AGILE Conference Reproducibility Review 2021

Daniel Nüst, AGILE 2021 Reproducibility Chair
daniel.nuest@uni-muenster.de
https://doi.org/10.5281/zenodo.4926269

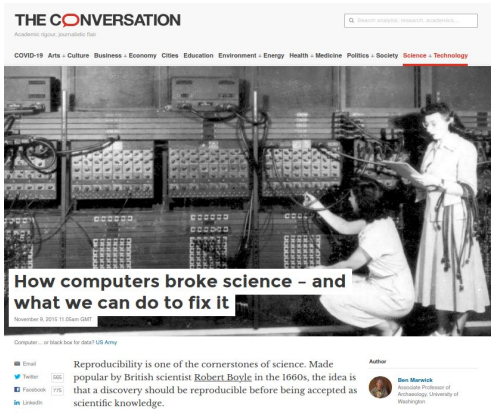
AGILE Association of Geographic Information Laboratories in Europe

REPRODUCIBLE AGILE

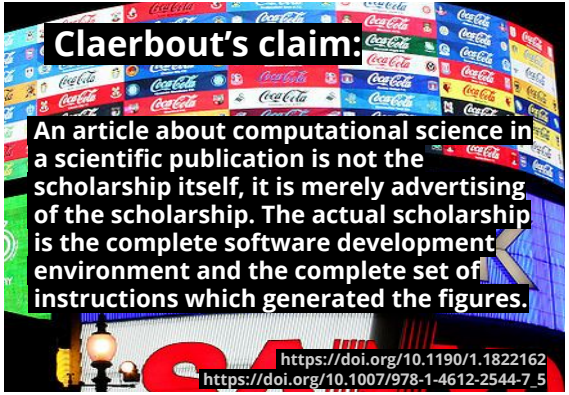
https://reproducible-agile.github.io/

2017, ‘18 & ‘19: Workshops on reproducibility
2019: Reproducible publications at AGILE conferences (initiative)
2020: First AGILE reproducibility review
2021: Second AGILE reproducibility review

https://giphy.com/gifs/usnationalarchives-nasa-scientist-scientists-1F1JGyGZhiSAA8Vuhn



https://theconversation.com/how-computers-broke-science-and-what-we-can-do-to-fix-it-49938



CC-BY 3.0, Sebastian Bertalan, Wikimedia Commons

https://giphy.com/gifs/with-computers-fascination-PxSFAnuubLkSA

https://giphy.com/gifs/david-hasselhoff-M3o3fL9nnxG4o

Crisis? Crisis of what?

Credibility crisis?

Replicability crisis?

Reproducibility crisis?

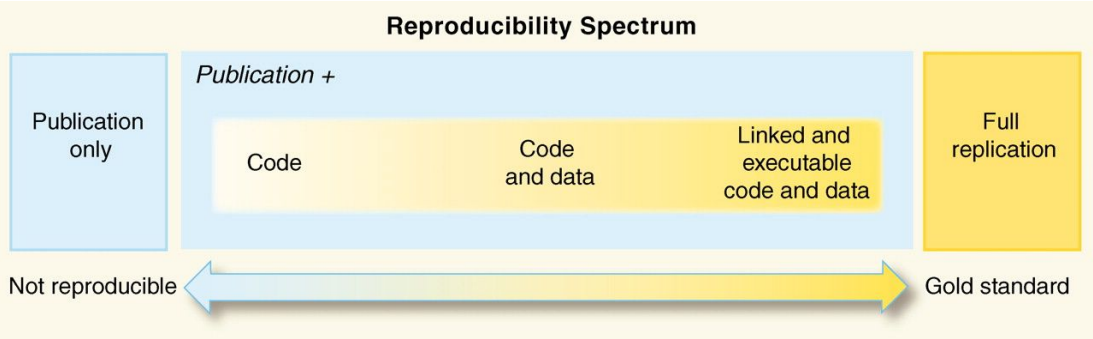
Robustness crisis?

Generalisability crisis?

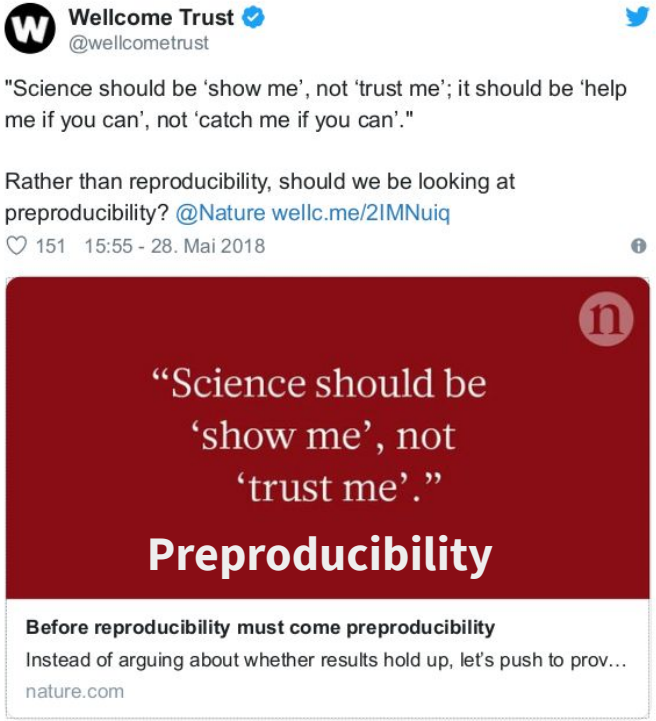
		Data	
		Same	Different
Analysis	Same	Reproducible	Replicable
	Different	Robust	Generalisable

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

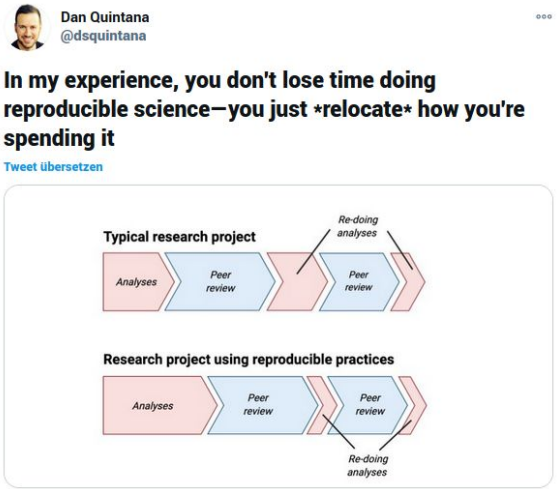
Reproducible Research & Open Science

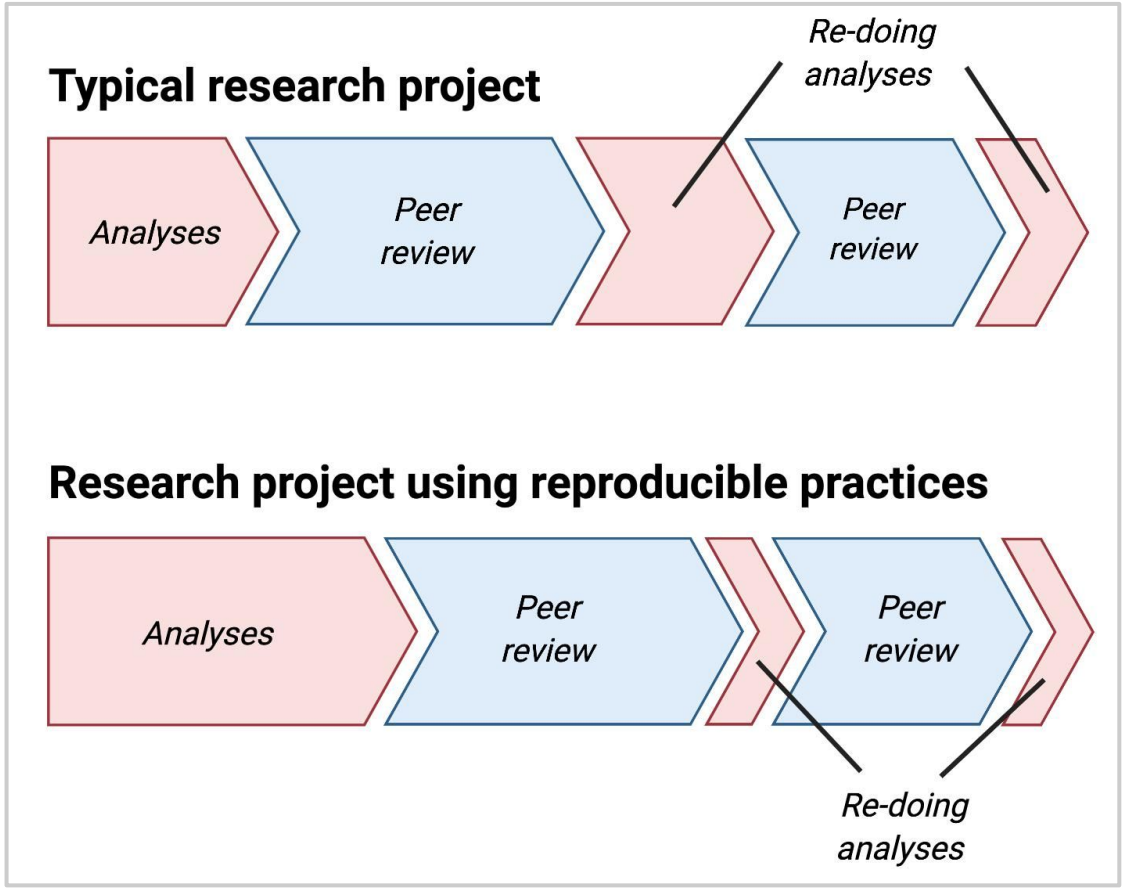


https://doi.org/10.1126/science.1213847



https://doi.org/10.1038/d41586-018-05256-0





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

AGILE Reproducible Paper Guidelines: Contents & First Revision

AGILE Reproducible Paper Guidelines

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

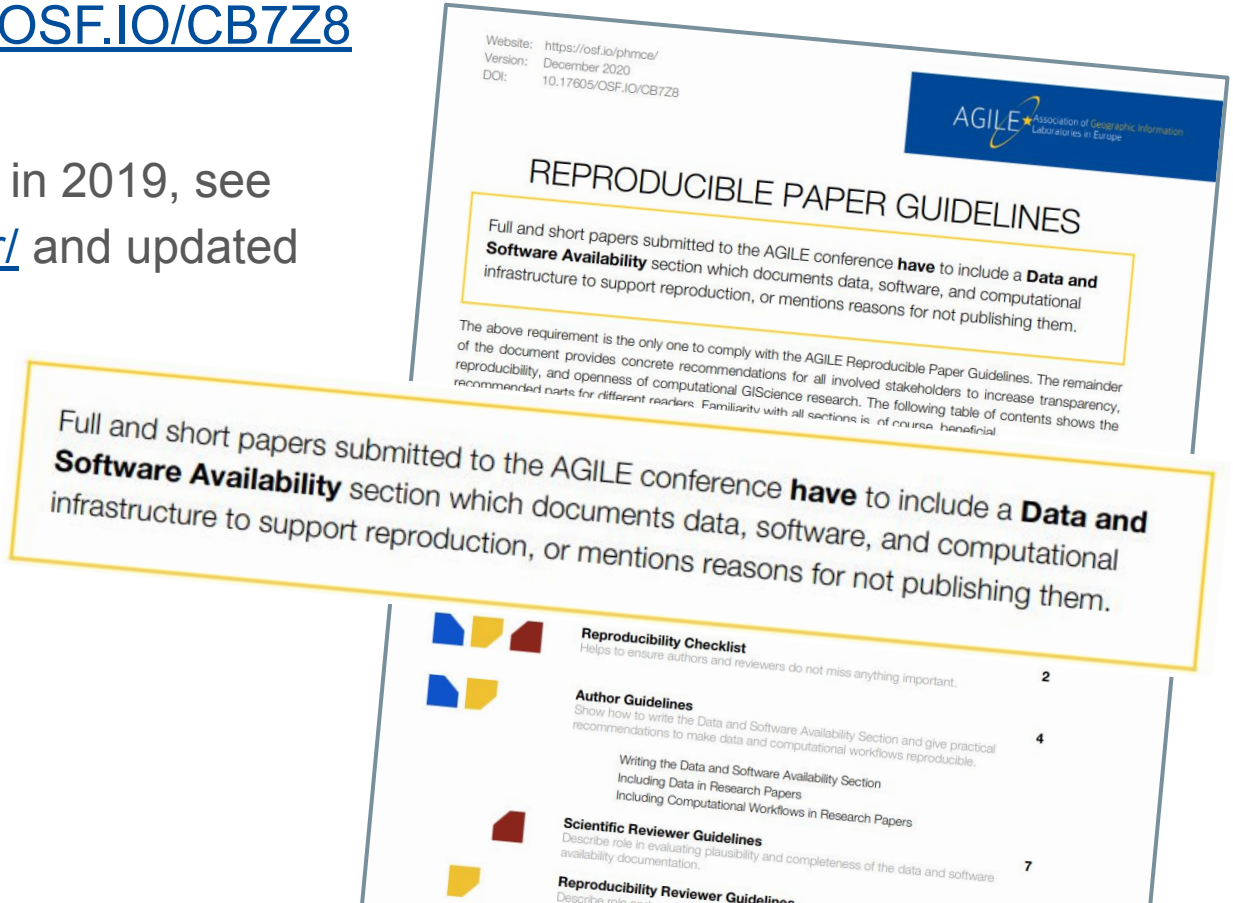
Created by AGILE Initiative in 2019, see report at https://osf.io/hupxr/ and updated in 2020

Transparency

Promotion

Acknowledge spectrum

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.

Notable updates in version 2 (December 2020)

Updated checklist (clarity + structure, extra items for after acceptance)

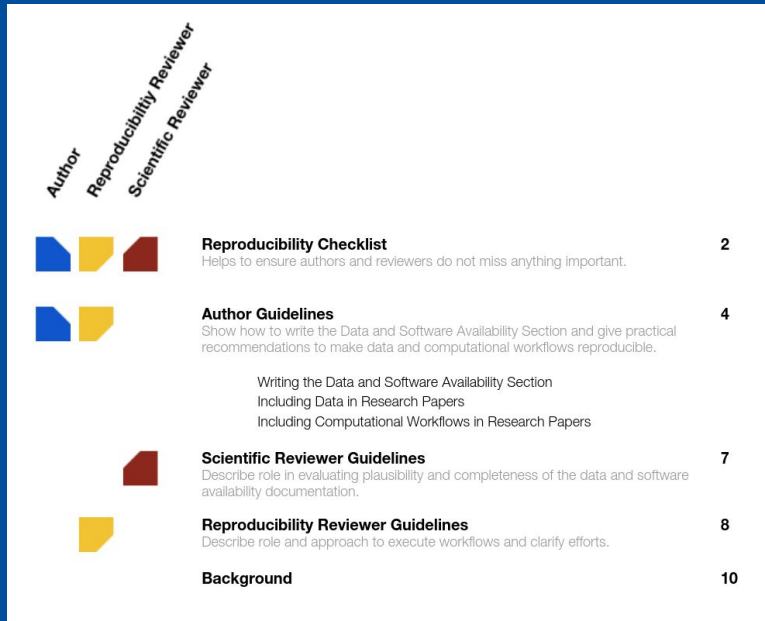
Reordering of contents (important stuff first!) and details moved to Wiki

TOC + visual aides for reader roles

Updated resources + language improvements

Removed “intermediate” level for computational workflows

NEW: Reproducibility reviewer guidelines



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

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

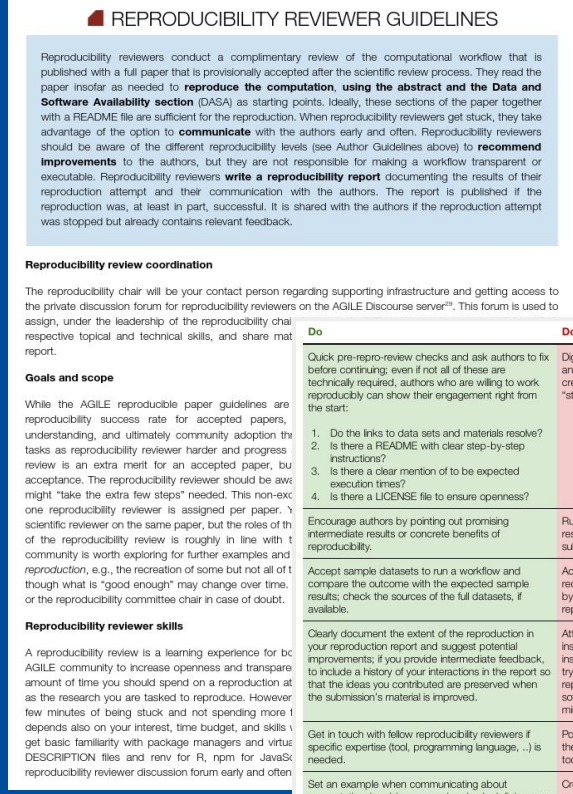
The guidelines for reproducibility reviewers

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.

Thanks!

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:
https://www.agile-giscience-series.net/review_process.html

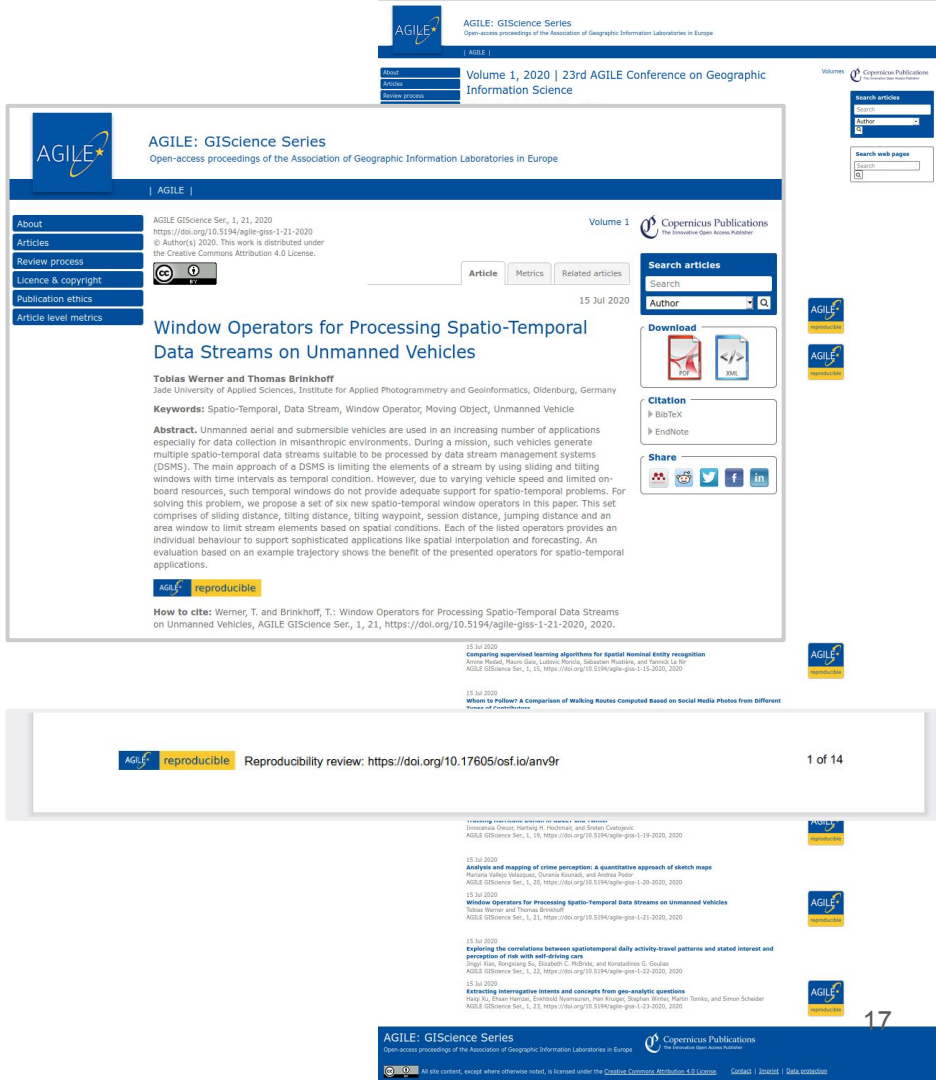
Process documentation:
https://osf.io/7rjpe/

Reproducibility review *after* accept/reject decisions

Reproducibility review & communication

Community conference & volunteers

Badges on proceedings website, article website with link, and first article page **(NEW! Thanks you, Copernicus!)**



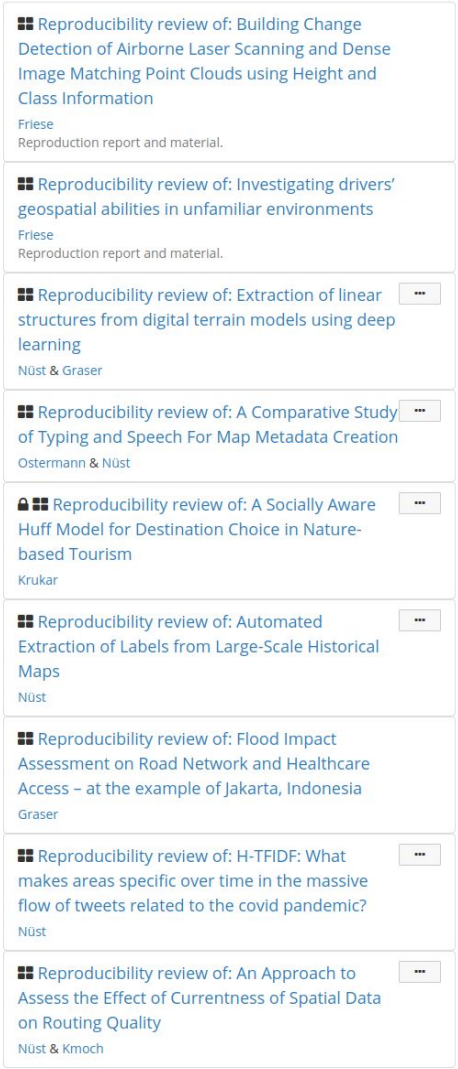
Reproducibility Review at AGILE Conference 2021

Reproducibility review results

9 reproducibility reports published (2020: 6)

8 not reproducible:

- 3 conceptual papers
- **data not shared (choice, licence)**
  - synthetic data! subsets!
- **code** not shared (choice) or proprietary software (repro reviewer matching failed)



https://osf.io/h64sd/

Reproducibility Reports

Published on OSF with a DOI
Title page, cites the paper

Paper links to report via URL
(no citation)

Automatically added to ORCID profile

Eventually indexed in GS

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

Philipp A. Friese

2021-06-07

REPRODUCIBLE AGILE

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://osf.io/dx92a. To cite the report use

Friese, Philipp A. (2021, May). Reproducibility review of: Investigating drivers' geospatial abilities in unfamiliar environments. https://doi.org/10.17605/OSF.IO/DX92A

Reviewed paper

Karkasina, D., Kokla, M., and Tomai, E.: Investigating drivers' geospatial abilities in unfamiliar environments, AGILE GIScience Ser., 2, 3, https://doi.org/10.5194/agile-giss-2-3-2021, 2021.

Summary

The updated submiss... tionnaires. The provi... , dataset and ques... paper and generates ...n...

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 was published at https://doi.org/10.17605/OSF.IO/DX92A.

ORCID
Connecting Research and Researchers

ABOUT FOR RESEARCHERS MEMBERSHIP DOCUMENTATION

Daniel Nüst

Biography

Daniel is a research software engineer and PhD student at th... producible geoscientific research in the project Opening Repr... producibility.

› 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 Typing ... Creation
Open Science Framework
2021 | other
DOI: 10.17605/osf.io/7fqtm
Source: DataCite

Reproducibility review of: An Approach to Assess ... Data on Routing Quality
Open Science Framework
2021 | other
DOI: 10.17605/osf.io/bdu28
Source: DataCite ★ Preferred source

Reproducibility review of: Automated Extraction of Labels from Large-Scale Historical Maps
Open Science Framework
2021 | other
DOI: 10.17605/osf.io/anv9r

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 ... (2021-06-08)
- Reproducibility review of: An Approach to Assess the Effec... (2021-06-08)
- Reproducibility review of: Automated Extraction of Labels f...
- Reproducibility review of: Extraction of linear structures fro... (2021-06-08)
- Reproducibility review of: H-TFIDF: What makes areas spe... to the covid pandemic? (2021-06-08)

...cholar

agile "reproducibility review of"

4 Ergebnisse (0,08 Sek.)

Nach Relevanz sortieren
Nach Datum sortieren

Beliebige Sprache
Seiten auf Deutsch

Patente einschließen
Zitate einschließen

Alert erstellen

[PDF] **Reproducibility review of**: Window operators for processing spatio-temporal data streams on unmanned vehicles
D Nüst, F Ostermann - 2020 - ris.utwente.nl
Page 1. **Reproducibility review of**: Window Operators for Processing Spatio-Temporal Data Streams on Unmanned Vehicles Daniel Nüst , Frank O. Ostermann 2020-07-13 This report is part of the reproducibility review at the **AGILE** conference ...
Zitiert von: 1 Alle 4 Versionen

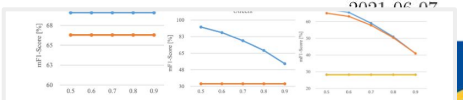
Reproducibility review:" Comparing supervised learning algorithms for Spatial Nominal Entity recognition
A Medad, M Gaio, L Moncla, S Mustière, Y Le Nir - research.utwente.nl
... For more information see https://reproducible-**agile**.github.io/ This document is published on OSF at https://osf.io/suwpj/ 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
I 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 ...

Reproducibility review reports AGILE 2021

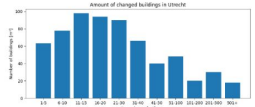
Reproducibility review of: Building Change Detection of Airborne Laser Scanning and Dense Image Matching Point Clouds using Height and Class Information

Philipp A. Friese



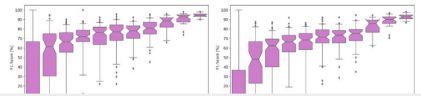
Graphic 1: Mean F1 Scores from Excel Sheet - corresponds to Figure 2-7 and methodologically to 12-15 in reproduced paper

After the authors provided an additional visualization script Figure 1 was reproduced. The generated image is shown in graphic 2.



Graphic 2: Generated Amounts of changed buildings - corresponds to Figure 1 in reproduced paper

The script generates Figure 10 and 11 automatically for each parameter permutation. An excerpt of the generated images are shown in graphic 3.



ILE conference. For more information see
olished on OSF at https://osf.io/RSF4M/.

Reviewed paper

Politz, F., Sester, M., and Brei
ning and Dense Image Matchi
GIScience Ser., 2, 10, https://

Graphic 4: Generated Prediction Images, visualised using QGIS, parameters 'jsd prob ct 0.7' - corresponds to Figure 8, 9, and 16 in reproduced paper

https://doi.org/10.17605/OSF.IO/RSF4M

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

Initial execution of the analysis script raised errors while generating Fig. 4. This was however resolved after contacting the authors and Fig. 4 was reproduced. The generated image is shown in Figure 1.

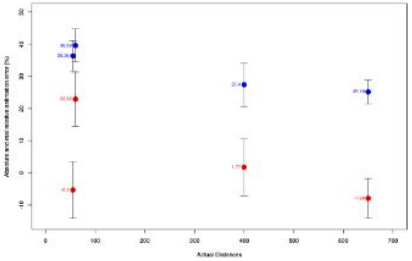


Figure 1: Average absolute (blue) and real relative (red) estimation error - corresponds to Figure 1 in reproduced paper

Table 1 and 2 were not generated automatically but instead the corresponding values where printed to the console. After extracting these values and exporting them to .csv files, both tables were reproduced. Refer to Table 1 and 2 in this report, which contain the raw values generated by the script.

Table 1: Spearman correlations Group 1 - corresponds to Table 1 in reproduced paper

type	S statistic	rho	p-value
SBSOD - Map Errors	463.24	-0.6197337	0.03160
SBSOD – Landmarks omitted	364.51	-0.2745132	0.38790
SBSOD – Road Segments mistakes	471.27	-0.6477894	0.02274
Landmarks omitted– Road Segments mistakes	160.86	0.4375473	0.15490
SBSOD – Direction estimates	233.45	0.1837559	0.56750
SBSOD – Distance estimates	342.00	-0.5545455	0.07665
Map Errors – Direction estimates	278.88	0.0249112	0.93870
Map Errors – Distance estimates	102.19	0.5354817	0.08958
Distance estimates – Direction estimates	205.81	0.0645223	0.85050

Table 2: Spearman correlations Group 2 - corresponds to Table 2 in reproduced paper

type	S statistic	rho	p-value
SBSOD - Map Errors	578.43	-0.5890887	0.03414
SBSOD – Landmarks omitted	520.67	-0.4304142	0.14210

This report is part of the reproduc
https://reproducible-agile.github.io
cite the report use

Friese, Philipp A. (2021, May
abilities in unfamiliar environ

Reviewed paper

Karkasina, D., Kokla, M., and
iar environments, AGILE GI
2021.

https://doi.org/10.17605/OSF.IO/DX92A

Reproducibility review reports AGILE 2021

Reproducibility review of: A Comparative Study of Typing and Speech For Map Metadata Creation

F.O. Ostermann and Daniel Nüst

Reproducibility reviewer notes

This review focuses on the reproduction of the analysis results. No in-depth examination
code was conducted, but it was confirmed that the provided code and be run and see
application used in the study. Using

```
# with npm version 6.14.8 and node version 14.13.0
npm install
npm start
```

we could run the application on http://localhost:8080, as shown in the screenshot bel

Figure 1: Screenshot of application executed locally

A clear license is missing in the repository. The most important information (soft
overview, exact questionnaire, maps used in the experiment) is also provided as supple

This report is part of the reproducibili
https://reproducible-agile.github.io/. T
cite the report use

Ostermann, F. O., & Nüst, D. (202
of Typing and Speech For Map M

Reviewed paper

Lai, P.-C. and Degbelo, A.: A Comparative Study of Typing and Speech For Map Metadata
Creation, AGILE GIScience Ser., 2, 7, https://doi.org/10.5194/agile-giss-2-7-2021, 2021.

Summary

The paper presents the results of a user experiment to improve GI-metadata using speech. A complete
reproduction is practically impossible to achieve. This reproducibility report therefore investigated two
components: First, whether sufficient information is provided to replicate the experiment elsewhere with

https://doi.org/10.17605/osf.io/7fqtm

Reproducibility review of: A Socially Aware Huff Model for Destination Choice in Nature-based Tourism

Jakub Krukar

2021-06-07

`Construction.ipynb` and can be verified by comparing the values from the output of chunk [12] to Table 8 in the paper. I expected that the values in the `Number of photos` column would stay the same (because the number of photos does not change), but the values in `outgoing/incoming trips` columns would change.

The location of the threshold variable is marked with an in-code comment in the file `Trip Construction.ipynb`:

```
if length.days > 4: #time threshold: average length of stay in both NPs
```

Table 8 Summary of attractions in Acadia National Park

Attraction	Number of photos	Outgoing trips	Incoming trips
Schoodic Institute	1119	53	64
Bass Harbor	2298	260	288
Southwest Harbor	723	109	111
Northeast Harbor	605	67	76
Bar Harbor	6259	433	357
Wild Gardens of Acadia	550	60	66
Cadillac Mountain	3285	349	345
Penobscot Peak	776	16	15
Bubble Rock	703	83	89
Jordan Pond	1250	227	250
Boulder Beach	536	85	102
Thunder Hole	977	167	185
Sand Beach	1253	216	177

Figure 1: Original Table 8 from the paper.

Places	a	b	c	d	e	f	g	h	i	j	k	l	m	total_out	total_in	cross_boundary	photos
Schoodic Institute	0	13	7	1	12	1	8	0	0	4	3	2	6	57	66	123	1119
Bass Harbor	12	0	34	9	64	13	53	4	8	25	12	15	21	268	295	563	2298
Southwest Harbor	3	44	0	6	30	3	15	4	1	4	1	2	2	115	117	232	723
Northeast Harbor	5	16	8	0	13	1	7	0	2	10	1	2	3	68	78	146	605
Bar Harbor	20	60	25	21	0	17	118	3	12	50	15	40	56	437	387	824	6259
Wild Gardens of Acadia	1	3	1	2	10	0	6	1	1	6	4	11	15	61	67	128	550
Cadillac Mountain	8	57	12	13	102	16	0	0	14	51	12	24	45	354	350	704	3285
Penobscot Peak	2	3	3	2	2	0	0	0	0	2	0	1	1	16	15	31	776

This report is part of the reproducil
https://reproducible-agile.github.io/.
the report use

Krukar, J. (2021, May 7). Reproducil
in Nature-based Tourism. https://dc

Reviewed paper

Shi, M., Janowicz, K., Cai, L., Mai, G
in Nature-based Tourism, AGILE G
2021.

Summary

The code, sample API query, and dov
working Binder link. All files containi

https://doi.org/10.17605/OSF.IO/4CPM3

Reproducibility review reports AGILE 2021

Reproducibility review of: Automated Extraction of Labels from Large-Scale Historical Maps

Daniel Nüst

2021-06-07

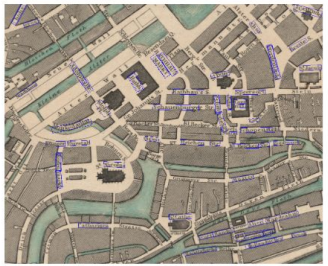


Figure 1: One output file from text recognition run.

I could run all cells in `String_Similarity_by_Levenshtein_Distance.ipynb`, and found the documentation to be extensive and a little bit raw, though very transparent, including tests by the author while developing the workflow etc. At first I got an error reading the `OCR_results.xlsx` file: `XLRDError: Excel xlsx file; not supported`, so I changed the data loading to use `openpyxl`.

Figure 6 seems to be created using QGIS, but no project file or georeferenced version of the base map was included in the repository.

I did not run the final notebook of step "06 Approximate georeferencing", because of the advertised run

This report is part of the reprodu
https://reproducible-agile.github.io
cite the report use

Nüst, D. (2021, May 6). Rep
Large-Scale Historical Maps.

Reviewed paper

Schlegel, I.: Automated Ext
GIScience Ser., 2, 12, https:/

Summary

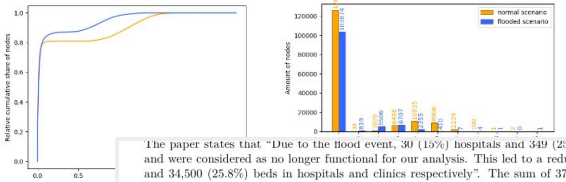
The provided workflow could be p
calculations could be executed and
providing all input and output data for each step, however, some manual steps were included which

Reproducibility review of: Flood Impact Assessment on Road Network and Healthcare Access – at the example of Jakarta, Indonesia

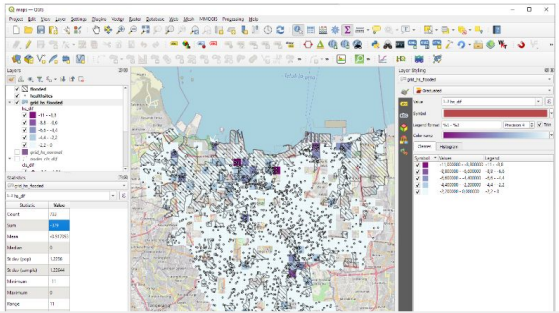
Anita Graser

2021-06-07

```
CDF results saved.
Histogram results saved.
```



The paper states that "Due to the flood event, 30 (15%) hospitals and 349 (25%) clinics were affected and were considered as no longer functional for our analysis. This led to a reduction of 12,000 (16.6%) and 34,500 (25.8%) beds in hospitals and clinics respectively". The sum of 379 affected heath service locations could be confirmed, as shown in the following screenshot:



This report is part
https://reproducible
cite the report use

Graser, A. (20
Network and H
17605/OSF.IO

The resulting CDF plots
Therefore, it is not strai

Reviewed paper

Klipper, I. G., Zipf, A., and Lautenbach, S.: Flood
Access at the example of Jakarta, Indonesia, AGI
giss-2-4-2021, 2021.

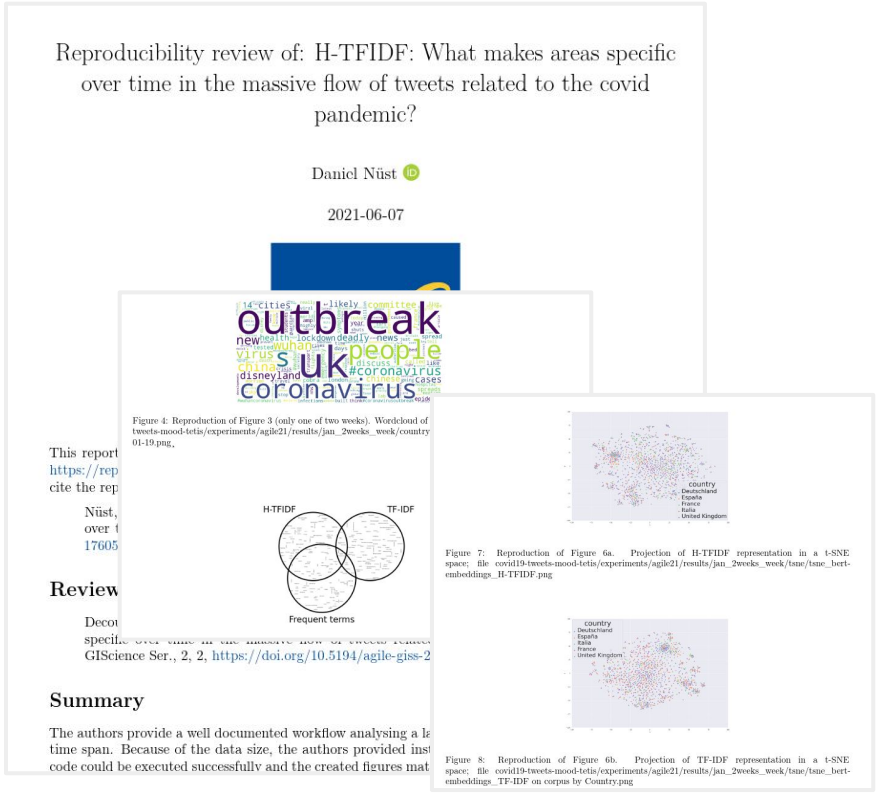
Summary

The provided workflow was **partially reproduc**
multiple sources (a Github.com repo, a GitLab r

https://doi.org/10.17605/osf.io/anv9r

https://doi.org/10.17605/OSF.IO/G4DCQ

Reproducibility review reports AGILE 2021



Reproducibility review of: H-TFIDF: What makes areas specific over time in the massive flow of tweets related to the covid pandemic?

Daniel Nüst

2021-06-07

Figure 4: Reproduction of Figure 3 (only one of two weeks). Wordcloud of tweets-mood-tetis/experiments/agile21/results/jan_2weeks_week/country 01-19.png.

Figure 7: Reproduction of Figure 6a. Projection of H-TFIDF representation in a t-SNE space; file covid19-tweets-mood-tetis/experiments/agile21/results/jan_2weeks_week/tsne/tsne_bert-embeddings_H-TFIDF.png

Figure 8: Reproduction of Figure 6b. Projection of TF-IDF representation in a t-SNE space; file covid19-tweets-mood-tetis/experiments/agile21/results/jan_2weeks_week/tsne/tsne_bert-embeddings_TF-IDF on corpus by Country.png

This report
https://rep
cite the rep

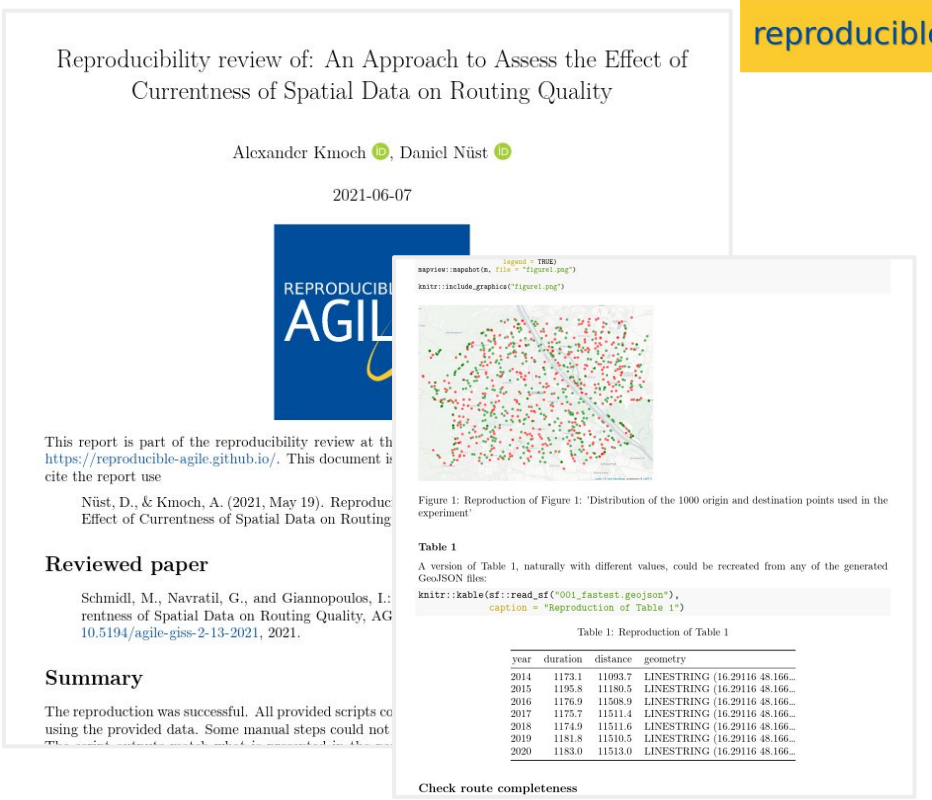
Nüst,
over t
17605

Review

Deco
speci
GIScience Ser., 2, 2, https://doi.org/10.5194/agile-giss-2

Summary

The authors provide a well documented workflow analysing a la
time span. Because of the data size, the authors provided inst
code could be executed successfully and the created figures mat



Reproducibility review of: An Approach to Assess the Effect of Currentness of Spatial Data on Routing Quality

Alexander Kmoch, Daniel Nüst

2021-06-07

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

Nüst, D., & Kmoch, A. (2021, May 19). Reproduc
Effect of Currentness of Spatial Data on Routing

Reviewed paper

Schmidl, M., Navratil, G., and Giannopoulos, I.:
rentness of Spatial Data on Routing Quality, AG
10.5194/agile-giss-2-13-2021, 2021.

Summary

The reproduction was successful. All provided scripts co
using the provided data. Some manual steps could not

```
legend = TRUE)
mapview::mapshot(m, file = "figure1.png")
knitr::include_graphics("figure1.png")
```


Figure 1: Reproduction of Figure 1: 'Distribution of the 1000 origin and destination points used in the experiment'

Table 1

A version of Table 1, naturally with different values, could be recreated from any of the generated GeoJSON files:

```
knitr::kable(sf::read_sf("001_fastest.geojson"),
             caption = "Reproduction of Table 1")
```

Table 1: Reproduction of Table 1

year	duration	distance	geometry
2014	1173.1	11093.7	LINESTRING (16.29116 48.166...
2015	1195.8	11180.5	LINESTRING (16.29116 48.166...
2016	1176.9	11508.9	LINESTRING (16.29116 48.166...
2017	1175.7	11511.4	LINESTRING (16.29116 48.166...
2018	1174.9	11511.6	LINESTRING (16.29116 48.166...
2019	1181.8	11510.5	LINESTRING (16.29116 48.166...
2020	1183.0	11513.0	LINESTRING (16.29116 48.166...

Check route completeness

https://doi.org/10.17605/rdnyu

https://doi.org/10.17605/bdu28

Reproducibility review reports AGILE 2021

Reproducibility review of: Extraction of linear structures from digital terrain models using deep learning

Daniel Nüst and Anita Graser

2021-06-07

This report is part of the reproducibility review at the AGILE conference. For more inf
https://reproducible-agile.github.io/. This document is published on OSF at https://osf.i
cite the report use

> Nüst, D., & Graser, A. (2021, April 30). Reproducibility review of: Extraction o
> structures from digital terrain models using deep learning. https://doi.org/10.17605
> 2sc7g

Reviewed paper

> Satari, R., Kazimi, B., and Sester, M.: Extraction of linear structures from digital
> models using deep learning, AGILE GIScience Ser., 2, 11, https://doi.org/10.5194/ag
> 2-11-2021, 2021.

Summary

The provided workflow was **partially reproduced**. Based on the provided test file and i
was able to recreate the computing environment and run the segmentation models. Relevan
the paper could be recreated. The training and validation part of the workflow is irreprodu

This finished within a minute! These values match the column `SegNet` of Table 1, within a level of precision to be expected from such a classification.

```
hrnet <- read.csv("agile-submission-2021/HRNetBinarySegmentation/files/evaluation_file.csv")
segnet <- read.csv("agile-submission-2021/SegnetBinarySegmentation//files/evaluation_file.csv")

suppressPackageStartupMessages(library("tidyverse"))
dplyr::full_join(hrnet, segnet) %>%
  knitr::kable()
```

acc	f1_m	loss	model_type	precision_m	recall_m	specific_name
0.9069678	0.7762662	0.2579104	hrnet	0.8187426	0.7393778	simple_binary
0.8787796	0.6977567	0.3094499	segnetCustomized	0.7688990	0.6399856	simple_binary

Run the next segmentation:

```
cd multiclassSegmentation
python3 evaluate.py --evaluation_file=evaluation_file.csv
```

This completes and recreates the data in Table 3 within reasonable numerical precision based on the file `multiClassEvaluation.csv`. It is unclear to me how Table 2 can be constructed from `evaluation_file.csv` of this segmentation, but I assume it can be.

```
multi <- read.csv("agile-submission-2021/multiclassSegmentation/files/multiClassEvaluation.csv")

rows <- lapply(c(0:5), function(class) {
  classValues <- multi %>%
    dplyr::select(dplyr::ends_with(as.character(class)))
  names(classValues) <- c("sparse_iou", "prediction", "recall", "f1.score", "support")
  c('Class label' = as.character(class), classValues)
})

dplyr::bind_rows(rows) %>%
  knitr::kable()
```

Class label	sparse_iou	prediction	recall	f1.score	support
0	0.8952831	0.9270560	0.9630006	0.9446866	12342971
1	0.2073642	0.4009994	0.3745212	0.3873083	247492
2	0.0324831	0.3139039	0.0373107	0.0666941	282037
3	0.7648531	0.8797419	0.8552982	0.8673479	1961849
4	0.4453632	0.6953347	0.5580455	0.6191711	830866
5	0.1593569	0.2697686	0.2787574	0.2741893	145345

https://doi.org/10.17605/osf.io/2sc7g

General observations and lessons learned

- **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

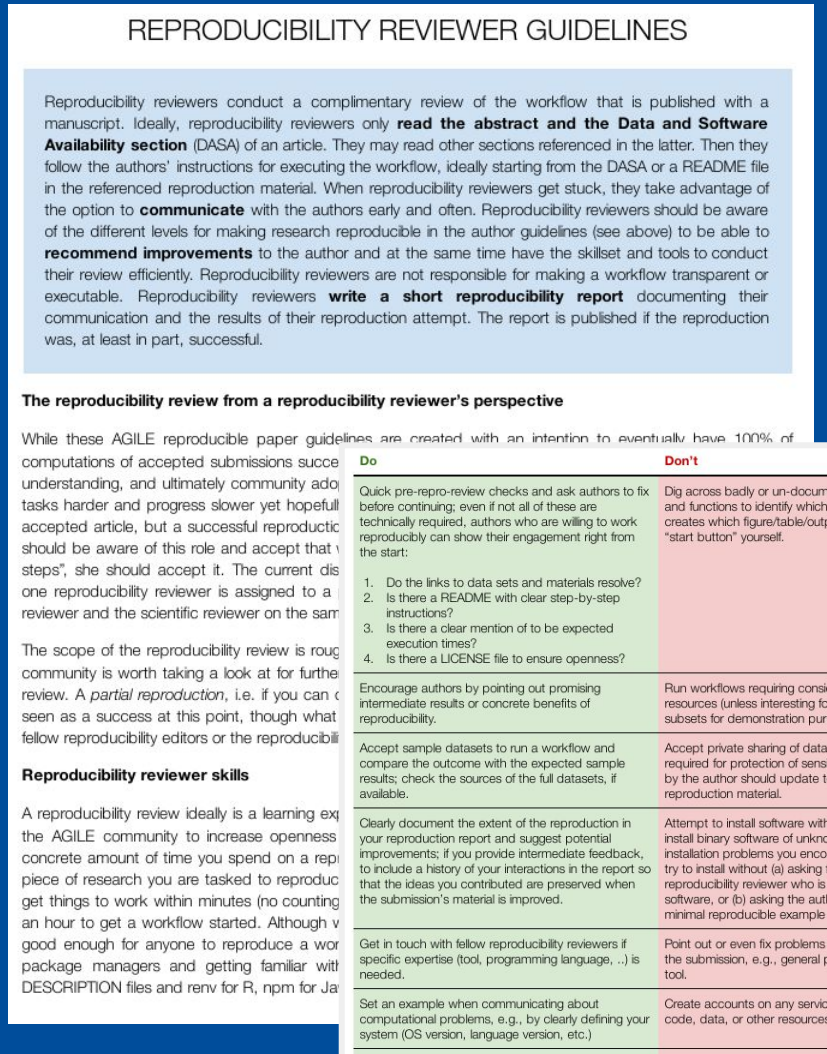
The guidelines for reproducibility reviewers

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!

Next steps

Do it again in 2022 🎉

🛠️ Revise guidelines? 🇮🇹 🇫🇷 🇨🇳

Grow reproducibility reviewer team

YOU!, opportunity ECRs (mentoring/workshops/…)

Continue meta-research 🕵️

Ostermann, F., Nüst, D., Granell, C., Hofer, B., & Konkol, M. (2020). *Reproducible Research and GIScience: an evaluation using GIScience conference papers*. EarthArXiv. https://doi.org/10.31223/x5zk5v

Continue community engagement towards opening scholarship

Scope
Requirements
Acceptance condition?

Open review if tenured
Format-free first submission
CRediT

Phase out when standard practice...

Thank you!

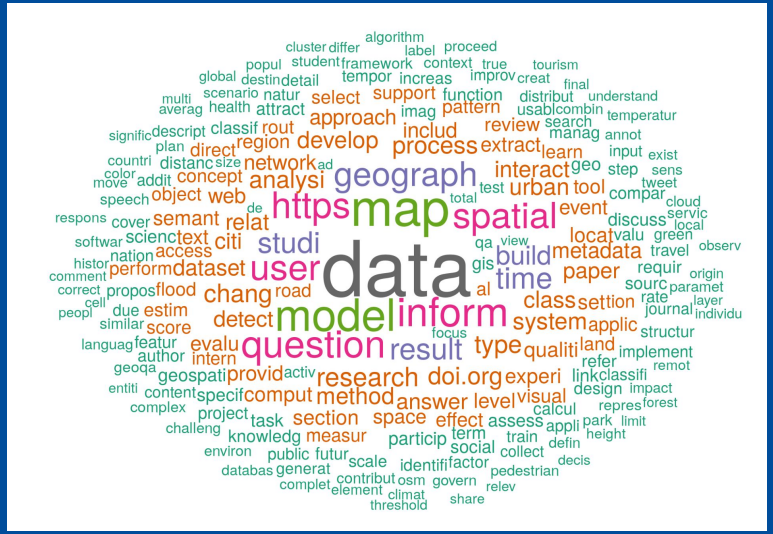
I look forward to your questions!
@nordhomen | d.n@wwu.de

Slides: **http://bit.ly/agile2021-repro-review**

Reproducibility Committee 2021
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)
Philipp Friese (Technical University of Munich, Germany)
Anita Graser (Austrian Institute of Technology, Austria)
Jakub Krukar (University of Münster, Germany)

REPRODUCIBLE AGILE

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



Word-stem cloud of all AGILE 2021 submissions (full/short/poster & accepted/rejected)