Reproducible Research is like riding a bike

Reproducibility is a fundamental pillar in science but it has recently been described as hard and challenging to achieve, as stated in numerous editorials and papers, some of which alert on a "reproducibility crisis". In this article we outline 1/ the approach taken to put Reproducible Research (RR) in the agenda of the GIScience community, 2/ first actions and initial lessons learned towards the discussion and adoption of RR principles and practices in the workflows and habits of researchers, and finally, we present 3/ our short-term strategy (two years) and specific actions to achieve the main goal of making RR an integral part of scientific workflows of the GIScience community.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27216v1 | CC BY 4.0 Open Access | rec: 16 Sep 2018, publ: 16 Sep 2018

Reproducible Research is like riding a bike

Carlos Granell[1], Daniel Nüst[2], Frank O. Ostermann[3], and Rusne Sileryte[4]

[1]**Institute of New Imaging Technologies, Universitat Jaume I de Castellón, Castellón, Spain**
[2]**Institute for Geoinformatics, University of Münster, Münster, Germany**
[3]**Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands**
[4]**Faculty of Architecture and the Built Environment, Delft University of Technology, Delft, The Netherlands**

Corresponding author:
Carlos Granell[1]

Email address: carlos.granell@uji.es

ABSTRACT

Reproducibility is a fundamental pillar in science but it has recently been described as hard and challenging to achieve, as stated in numerous editorials and papers, some of which alert on a "reproducibility crisis". In this article we outline 1/ the approach taken to put Reproducible Research (RR) in the agenda of the GIScience community, 2/ first actions and initial lessons learned towards the discussion and adoption of RR principles and practices in the workflows and habits of researchers, and finally, we present 3/ our short-term strategy (two years) and specific actions to achieve the main goal of making RR an integral part of scientific workflows of the GIScience community.

INTRODUCTION

Science Europe, an Brussels-based association of European Research Funding Organisations and Research Performing Organisations (https://www.scienceeurope.org/), recently announced the launch of "cOAlition S" (https://www.scienceeurope.org/coalition-s/) or simply "Plan S" , an initiative to make full and immediate Open Access to research publications a reality by January 2020 (Enserink, 2018). On the occasion of the launch of Plan S, Marc Schiltz, President of Science Europe, puts it into context and describes its key principles (Schiltz, 2018) remarking that "only results that can be discussed, challenged, and, where appropriate, tested, and reproduced by others qualify as scientific". Beyond of mere expressions of interest, we concur with the need for concrete initiatives and actions towards the support of Open Access in science like Plan S does. New scientific discoveries build on previously established scientific results. Schiltz (2018) continues, "science can therefore only function properly if research results are made openly available to the community so that they can be submitted to the test and scrutiny of other researchers", thereby acknowledging reproducibility as a fundamental principle in science.

Despite the wide recognition of reproducibility as a key principle in science by the scientific community, it is scarcely valued when it comes to practice. Some voices claimed that we are in an era of a "reproducibility crisis" that includes "hard" sciences, social sciences, and humanities. Take as example the media coverage produced in only one month (August 2018) with respect to several reproducibility failures of experiments and/or studies that have made headings as extensively reported on scientific media news and high-profile journals editorials (Kaiser, 2018; Grens, 2018b,a; Azvolonsky, 2018; Law, 2018). Other scientific disciplines such as humanities and social sciences are starting to look at reproducible research as they increasingly rely on the use of computer and computational analyses in their scientific work (Peels and Bouter, 2018). Once the computer becomes an indispensable part of a scientific project, the narrative of the "materials section" –i.e. data and methods– of a traditional scientific paper falls short to provide the required information to reproduce the results (Baker, 2016). Due to the increasing use of computational methods and analysis, the nascent fields of social sciences and humanities such as the digital humanities (https://eadh.org/), geohumanities (http://geohumanities.org/) and computational social sciences (Lazer et al., 2009) are also debating the idea of adopting computational reproducibility practices

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27216v1 | CC BY 4.0 Open Access | rec: 16 Sep 2018, publ: 16 Sep 2018

in their daily scientific work (Peels and Bouter, 2018). We argue that GIScientists should follow the path of computational reproducibility so that our scientific results can be "discussed, challenged, and, where appropriate, tested, and reproduced by others", as Schiltz (2018) argued.

Our ongoing initiative to promote Reproducible Research (RR) in the GIScience domain goes hand in hand with a series of workshops held in conjunction with the Association of Geographic Information Laboratories in Europe's (AGILE) annual conference series (see http://o2r.info/reproducible-agile/). We deliberately chose AGILE conference/association as a "starting point" of our activities, as a defining aspect of our community-driven approach to reproducibility that's explained below. Next, we report on the results and lessons learned during workshops held at two consecutive AGILE conferences and how they inform the upcoming initiatives to introduce RR in the AGILE community. Finally, based on the workshops' discussion and experience, we propose a short-term strategy with a set of actions to achieve the goal of making RR an integral part of the scientific workflows of the AGILE community, i.e. individual researchers and research laboratories. Mastering the production and reading of reproducible papers requires time. The benefit, though, is that newcomers to RR may become self-learners after first successes through guided materials in order to progressively adapt RR practices into their daily scientific work. It is like riding a bike: you can only learn it by trying and maybe falling down a few times. Without one's own experience to successfully ride the first few meters, one can read many books about riding but never manage to do it. However, once engrained in scientific practice, you never "unlearn" it. The community-driven approach and workshop series at AGILE conference for RR are designed to help researchers into the saddle.

COMMUNITY-DRIVEN, PROCESS-BASED APPROACH TO REPRODUCIBILITY

Even though individual researchers are vital, they are not the only actors in making reproducibility research a reality. Indeed, the challenge of establishing RR practices comprises a diversity of actors and stakeholders – funding agencies, research institutions, graduate study programs, publishers, journals, professional/academic associations, conferences, etc. – which all together are responsible for promoting and acknowledging RR practices in the long term.

Inspired by Tennant (2018), we begin with the idea that all these actors form a sort of complex interrelated ecosystem layered in hierarchical levels (Figure 1). Each level looks at RR from a different perspective, and interprets its own reality, needs and goals with respect to the adoption of RR into scientific workflows. While individual researchers worry about the lack of motivation and of supporting tools to facilitate reproducible research practices, journals look at it under a different angle as confirmed by a recent study (Vasilevsky et al., 2017) with biomedical journals in which "a significant association between higher Impact Factors and journals with a data sharing requirement" existed, for instance.

A top-down approach cannot successfully achieve a common consensus in the adoption of RR since it seems unrealistic to believe that "more than 10 million scientists, highly educated and intelligent, would agree with some rules created for them by a small number of people" (Tennant, 2018). While some policy aspects or norms of the scientific endeavour may be proposed by funding agencies and bodies following a top-down approach, like Plan S (Schiltz, 2018), these norms however are normally preceded by an intensive consultation process among the involved parties, and the nature of these norms mainly affect administrative tasks or procedures of the scientific process rather than the scientist process itself.

At the other end of the spectrum, a crowdsourced, bottom-up approach to consolidate RR practices is also quite unlikely to succeed given the varied perspectives and practices of the research community about RR. There exist many geographical, disciplinary and actor differences and subtleties which altogether make it impractical to reach a consensus framework for RR practices that would fit all scientific domains, fields and disciplines.

As an alternative to the top-down versus bottom-up dichotomy, our approach to leveraging RR practices in the GIScience community is to propose motivating incentives at the community level which may "drag" other actors (individuals, groups, etc.) in the adoption of RR practices (Figure). In Nüst et al. (2018), we *partly* explored this strategy considering the AGILE conference/association as an actor who can provide strong incentives for RR practices at the community level. As a scholarly association, AGILE's activities may be quickly acquired by directly related actors such as AGILE member labs, individual researchers, as well as institutional-level actors like publishers and related journals. Rather than a linear interaction among the involved actors, the proposed community-based approach favours

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27216v1 | CC BY 4.0 Open Access | rec: 16 Sep 2018, publ: 16 Sep 2018

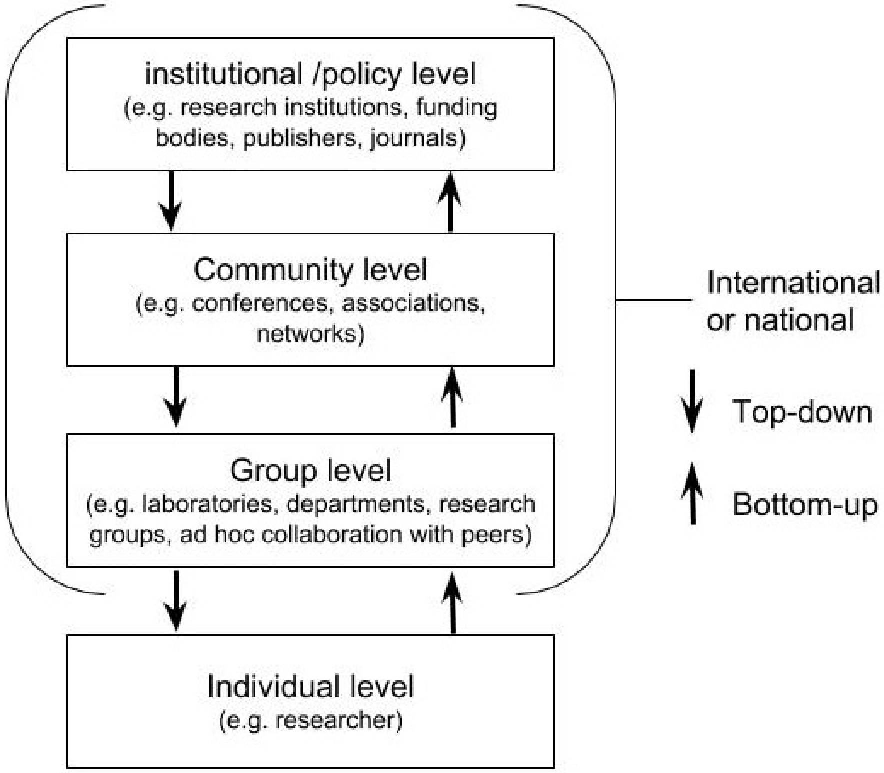

Figure 1. Actors ecosystem layered in hierarchical levels.

networked interaction as the AGILE conference/association acts as a community-level stimulus (like a network hub) to influence research groups and individual researchers to incorporate RR practices in their scientific workflows.

We earlier said that we "partly explored this strategy" because the focus was only on the community aspect. To be successful, RR must also be an intrinsic part of the scientific workflow and practices. That is, RR must be understood as a dynamic, evolving process by which researchers are continuously adapting, consolidating and improving methods and techniques to make their research reproducible, as well as reflecting on the decisions made in each reproducible project, in line with the proposal of Shannon and Walker (2018) to open GIScience. A checklist (bicycle, helmet, etc.) is useful, but it will never allow one to learn to ride a bicycle. Only through a trial-and-error process one can learn it. The teacher is the experience. Similarly, checklists for RR are useful, but researchers must rely on a process-based approach to acquire, engage and reflect on RR practices.

While the process-based approach to RR has been less explored, we recognised its importance from the outset (see next section) and it is a milestone in our future roadmap. Next, we overview the first actions and initial lessons learned towards the discussion and adoption of RR principles by the GIScience domain, putting the emphasis on the case of the (community-level) AGILE conference/association.

WORKSHOPS AS A MEAN TO SPREAD KNOWLEDGE

The community-driven approach is necessarily born out of and grows around a community event like the AGILE conference. Our ongoing initiative to promote RR revolves around a series of workshops (http://o2r.info/reproducible-agile/) held in conjunction with the AGILE conferences (https://agile-online.org/). The first workshop "Reproducible Geosciences Discussion Forum" (https://o2r.info/reproducible-agile/2017) was held at AGILE 2017 conference in Wageningen, The Netherlands. Outcomes of the first workshop included a broad discussion about RR in the geospatial domain, and the consolidation of a

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27216v1 | CC BY 4.0 Open Access | rec: 16 Sep 2018, publ: 16 Sep 2018

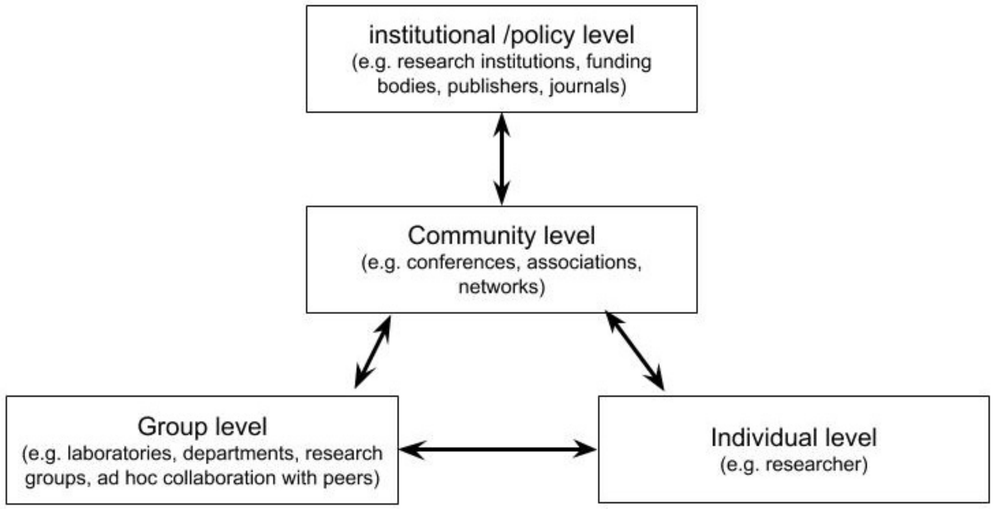

Figure 2. Actors ecosystem *dragged* by a community-level hub.

collaborative effort (Nüst et al., 2018) to analyse nominees for best short and full papers of past AGILE conferences (2008-2017). Data, computational analysis, and results of the analysis and visualisations are publicly available as a research compendium (Nüst, 2018).

The analysis conducted in Nüst et al. (2018) discerned the level of "preproducibility" of AGILE papers, understanding a preproducibible scientific paper as one that "has been described in adequate details for others to undertake it" (Stark, 2018). In our analysis, we did not reproduce the papers but assess their level of preproducibility in terms of analysing whether or not a paper provided sufficient details according to a set of criteria to enable its reproduction. The authors of the analyzed AGILE papers were asked to fill in a survey to comment on the proposed level of preproducibility of their papers and to give their opinion and suggestions to improve reproducibility (full details in Nüst et al. (2018)). With respect to the latter, survey respondents suggested that they were generally aware of the need for reproducibility in their paper and that they *knew how to improve reproducibility in their work*. However, many did not consider it a priority due to the lack of motivation or the additional effort required to do so, which was disproportionately large compared to the added value; such an argument is a recurring theme widely studied and evaluated in the literature (Tenopir et al., 2011, 2015; Thursby et al., 2018). Here, we explicitly mention this statement because "the lack of perceived motivation" reinforces the need for the proposed community-driven approach to focus on a well identified incentive (AGILE conference publication) for (regular AGILE) researchers, and research laboratories that can strongly motivate them to incorporate reproducibility practices in their work (submitted to AGILE conference). With respect to the "[authors] knew how to improve reproducibility in their work", we do not doubt about the respondents' claim, but our overall perception is that reproducibility is still seen as set of specific characteristics and "additional" tasks –e.g. make code available, publish data, etc.–, like a checklist for paper submission, rather than an intrinsic, evolving process at the core of the scientific method from the outset of a research project. That is why we put the emphasis on the process to consolidate and strengthen practices and ways of working for RR.

The second workshop "Reproducible Research Publication" (https://o2r.info/reproducible-agile/2018) took place during AGILE 2018 conference in Lund, Sweden. It focussed on a hands-on session to understand the technical challenges encountered while reproducing a research paper. We elicited feedback from the workshop participants to identify their experiences and needs regarding the reproduction process, complemented with an on-line survey to all registered participants sent one week after the AGILE workshop (4 respondents). Hands-on experience on reproducing the computational analysis of a published paper (i.e. Nüst (2018)) was mentioned as the most useful part of the workshop, although it was not without flaws. The participants would have preferred to reproduce a paper in a programming language

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27216v1 | CC BY 4.0 Open Access | rec: 16 Sep 2018, publ: 16 Sep 2018

they were already familiar with. Due to the lack of experience with R, some found the example of computational analysis too complicated. Indeed, the biggest obstacle faced by most workshop participants was to deal with the missing dependencies during the reproduction exercise, which is consistent with the main problems identified by Konkol et al. (2018), where a major technical problem was generated by calling a library that was not installed in the reproduction environment. It is noteworthy almost all participants were eventually successful.

SHORT-TERM STRATEGY AND ACTIONS

In this section we outline our short-term strategy (two years) and specific actions towards the introduction of RR practices within the scientific workflows of the GIScience community. First, we present ongoing action, followed by a discussion of foreseen actions.

With respect to ongoing actions, we concentrate on two. The first one is the planning of the third workshop at AGILE as part of the RR workshops series. For the third edition, participants of the second workshop would prefer to obtain more information on the possible concepts of reproducibility (e.g. data *versus* methods, open *versus* non-open, repeatability *versus* reproducibility, processing *versus* interpretation) and be introduced/tasked with less complicated examples of reproducible papers. Besides, a review and discussion on possible tools and their pros and cons (open/free *versus* proprietary/commercial, community-driven *versus* company-driven, local/institutional *versus* global providers, etc.) has been mentioned as a suggestion for the upcoming workshops, like running a sort of Carpentry workshops at conferences.

The second ongoing action is the submission of an AGILE initiative proposal (https://agile-online.org/funding-initiatives), which is under evaluation by the AGILE Council, to make next editions of the AGILE conference more reproducible by updating the Call for Papers. The updated call would provide clear and concrete guidelines about how to submit and review reproducible papers. If approved, the work to be done will be mostly based on the set of recommendations and suggestions for the AGILE conference/association described in Nüst et al. (2018). Again, by following a community-driven stimulus, we expect a snowball effect that may lead to a change in practises in the actors within the AGILE community, and even influence other community-level actors such as sister conferences (e.g. OGRS) and associations.

Looking at the future, among the next actions reproducibility in teaching is a priority. The last workshop discussion showed that the detailed manual reproduction was an important learning experience when starting with reproducible research. Indeed, RR guidelines or "recipes" may be viewed as initial seeds for designing and creating open educational resources and materials to help early-stage researchers and established researchers alike understand the main concepts of RR and Open Science and apply them in geospatial research. However, understanding the challenges and pitfalls a reader might have is an important prerequisite and motivation for changing one's own habits (Nüst et al., 2018). Like learning to ride a bicycle, an evolving trial-and-error process is fundamental to understand and overcome the barriers for open and reproducible research (cf. (Konkol et al., 2018)).

Workshop participants also expressed interest in keeping informed about the future activities and the process made by the workshop organisers (authors of the paper) to promote reproducibility. Most would even like to contribute to the effort of revision of materials (e.g. teaching materials) or active participation in their preparation. A key takeaway message from our experiences during the workshops was the participants' perception about the importance of educational resources and teaching materials for reproducible research. Remarkable educational resources for researchers are available on-line or under development: Open Science MOOC https://opensciencemooc.github.io/site/; Digital open science MOOC (Toelch and Ostwald, 2018); The Carpentries initiative (https://carpentries.org/), which is a renowned example by the way teaching materials (lessons) are created, taught, and delivered via on-site workshops, such as the Reproducible Research in R Workshop Overview (https://datacarpentry.org/rr-workshop/) or the Geospatial Data Workshop (https://datacarpentry.org/lessons); and the Teaching Tech Together, http://teachtogether.tech/en/, which is related to the Software Carpentry instructor training program.

Nevertheless, the development of educational materials for reproducible research still faces open questions pertinent for example to the format, content and way of delivering them: How are research challenges related to teaching methods? Is there any research about the relation between skills of teachers and students? How many teachers at uni know/practice reproducibility? Is RR relevant for BSc or only for MSc? What are the challenges of teaching technology *versus* teaching set-up at universities (e.g.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27216v1 | CC BY 4.0 Open Access | rec: 16 Sep 2018, publ: 16 Sep 2018

Carpentries require multiple instructors and small classes, etc.)? The controversy continues and requires a broader discussion among the involved actors.

FINAL THOUGHTS

Mastering the production and reading of reproducible papers requires time. Reader, do not get us wrong: reproducibility research can be time-consuming and is becoming incredibly complex because neither software nor data are static (Perkel, 2018). In addition, there are limits to the level of detail a scholarly article can provide due to technical restrictions and privacy concerns, so not "everyone" may be able to reproduce a paper, nor it may be fully reproduced only based on the article text itself. It should be noted that distinct levels of reproduction (Peng, 2011) are perfectly fine beyond of a binary black-and-white classification. Equally important are the potential interactions a reproducible article allows (reproduce a chart, apply method to different data, conduct a review) which may differ between readers and thus require different techniques and practices for each individual.

Unfortunately, there are no "common" guidelines and "typical" research projects. RR requires an evolving process that changes as the nature and characteristics of software ,data, and research questions do. Author, try to fit reproducible research practices in your regular scientific workflow, adapt them and incorporate more as you get more confident in your reproducible scientific work. As recently said in an editorial, code and "data sharing is not only a way to improve the reproducibility and robustness of the science that is taking today, but can drive new science for tomorrow" (Editorial, 2018).

ACKNOWLEDGMENTS

We thank both all participants of the two AGILE workshops (2017, 2018) and respondents to the survey for their detailed comments and valuable suggestions.

REFERENCES

Azvolonsky, A. (2018). Studies unable to reproduce results of two diabetes papers. https://www.the-scientist.com/news-opinion/studies-unable-to-reproduce-results-of-two-diabetes-papers-64630.

Baker, M. (2016). Why scientists must share their research code. *Nature*.

Editorial (2018). Data sharing and the future of science. *Nature Communications*, 9(1).

Enserink, M. (2018). European science funders ban grantees from publishing in paywalled journals. *Science*.

Grens, K. (2018a). Latest effort to replicate psych studies yields 62 percent success. https://www.the-scientist.com/news-opinion/latest-effort-to-replicate-psych-studies-yields-62-percent-success-64713.

Grens, K. (2018b). Latest reproducibility project study fails to replicate. https://www.the-scientist.com/the-nutshell/latest-reproducibility-project-study-fails-to-replicate-29961.

Kaiser, J. (2018). Plan to replicate 50 high-impact cancer papers shrinks to just 18. *Science*.

Konkol, M., Kray, C., and Pfeiffer, M. (2018). Computational reproducibility in geoscientific papers: Insights from a series of studies with geoscientists and a reproduction study. *International Journal of Geographical Information Science*, page 1–22.

Law, Y.-H. (2018). Replication failures highlight biases in ecology and evolution science. https://www.the-scientist.com/features/replication-failures-highlight-biases-in-ecology-and-evolution-science-64475.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., and et al. (2009). Social science: Computational social science. *Science*, 323(5915):721–723.

Nüst, D., Boettiger, C., and Marwick, B. (2018). How to Read a Research Compendium. *ArXiv e-prints 1806.09525*.

Nüst, D. (2018). Reproducibility package for "reproducible research and giscience: An evaluation using agile conference papers". https://zenodo.org/record/1227260.

Nüst, D., Granell, C., Hofer, B., Konkol, M., Ostermann, F. O., Sileryte, R., and Cerutti, V. (2018). Reproducible research and giscience: an evaluation using agile conference papers. *PeerJ*, 6:e5072.

Peels, R. and Bouter, L. (2018). The possibility and desirability of replication in the humanities. *Palgrave Communications*, 4(1).

Peng, R. D. (2011). Reproducible research in computational science. *Science*, 334(6060):1226–1227.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27216v1 | CC BY 4.0 Open Access | rec: 16 Sep 2018, publ: 16 Sep 2018

Perkel, J. M. (2018). A toolkit for data transparency takes shape. *Nature*, 560(7719):513–515.
Schiltz, M. (2018). Science without publication paywalls: coalition s for the realisation of full and immediate open access. *Frontiers in Neuroscience*, 12.
Shannon, J. and Walker, K. (2018). Opening giscience: A process-based approach. *International Journal of Geographical Information Science*, 32(10):1911–1926.
Stark, P. B. (2018). Before reproducibility must come preproducibility. *Nature*, 557(7707):613–613.
Tennant, J. (2018). Foundations for open scholarship strategy development: First formal release.
Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M., and Frame, M. (2011). Data sharing by scientists: Practices and perceptions. *PLoS ONE*, 6(6):e21101.
Tenopir, C., Dalton, E. D., Allard, S., Frame, M., Pjesivac, I., Birch, B., Pollock, D., and Dorsett, K. (2015). Changes in data sharing and data reuse practices and perceptions among scientists worldwide. *PLOS ONE*, 10(8):e0134826.
Thursby, J. G., Haeussler, C., Thursby, M. C., and Jiang, L. (2018). Prepublication disclosure of scientific results: Norms, competition, and commercial orientation. *Science Advances*, 4(5):eaar2133.
Toelch, U. and Ostwald, D. (2018). Digital open science—teaching digital tools for reproducible and transparent research. *PLOS Biology*, 16(7):e2006022.
Vasilevsky, N. A., Minnier, J., Haendel, M. A., and Champieux, R. E. (2017). Reproducible and reusable research: are journal data sharing policies meeting the mark? *PeerJ*, 5:e3208.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27216v1 | CC BY 4.0 Open Access | rec: 16 Sep 2018, publ: 16 Sep 2018