

REPRODUCIBILITY

Enhancing reproducibility for computational methods

Data, code, and workflows should be available and cited

By Victoria Stodden,¹ Marcia McNutt,² David H. Bailey,³ Ewa Deelman,⁴ Yolanda Gil,⁴ Brooks Hanson,⁵ Michael A. Heroux,⁶ John P.A. Ioannidis,⁷ Michela Tauber⁸

Over the past two decades, computational methods have radically changed the ability of researchers from all areas of scholarship to process and analyze data and to simulate complex systems. But with these advances come challenges that are contributing to broader concerns over irreproducibility in the scholarly literature, among them the lack of transparency in disclosure of computational methods. Current reporting methods are often uneven, incomplete, and still evolving. We present a novel set of Reproducibility Enhancement Principles (REP) targeting disclosure challenges involving computation. These recommendations, which build upon more general proposals from the Transparency and Openness Promotion (TOP) guidelines (1) and recommendations for field data (2), emerged from workshop discussions among funding agencies, publishers and journal editors, industry participants, and researchers representing a broad range of domains. Although some of these actions may be aspirational, we believe it is important to recognize and move toward ameliorating irreproducibility in computational research.

Access to the computational steps taken to process data and generate findings is as important as access to data themselves. Computational steps can include information that details the treatment of outliers and missing values or gives the full set of model parameters used. Unfortunately, reporting of and access to such information is not routine in the scholarly literature (3). Although independent reimplementations of an experiment can provide important scientific evidence regarding a discovery and is a practice we wish to encourage, access to the underlying software and data is key

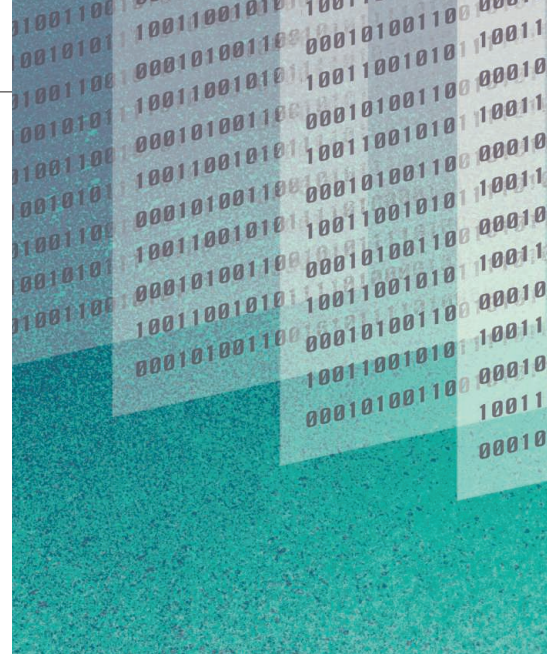
to understanding how computational results were derived and to reconciling any differences that might arise between independent replications (4). We thus focus on the ability to rerun the same computational steps on the same data the original authors used as a minimum dissemination standard (5, 6), which includes workflow information that explains what raw data and intermediate results are input to which computations (7). Access to the data and code that underlie discoveries can also enable downstream scientific contributions, such as meta-analyses, reuse, and other efforts that include results from multiple studies.

RECOMMENDATIONS

Share data, software, workflows, and details of the computational environment that generate published findings in open trusted repositories. The minimal components that enable independent regeneration of computational results are the data, the computational steps that produced the findings, and the workflow describing how to generate the results using the data and code, including parameter settings, random number seeds, make files, or function invocation sequences (8, 9).

Often the only clean path to the results is presented in a publication, even though many paths may have been explored. To minimize potential bias in reporting, we recommend that negative results and the relevant spectrum of explored paths be reported. This places results in better context, provides a sense of potential multiple comparisons in the analyses, and saves time and effort for other researchers who might otherwise explore already traversed, unfruitful paths.

Persistent links should appear in the published article and include a permanent identifier for data, code, and digital artifacts upon which the results depend. Data and code underlying discoveries must be discoverable from the related publication, accessible, and reusable. A unique identifier should be assigned for each artifact by the article publisher or repository. We recommend digital object identifiers (DOIs) so that it is possible to discover related data sets and code through the DOI structure itself, for example, using a hierarchical schema. We advocate sharing digital scholarly objects in open trusted repositories that are crawled by search engines.



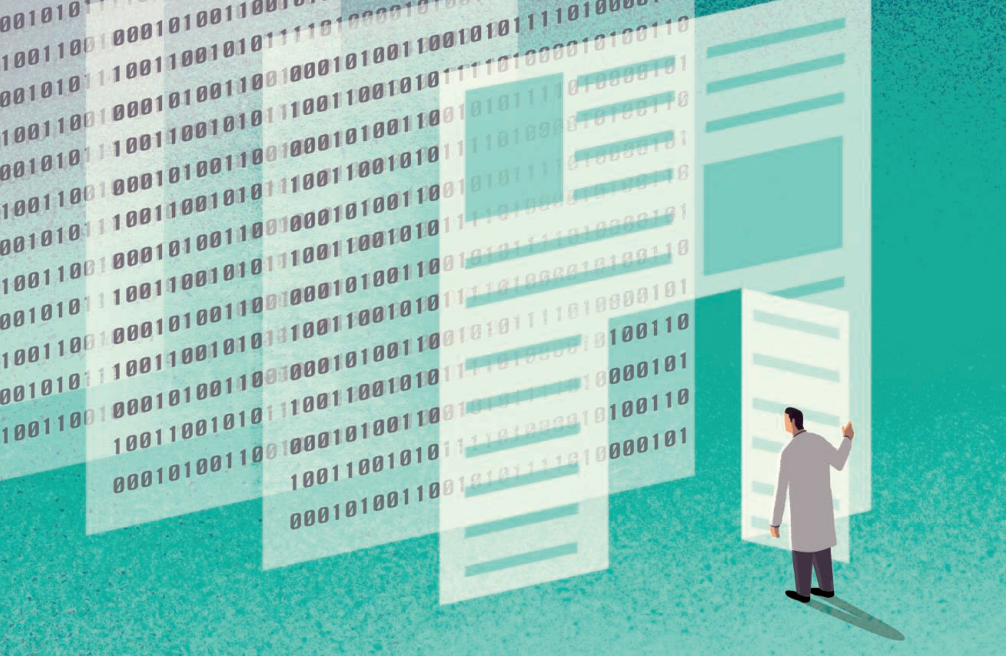
Sufficient metadata should be provided for someone in the field to use the shared digital scholarly objects without resorting to contacting the original authors (i.e., <http://bit.ly/2fVwjPH>). Software metadata should include, at a minimum, the title, authors, version, language, license, Uniform Resource Identifier/DOI, software description (including purpose, inputs, outputs, dependencies), and execution requirements.

To enable credit for shared digital scholarly objects, citation should be standard practice. All data, code, and workflows, including software written by the authors, should be cited in the references section (10). We suggest that software citation include software version information and its unique identifier in addition to other common aspects of citation.

To facilitate reuse, adequately document digital scholarly artifacts. Software and data should include adequate levels of documentation to enable independent reuse by someone skilled in the field. Best practice suggests that software include a test suite that exercises the functionality of the software (10).

Use Open Licensing when publishing digital scholarly objects. Intellectual property laws typically require permission from the authors for artifact reuse or reproduction. As author-generated code and workflows fall under copyright, and data may as well, we recommend using the Reproducible Research Standard (RRS) to maximize utility to the community and to enable verification of findings (11). The RRS recommends attribution-only licensing, e.g., the MIT License or the modified Berkeley Software Distribution (BSD) License for software and workflows; the Creative Commons Attribution (CC-BY) license for media; and public domain dedication for data. The RRS and principles of open licensing should be clearly explained to authors by journals, to ensure long-term open access to digital scholarly artifacts.

¹University of Illinois at Urbana-Champaign, Champaign, IL 61801, USA. ²National Academy of Sciences, Washington, DC 20418, USA. ³University of California, Davis, CA 95616, USA. ⁴University of Southern California, Los Angeles, CA 90007, USA. ⁵American Geophysical Union, Washington, DC 20009, USA. ⁶Sandia National Laboratories, Avon, MN 56310, USA. ⁷Stanford University, Stanford, CA 94305, USA. ⁸University of Delaware, Newark, DE 19716, USA. Email: vcs@stodden.net



Journals should conduct a reproducibility check as part of the publication process and should enact the TOP standards at level 2 or 3. Such a check asks whether the data, code, and computational steps upon which findings depend are available in an open trusted repository in a discoverable and persistent way, with links provided in the publication. And have all digital artifacts been openly licensed? Is documentation and workflow information available for a reader to follow the discovery process? Are all digital scholarly objects used in the discovery process cited in the manuscript's reference section? Could the published computational findings be reproduced on an independent system by using the data and code provided?

The last item is arguably the most time-consuming for reviewers and difficult to carry out, and many journals may choose not to adopt it or may perform partial reproduction for only some of the computational findings. The journal article should specify which of these items have been checked and, if so, whether they are fully or partially fulfilled.

Journals should strive to enact level 2 or 3 of the TOP standards on "Data transparency" and "Analytic methods (code) transparency." Level 3 recommends an independent reproduction of findings. Some journals are already taking steps in this direction (12, 13).

To better enable reproducibility across the scientific enterprise, funding agencies should instigate new research programs and pilot studies. Resolving some barriers to reproducibility may be straightforward; however, others may take time and community effort to overcome. We recommend enacting research programs to advance our understanding of reproducibility in computationally enabled research. Topics might include methods for verifying queries on confidential data; extending validation, verification, and uncertainty quantification to encompass reproducibility;

numerical reproducibility and sensitivity to small variations in computation (14); testing standards for code, including closed or proprietary codes; cyberinfrastructure that supports reproducibility, as well as innovative computational work; pilot efforts to create "instruction manuals" for manuscript submission (e.g., <http://libguides.caltech.edu/authorcarpentry>); policy research on intellectual property law and software patenting; costs and benefits to reproducibility in different settings, for example, in industry collaboration; provenance and workflow repositories; and exploring how to make investments regarding the preservation of various digital artifacts. Funding bodies could support efforts to reproduce results in different computational settings to better understand sources of error in computational findings.

BARRIERS, EXCEPTIONS, ONGOING EFFORTS

We recognize that there are challenges to the implementation of these recommendations. There will necessarily be exceptions in the near term and possibly indefinitely, for example, analysis and data involving human subjects or proprietary codes. However, we believe that creative ways to manage exceptions could be developed in such cases and that exceptions should be explained in the article. For example, if data or code cannot be made publicly accessible, the research team or journals could have infrastructure, policies, and procedures in place for rapidly giving reviewers access to information necessary to perform a review (13, 15).

It may not be possible to fully disclose, or even license, all proprietary software used in the discovery pipeline. However, scripts designed to be executed by propriety software such as MATLAB may be openly licensed by the script authors under the RRS. We also feel there are broad benefits to code release, for example, allowing for inspection, even if

the code cannot be executed (16).

Beyond the reproducibility check described above, journals can improve review of computational findings by rewarding reviewers who take extra effort to verify computational findings. Authors that facilitate such a review could be rewarded with badging of their published article (e.g., <http://bit.ly/Badging2gP>). Best practices for reviewers of reproducible publications need to be formulated. Funding agencies may encourage, request, and reward reproducible research practices in the scientific investigations that they review and fund.

Appropriate methodology to facilitate reproducibility should be taught to students who will use computational techniques in research. Best practices of digital scholarship should be required and incorporated into curricula and should include discussions of ethics, use of repositories, and version control, for example. Key societies or communities should consider short courses, best practices publications, and awards to promote these skills. Groups or research areas with limited experience in reproducible research practices could focus initially on a few seminal articles to demonstrate and promote reproducibility.

We believe that as these efforts become commonplace, practices and tools will continue to emerge that reduce the amount of time and resource investment necessary to facilitate reproducibility and support increasingly ambitious computational research. ■

REFERENCES AND NOTES

1. B. A. Nosek et al., *Science* **348**, 1422 (2015).
2. M. McNutt et al., *Science* **351**, 1024 (2016).
3. A. A. Alsheikh-Ali et al., *PLOS ONE* **6**, e24357 (2011).
4. D. Donoho et al., *IEEE Comput. Sci. Eng.*, **11**, 8 (2009).
5. V. Stodden, *JMS Bull. Online*, 17 November (2013); <http://bit.ly/BullJMSStat2013>.
6. D. H. Bailey, J. M. Borwein, V. Stodden, *Notices Amer. Math. Soc.* **60** (6), 679 (2013).
7. D. Garijo et al., *PLOS ONE* **8**, e80278 (2013).
8. D. Donoho, V. Stodden, in *The Princeton Companion to Applied Mathematics*, N. J. Higham, Ed. (Princeton Univ. Press, Princeton, NJ, 2016), pp. 916–925.
9. R. Gentleman, D. Temple Lang, *J. Comput. Graph. Stat.* **16**, 1 (2007).
10. V. Stodden, S. Miguez, *J. Open Res. Softw.* **2**, e21 (2014).
11. V. Stodden, *Comput. Sci. Eng.* **11**, 35 (2009).
12. V. Stodden, P. Guo, Z. Ma, *PLOS ONE* **8**, e67111 (2013).
13. M. Heroux, *ACM Trans. Math. Softw.* **41**(3), art13 (2015).
14. D. H. Bailey, J. M. Borwein, V. Stodden, in *Reproducibility: Principles, Problems, Practices*, H. Atmanspacher and S. Maasen, Eds. (Wiley, New York, 2015), pp. 205–232.
15. M. Fuentes, *AMSTAT News*, July 2016; <http://bit.ly/JASA2g>.
16. R. J. LeVeque, *SIAM News* **46**, April 2013.

ACKNOWLEDGMENTS

These recommendations emerged from a workshop held at the American Association for the Advancement of Science (AAAS), Washington, DC, 16 and 17 February 2016, funded by the Laura and John Arnold Foundation (<http://bit.ly/AAAS2016Arnold>). Workshop participants are identified in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/354/6317/1240/suppl/DC1

10.1126/science.aah6168

Enhancing reproducibility for computational methods

Victoria Stodden, Marcia McNutt, David H. Bailey, Ewa Deelman, Yolanda Gil, Brooks Hanson, Michael A. Heroux, John P.A. Ioannidis and Michela Taufer

Science **354** (6317), 1240-1241.
DOI: 10.1126/science.aah6168

ARTICLE TOOLS

<http://science.sciencemag.org/content/354/6317/1240>

SUPPLEMENTARY MATERIALS

<http://science.sciencemag.org/content/suppl/2016/12/07/354.6317.1240.DC1>

RELATED CONTENT

<http://science.sciencemag.org/content/sci/355/6323/357.3.full>

REFERENCES

This article cites 11 articles, 2 of which you can access for free
<http://science.sciencemag.org/content/354/6317/1240#BIBL>

PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2016, American Association for the Advancement of Science