

Best Practices for Researchers Publishing Computational Results

From StoddenWiki

This guide is meant both as a living, evolving, and changing document, and as a generic approach with the goal of being as broadly useful as possible to computational scientists wishing to practice reproducible research. Many fields have specialized and accepted ways of implementing the principles embodied in these best practices. I will elucidate these when I can, providing more detail over time. In the meantime, this is meant to provide a rough guide to producing and disseminating really reproducible computational research.

Suggestions and collaborations are welcome. Just create an account before editing this wiki and/or direct any email comments to me at BestPracticesATstodden.net.

A fixed citable version of this document is available [here](#). If you have found this information useful in your own practice of reproducible research, please cite the document: V. Stodden, "Best Practices for Researchers When Publishing Computational Results," SSRN, Sept 1, 2012.

Contents

- 1 Principles and Best Practices
 - 1.1 Best Practices at Publication
 - 1.2 Pre-publication Best Practices
 - 1.3 Influences from Sources External to the Research Process
- 2 References and further reading
- 3 Open Questions and Challenging Cases

Principles and Best Practices

In this document we envision reproducibility as a digital concept, to distinguish it from reproducibility from first principles (Notions of reproducibility), or repeating an experiment from scratch. Our goal is to facilitate the verification of computational results from the digital data files and the computer instructions required to generate the tables and figures in published articles.

This document assumes you have the right to make the data and code publicly available. Usually this means you either generated the data and code yourself, or have permission from others who did so.

Best Practices at Publication

1. **Data must be available and accessible.** In this context the term "data" means the raw data files used as a basis for the computations, that are necessary for others to regenerate published computational findings.

2. **Code and methods must be available and accessible.** The traditional methods section in a typical publication does not communicate sufficient detail for a knowledgeable reader to replicate computational results. A necessary action is making the complete set of instructions, typically in the form of computer scripts or workflow pipelines, conveniently available.
3. **Citation.** Do it. If you use data you did not collect from scratch, or code you did not write, however little, cite it. Citation standards for code and data are discussed but it is less important to get the citation perfect than it is to make sure the work is cited at all.
4. **Copyright and Publisher Agreements.** Publishers, almost uniformly, request that authors transfer all ownership rights over the article to them. All they really need is the authors' permission to publish.
5. **Supplemental materials.** Publishers should establish style guides for supplemental sections, and authors should organize their supplemental materials following best practices.

Pre-publication Best Practices

- Data management best practices
- Programming Best Practices
- Provenance, Workflow Tracking, and Publishing Environments

Influences from Sources External to the Research Process

1. **Funded Research Requirements.**
2. **Publisher Requirements.**
3. **Institutional Intellectual Property Requirements.**

References and further reading

- Tools and technology
- High performance computing and reproducibility
- Influencing policy and changing the culture
- Education, courses, and training
- References and Further Reading
- Other Online Best Practice Guides for Research

Open Questions and Challenging Cases

1. Case studies in reproducible results

Retrieved from "http://wiki.stodden.net/index.php?title=Best_Practices_for_Researchers_Publishing_Computational_Results&oldid=131"

Categories: Reproducible Research | Computational Reproducibility

-
- This page was last modified on 4 May 2013, at 14:03.
 - This page has been accessed 6,913 times.
 - Content is available under Creative Commons Attribution.