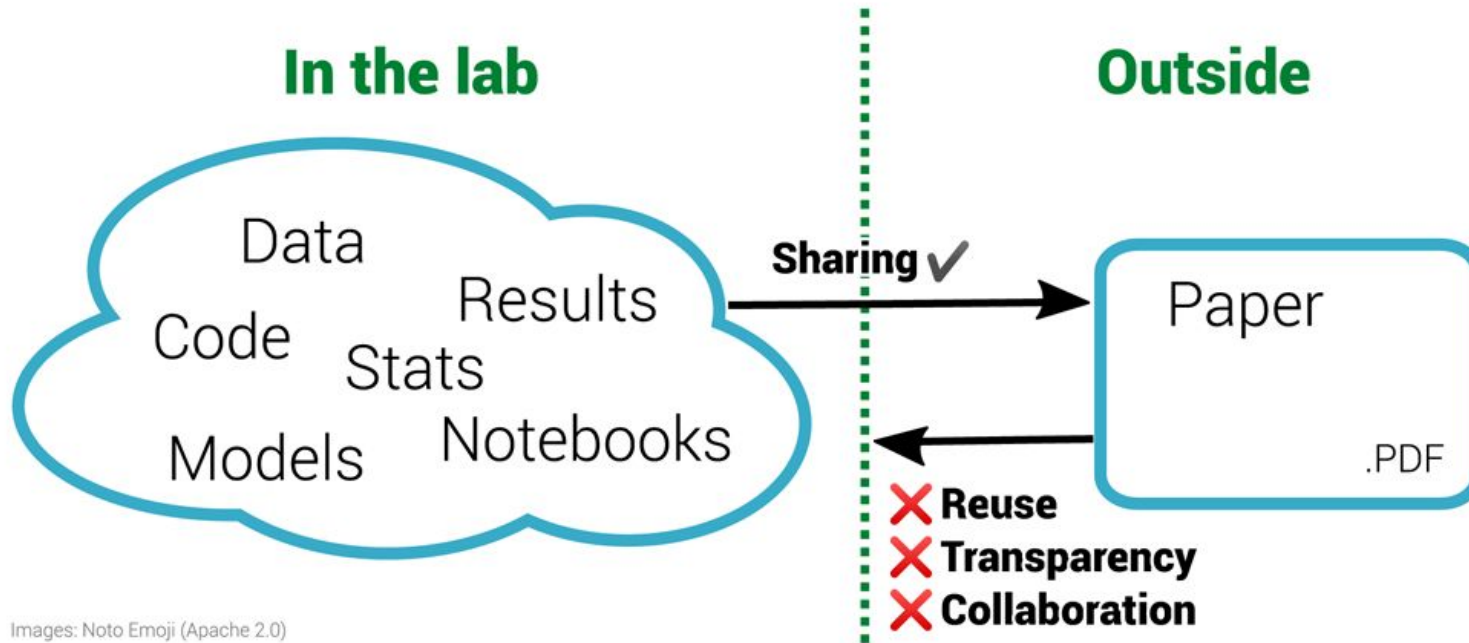
Computer & method skills  
statistics, reproducibility,  
programming, data science



**</Excursion>**

# **CODECHECK**

**[codecheck.org.uk](https://codecheck.org.uk)**



**The inverse problem in reproducible research. Figure 1 of <https://doi.org/10.12688/f1000research.51738.1>**

The left half of the diagram shows a diverse range of materials used within a laboratory. These materials are often then condensed for sharing with the outside world via the research paper, a static PDF document. Working backwards from the PDF to the underlying materials is impossible. This prohibits reuse and is not only non-transparent for a specific paper but is also ineffective for science as a whole. By sharing the materials on the left, others outside the lab can enhance this work.



# Research Compendia

- = programming language packaging + science stuff
- = templates
- = community practices  
(lab, discipline, language, method)

**research-compendium.science**



# Who will look into the research compendia?





# Professionalisation



#





Independent execution of computations underlying research articles.



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

## Principles

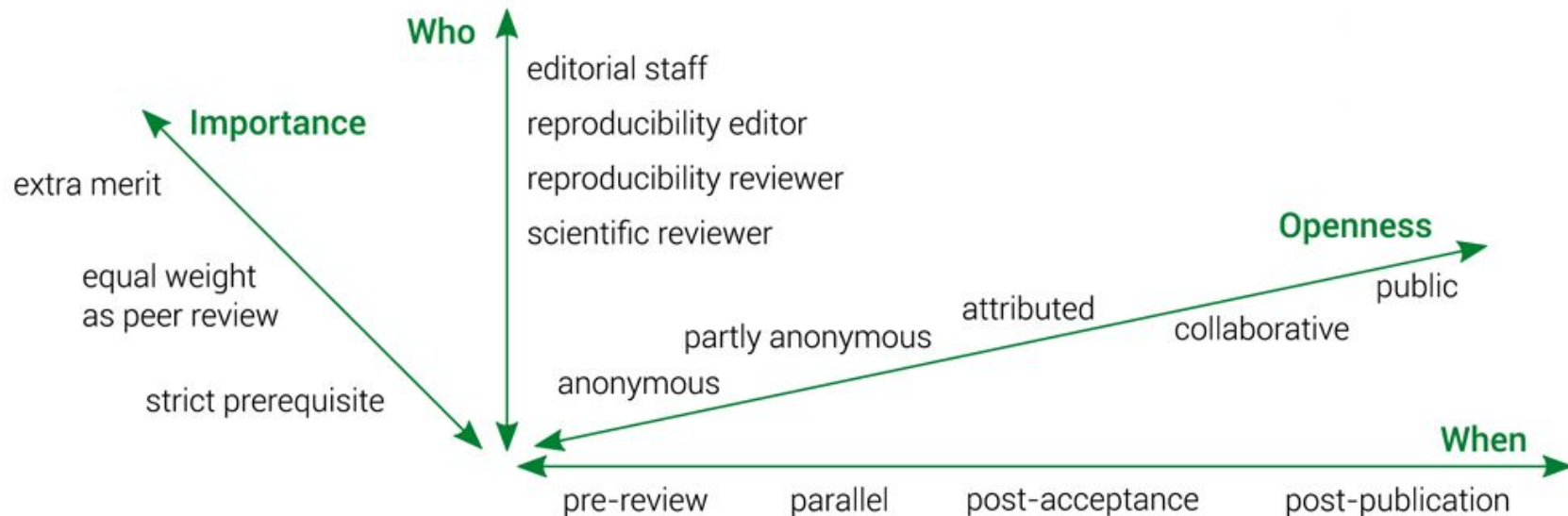
1. Codecheckers **record** but don't investigate or fix.
2. **Communication** between humans is key.
3. **Credit** is given to codecheckers.
4. Workflows must be **auditable**.
5. **Open** by default and **transitional** by disposition.



50+ Certificates

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

# One re-execution of computational workflow by codechecker during peer review



The CODECHECK variations. Figure 1 of <https://doi.org/10.12688/f1000research.51738.1>

# Reproducible AGILE



<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>

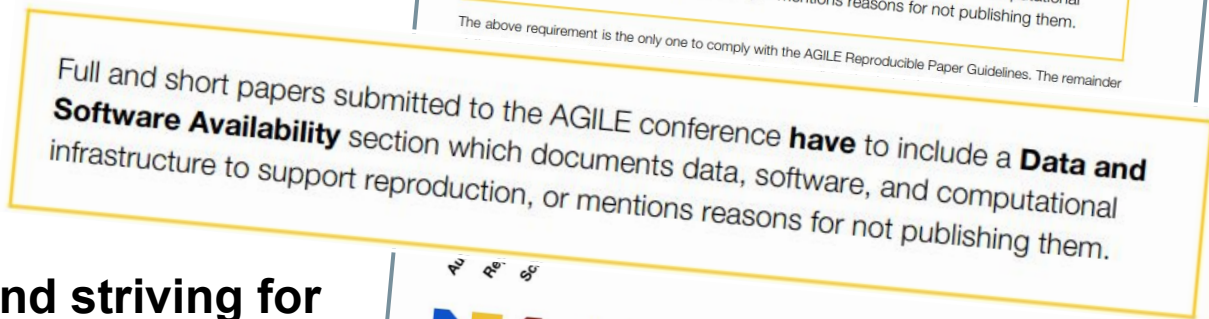
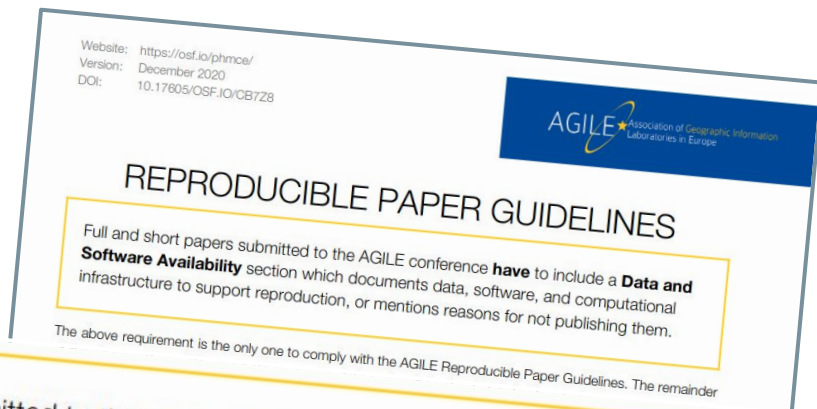
Created by AGILE Initiative in 2019 (see report at <https://osf.io/hupxr/>) and updated in 2020








**Transparency over enforcement**

**Promote, don't exclude**

**Acknowledge spectrum and striving for ideal**

**Tailored to GIScience**



	Au	Re	Sc		
				<b>Reproducibility Checklist</b> Helps to ensure authors and reviewers do not miss anything important.	2
				<b>Author Guidelines</b> Show how to write the Data and Software Availability Section and give practical recommendations to make data and computational workflows reproducible.	4
				Writing the Data and Software Availability Section Including Data in Research Papers Including Computational Workflows in Research Papers	
				<b>Scientific Reviewer Guidelines</b> Describe role in evaluating plausibility and completeness of the data and software availability documentation.	7
				<b>Reproducibility Reviewer Guidelines</b> Describe role in evaluating plausibility and completeness of the data and software availability documentation.	

# 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

## 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.

Author Reproducibility Reviewer Scientific Reviewer			
	<b>Reproducibility Checklist</b> Helps to ensure authors and reviewers do not miss anything important.		2
	<b>Author Guidelines</b> Show how to write the Data and Software Availability Section and give practical recommendations to make data and computational workflows reproducible.  Writing the Data and Software Availability Section Including Data in Research Papers Including Computational Workflows in Research Papers		4
	<b>Scientific Reviewer Guidelines</b> Describe role in evaluating plausibility and completeness of the data and software availability documentation.		7
	<b>Reproducibility Reviewer Guidelines</b> Describe role and approach to execute workflows and clarify efforts.		8
	<b>Background</b>		10

### 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](#).

# Review process



Proceedings: [agile-giscience.net/review\\_process.html](https://agile-giscience.net/review_process.html)  
Process documentation: [osf.io/7rjpe/](https://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!)**

The screenshot shows the article page for 'Window Operators for Processing Spatio-Temporal Data Streams on Unmanned Vehicles' by Tobias Werner and Thomas Brinkhoff. The page includes a navigation menu on the left, a search bar at the top right, and a list of related articles at the bottom. The article title is highlighted in blue. The 'reproducibility' badge is visible in the top right corner of the article content area. The page number '24' is visible in the bottom right corner.



# 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 this report use

Decoupes, R. (2022, May 23). 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

Cui, 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-gisc-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 Training* 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).

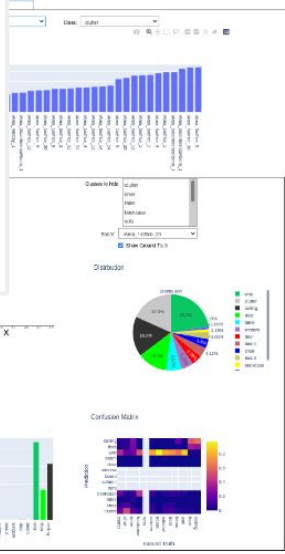


Figure 2: visual-exploration-dash-board-pt2.ipynb: Tab2. Partially reproduction of Figure 4.

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

Jakub Krnkar   
2022-06-09



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/MVQCW>. To cite this report use

Krnkar, J. (2022, May 23). Reproducibility review of: Unlocking social network analysis methods for studying human mobility. <https://doi.org/10.17605/OSF.IO/MVQCW>

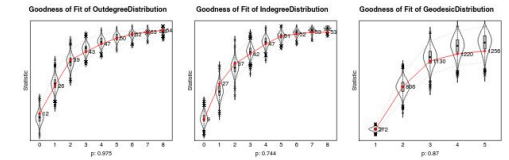
## Reviewed paper

Wiseman, N., Martin, H., and Raubal, M.: Unlocking social network analysis methods for studying human mobility, AGILE GIScience Ser., 3, 19, <https://doi.org/10.5194/agile-gisc-3-19-2022>

## Summary

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

files seen 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.

# Reproducibility Reports

Published with a DOI

Title page, cites the paper


Paper links to report via URL/badge (no citation)

Automatically added to ORCID profile

Reproducibility review of: Investigating drivers' geospatial abilities in unfamiliar environments

Philipp A. Friese

2021-06-07



#### 2.4 Data and Software Availability

Questionnaires and sketches were collected anonymously. All statistical analyses, which results are detailed in the following section, have been performed in R (R Core Team, 2021) using the tidyverse package (Wickham et al., 2019). Driving directions given to participants, an Exemplary Questionnaire in English, the collected survey data in tabular form, the R code of the statistical analysis workflow, and all necessary metadata supporting this publication, are available on figshare and are accessible via the following DOI: <https://doi.org/10.6084/m9.figshare.14460102.v4>. The workflow underlying this paper was successfully reproduced by an independent reviewer during the AGILE reproducibility review and a [reproducibility report](https://doi.org/10.17605/OSF.IO/DX92A) was published at <https://doi.org/10.17605/OSF.IO/DX92A>.

#### 3 Results

Three measures were evaluated corresponding to the tasks performed: map sketching, distance estimates, and direction estimates. The results of the SBSOD

at the AGILE conference. For more information see the report published on OSF at <https://osf.io/dx92a>. To

Reproducibility review of: Investigating drivers' geospatial abilities in unfamiliar environments /doi.org/10.17605/OSF.IO/DX92A

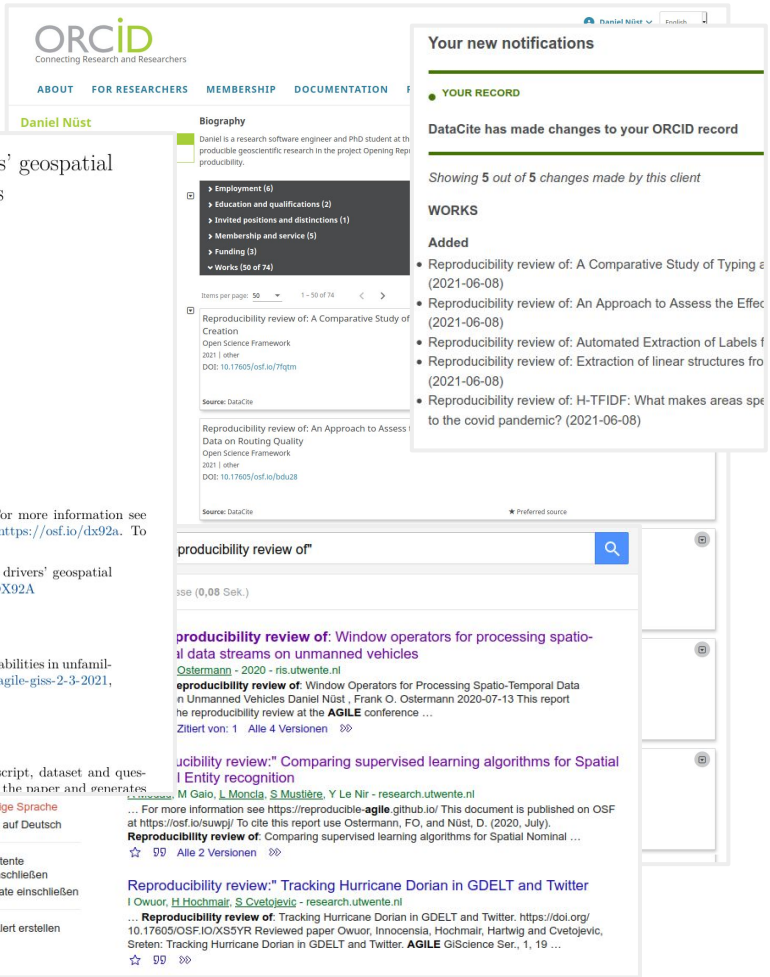
Investigating drivers' geospatial abilities in unfamiliar environments, 3, <https://doi.org/10.5194/agile-giss-2-3-2021>, 2021. doi:10.5194/agile-giss-2-3-2021

and provides the analysis script, dataset and statistical analyses presented in the paper and generates

Bellebige Sprache  
Seiten auf Deutsch

Patente einschließen  
 Zitate einschließen

Alert erstellen



The screenshot shows the ORCID profile of Daniel Nüst, a research software engineer and PhD student. The profile includes a biography, a list of works, and a notification from DataCite. A search for 'reproducibility review of' yields results for 'Window operators for processing spatial data streams on unmanned vehicles' and 'Comparing supervised learning algorithms for Spatial Entity recognition'. The search interface includes a search bar, a results list, and a sidebar with filters.

**ORCID**  
Connecting Research and Researchers

ABOUT FOR RESEARCHERS MEMBERSHIP DOCUMENTATION

Daniel Nüst

**Biography**

Daniel is a research software engineer and PhD student at the University of Applied Sciences, Technikum Wien, where he produces geoscientific research in the project Opening Reproducibility.

- Employment (6)
- Education and qualifications (2)
- Invited positions and distinctions (1)
- Membership and service (5)
- Funding (3)
- Works (50 of 74)

Items per page: 50 1 - 50 of 74

Reproducibility review of: A Comparative Study of Open Science Framework Creation  
2021 | other  
DOI: 10.17605/OSF.IO/7RQPM  
Source: DataCite

Reproducibility review of: An Approach to Assess the Quality of Data on Routing Quality  
Open Science Framework  
2021 | other  
DOI: 10.17605/OSF.IO/bdu28  
Source: DataCite

**Your new notifications**

**YOUR RECORD**

**DataCite has made changes to your ORCID record**

Showing 5 out of 5 changes made by this client

**WORKS**

**Added**

- Reproducibility review of: A Comparative Study of Typing and Writing (2021-06-08)
- Reproducibility review of: An Approach to Assess the Effectiveness of Open Science Frameworks (2021-06-08)
- Reproducibility review of: Automated Extraction of Labels from Geospatial Data (2021-06-08)
- Reproducibility review of: Extraction of linear structures from Geospatial Data (2021-06-08)
- Reproducibility review of: H-TFIDF: What makes areas special to the covid pandemic? (2021-06-08)

reproducibility review of"

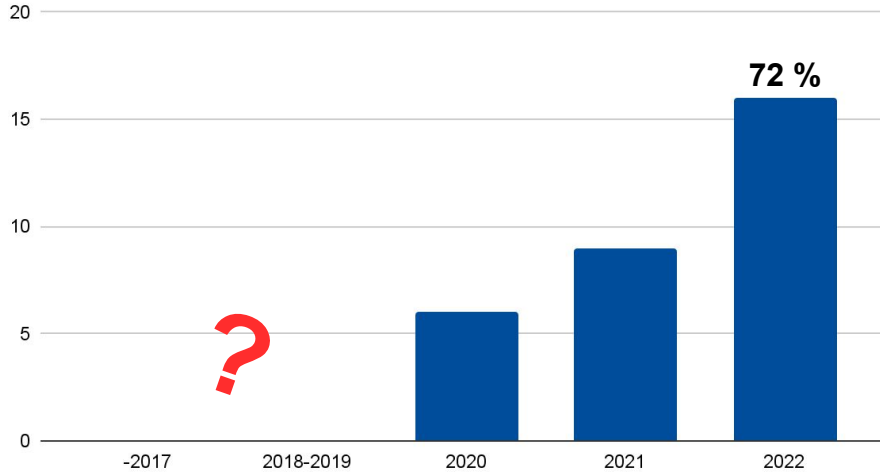
Reproducibility review of: Window operators for processing spatial data streams on unmanned vehicles  
Ostermann - 2020 - ris.utwente.nl  
Reproducibility review of: Window Operators for Processing Spatio-Temporal Data from Unmanned Vehicles Daniel Nüst, Frank O. Ostermann 2020-07-13 This report is a reproducibility review at the AGILE conference ...  
Zitiert von: 1 Alle 4 Versionen

Reproducibility review: "Comparing supervised learning algorithms for Spatial Entity recognition"  
M Galo, L Mondrago, S Mustiata, Y Le Nir - research.utwente.nl  
... For more information see <https://reproducible-agile.github.io/> This document is published on OSF at <https://osf.io/suwp/> To cite this report use Ostermann, FO, and Nüst, D. (2020, July).  
Reproducibility review of: Comparing supervised learning algorithms for Spatial Nominal ...  
Alle 2 Versionen

Reproducibility review: "Tracking Hurricane Dorian in GDELT and Twitter"  
Owuor, H Hochmair, S Cvetojevic - research.utwente.nl  
...  
Reproducibility review of: Tracking Hurricane Dorian in GDELT and Twitter. <https://doi.org/10.17605/OSF.IO/XS5YR> Reviewed paper Owuor, Innocensia, Hochmair, Hartwig, and Cvetojevic, Sreten: Tracking Hurricane Dorian in GDELT and Twitter. *AGILE GIScience Ser.*, 1, 19 ...  
Alle 2 Versionen

# Reproducibility review results

## Reproduced full papers



## 6 not reproducible:

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

Reproducibility review of: GeoXTag: Relative Spatial Information Extraction and Tagging of Unstructured Text Krukar	...
Reproducibility review of: Experimental evaluation of using BLE beacon for outdoor positioning in GPS-denied environment Nüst	...
Reproducibility review of: A method to produce metadata describing and assessing the quality of spatial landmark datasets in mountain area Friese Reproduction report and material.	...
Reproducibility review of: Understanding COVID-19 Effects on Mobility: A Community-Engaged Approach Friese Reproduction report and material.	...
Reproducibility Review of: Traffic Regulation Recognition using Crowd-Sensed GPS and Map Data: a Hybrid Approach Koukouraki	...
Reproducibility review of: Exploratory Analysis and Feature Selection for the Prediction of Nitrogen Dioxide Koukouraki	...
Reproducibility review of: Spatial Disaggregation of Population Subgroups Leveraging Self-Trained Multi-Output Gradient Boosted Regression Trees Ostermann	...
Reproducibility review of: Landmark Route - A Comparison to the Shortest Route Ostermann	...
Reproducibility review of: Unlocking social network analysis methods for studying human mobility Krukar	...
Reproducibility review of: A machine learning based approach for predicting usage efficiency of shared e-scooters using vehicle availability data Graneli	...
Reproducibility review of: Optimizing Electric Vehicle Charging Schedules Based on Probabilistic Forecast of Individual Mobility Graneli	...
Reproducibility review of: The Impact of Built Environment on Bike Commuting: Utilising Strava Bike Data and Geographically Weighted Models Decoupes	...
Reproducibility review of: Understanding the Imperfection of 3D point Cloud and Semantic Segmentation algorithms for 3D Models of Indoor Environment Decoupes	...
Reproducibility review of: Benchmarking Invasive Alien Species Image Recognition Models for a Citizen Science Based Spatial Distribution Monitoring Nüst	...
Reproducibility review of: Geoparsing: Solved or Biased? An Evaluation of Geographic Biases in Geoparsing Nüst & Tomai	...
Reproducibility review of: What are intersections for pedestrian users? Tomai & Nüst Reproduction report and material	✕



<https://osf.io/r5w79/>

# From **AGILE** to **EVERYWHERE**

CODECHECK is **transferable to your discipline:**  
human-centered process, communicative focus, supportive  
framework of tools and infrastructure, and candidates.



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



Learn crafts by **doing**. Be kind. Help.



**Human(with RSE skills)-centered process 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 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

A top-down view of a white workbench or toolbox filled with a wide variety of tools. The tools are organized in rows and columns. At the top, there are many wrenches of different sizes. Below them are pliers, screwdrivers, and other hand tools. Some tools have blue handles, while others are made of metal. The background is a plain white surface, making the tools stand out.

Faster codechecking with...

Notebooks

Binder (BinderHub)

Open methods

Missing?!

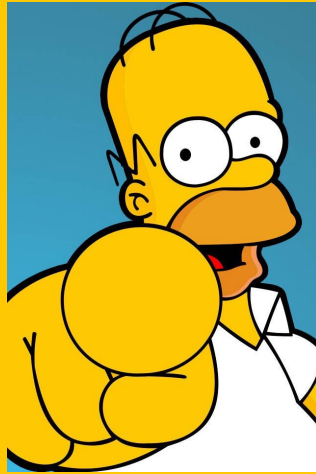
- free Binder-ready notebooks **close** to the data
- **metadata** connectivity > credit & partial re-building
- (small) reference datasets & demo pipelines



# Candidates to apply CODECHECK?



What can you do today?



## What can you do?

**Strive** for best reproducibility when sharing research

Use RSEng expertise as peer-reviewer and **be demanding**

Initiate **discourse** on code execution if you are an author, reviewer, editor, ...

**Become** a CODECHECKER (and ~~wait until~~ help us to get the thing flying)

# What do you get?

You can **help** people to write code that can be understood by others, can be reused, and can be improved. Your feedback will be **welcome** and valued by researchers.

You can give back to the **community**. You can **reduce** people's **pain** by helping to improve "academic" codebases. Spread **good practices** in academic software use and development. You help to increase **availability** and **usefulness** of code and data in your discipline and others.

It **connects** you with a journal's community, experiences in peer review (seeing the latest stuff) & **reviewer credit**, gives **recognition** in academia.

You get to know all the **package managers**. All of them. You **learn** useful tools and practices for your **own work**.

Understanding other people's code trains your **skills**, peeking into fields of research outside of your own.

You **meet** possible collaborators and foster collaboration & reuse over **reinventing** the wheel.

You can **push** publishing practices in the right direction and contribute to establishing a more open and friendly **culture** of reproducible research.

These statements are inspired by motivations shared by the awesome volunteer codechecker community:  
<https://github.com/codecheckers/codecheckers>

*improves peer  
review to foster*

**BETTER  
SOFTWARE  
BETTER  
RESEARCH**

*checking of  
by RSEs*

Thank you!

Questions?

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

<https://bit.ly/hirse-codecheck>

10.5281/zenodo.7919697

Slides published under CC BY 4.0



WHO'S AWESOME?

