

# What's holding us back on our way towards *Open Science* ?

Willi Rath | [wrath@geomar.de](mailto:wrath@geomar.de)

Based on discussions with: *Martin Claus, Claus Böning, Torge Martin, Markus Scheinert, Klaus Getzlaff, Christina Roth, Katharina Höflich, Rafael Abel, Arne Biastoch, Kristin Burmeister, Julia Getzlaff, Inga Koszalka, Carsten Schirnick, Claas Faber, Knut Günther, GEOMAR Data-Management Team, GEOMAR IT Department, ...*

online slides — <https://willirath.github.io/> → Talks

This is *not* a talk.

# Open Science

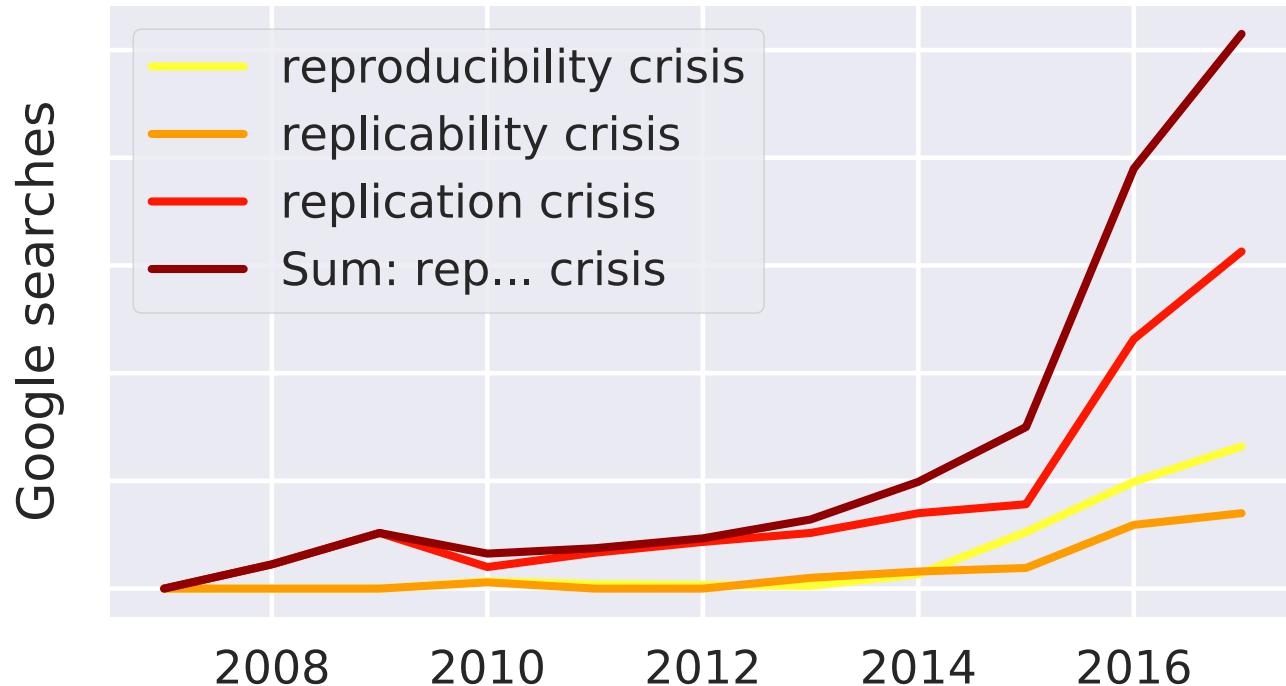
- ...
- *Open Educational Resources*
- *Open Access*
- *Open Peer Review*
- *Open Methodology*
- *Open Source*
- *Open Data*
- ...

[https://en.wikipedia.org/wiki/Open\\_science](https://en.wikipedia.org/wiki/Open_science)

# Open Science

- ...
- *(Open Educational Resources)*
- *(Open Access)*
- *(Open Peer Review)* ← Might touch the same issues.
- *Open Methodology* ← I'll talk about this.
- *Open Source* ← I'll talk about this.
- *Open Data* ← I'll talk about this.
- ...

[https://en.wikipedia.org/wiki/Open\\_science](https://en.wikipedia.org/wiki/Open_science)



[This notebook has details.](#)

# IS THERE A REPRODUCIBILITY CRISIS?



©nature

# ... not my department?

*Public debate* mostly focused on *fraud prevention* in the medical sciences.

— I'll argue that it's      who'd benefit from reproducibility.

# ← Your Boss

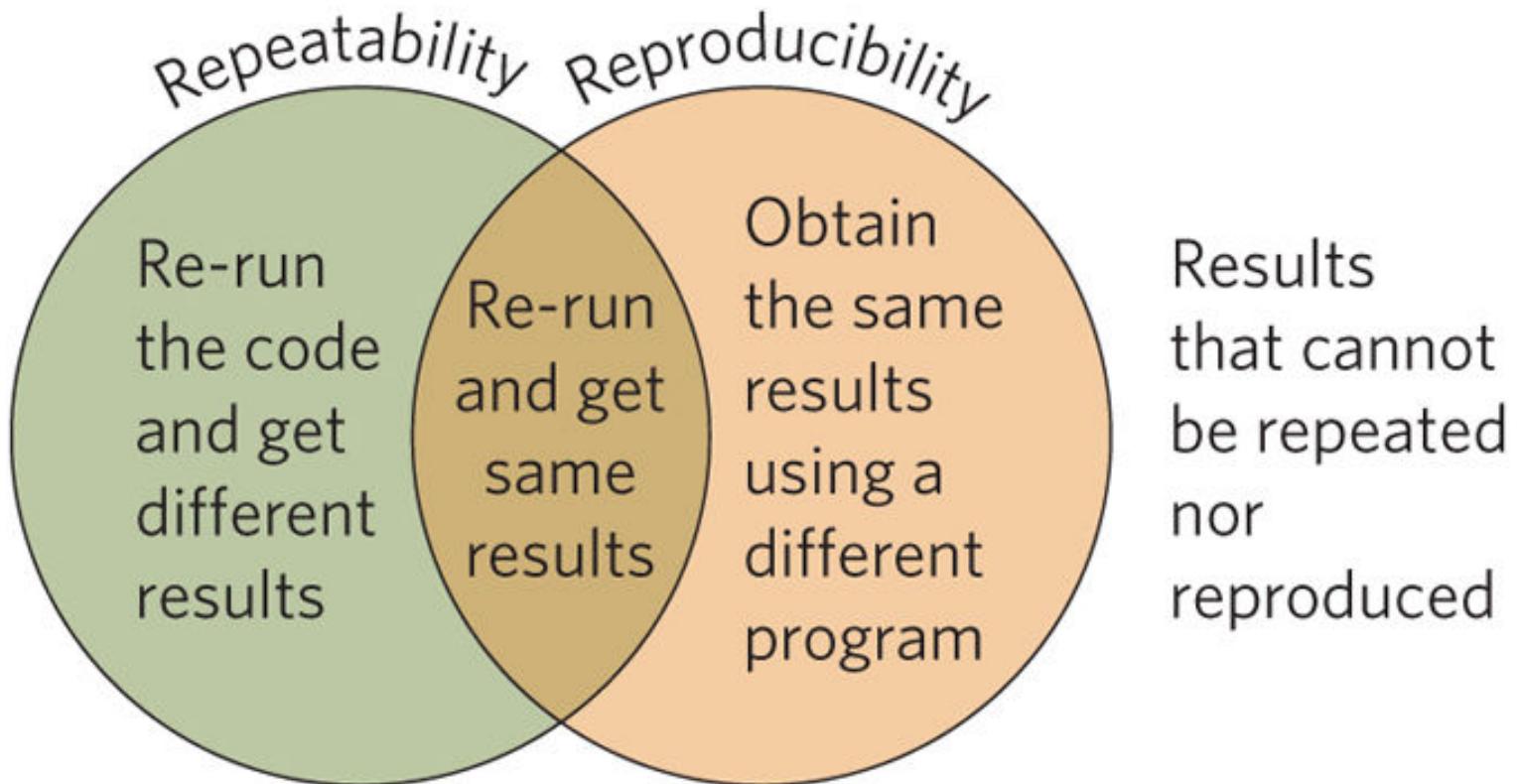


I'll prepare the talk on the plane. Can you send an update of the plot from our 2016 paper?



... 2h to take-off.

Jetzt

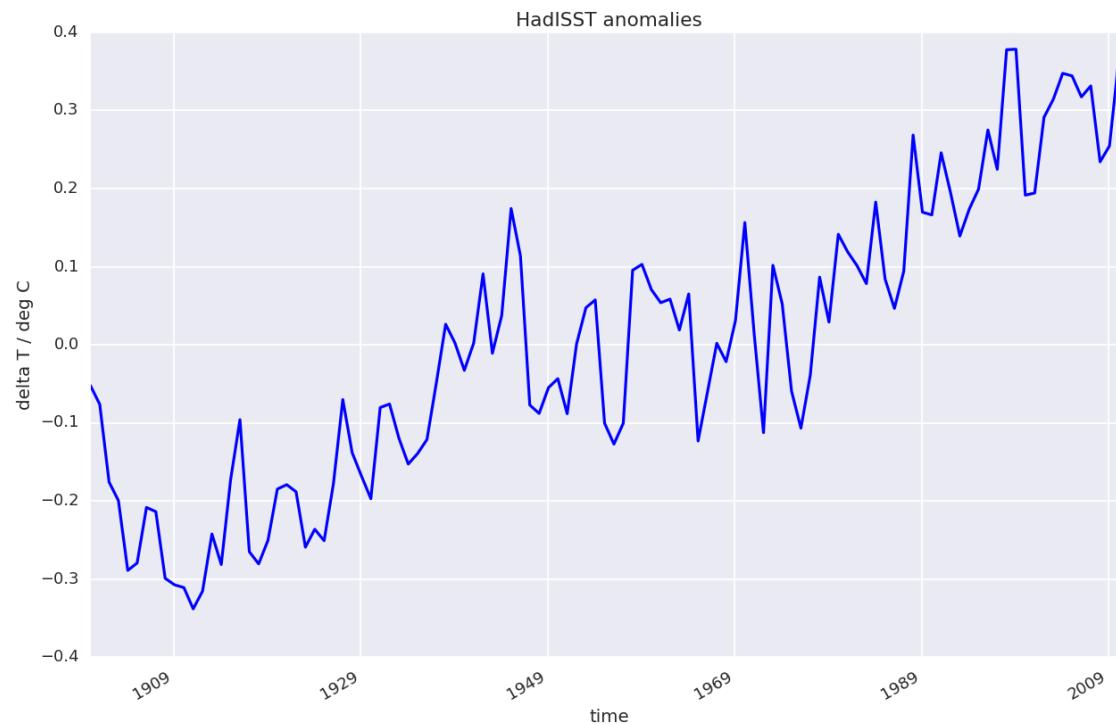


From Easterbrook (2014)

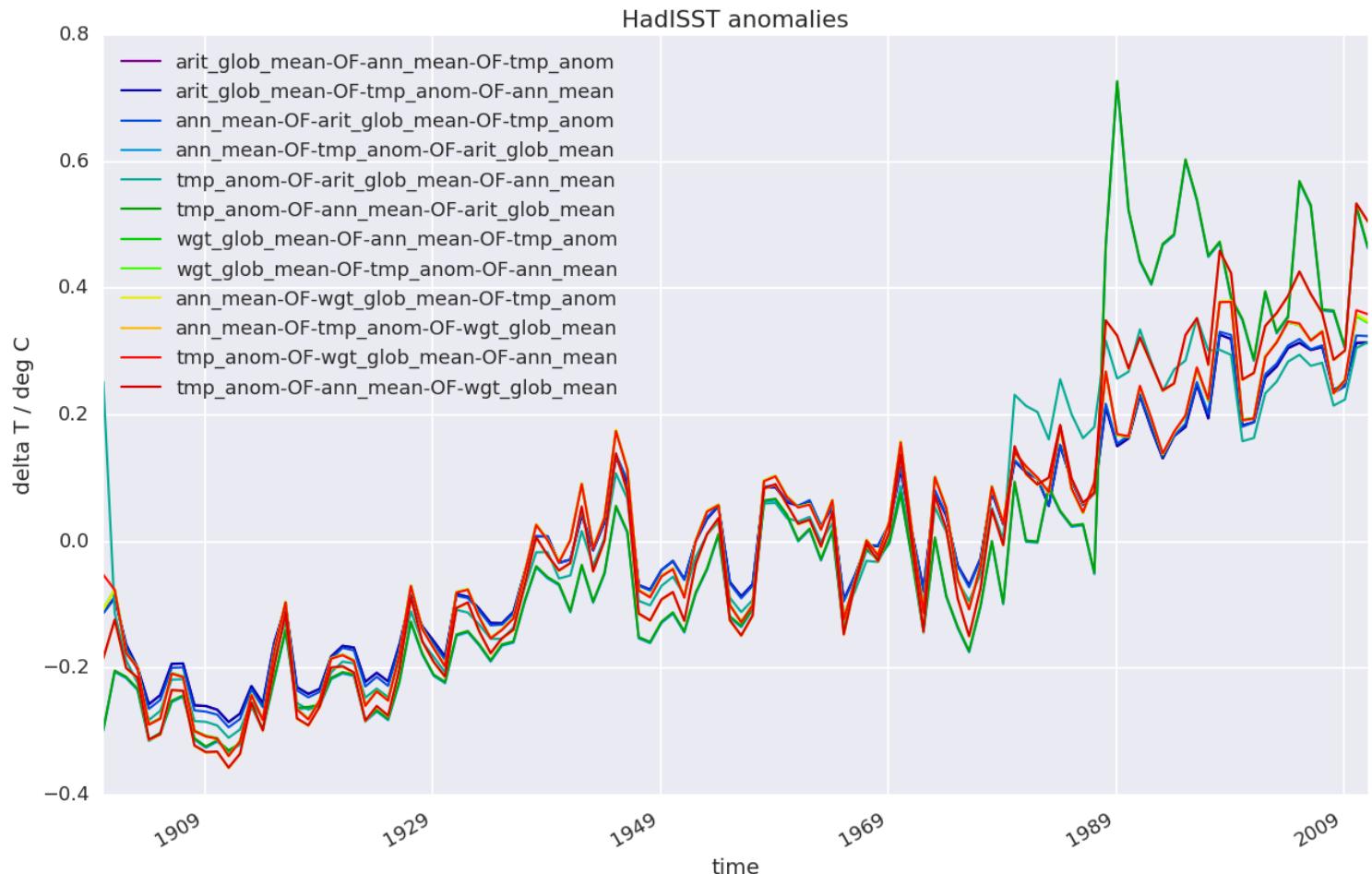
# Repeatability *Reproducibility*

Let's say an analysis is *repeatable*, if for any *sufficiently skilled* reader it is *in principle* possible to *completely understand* and *repeat all steps* the authors took from their initial idea to the final conclusions.

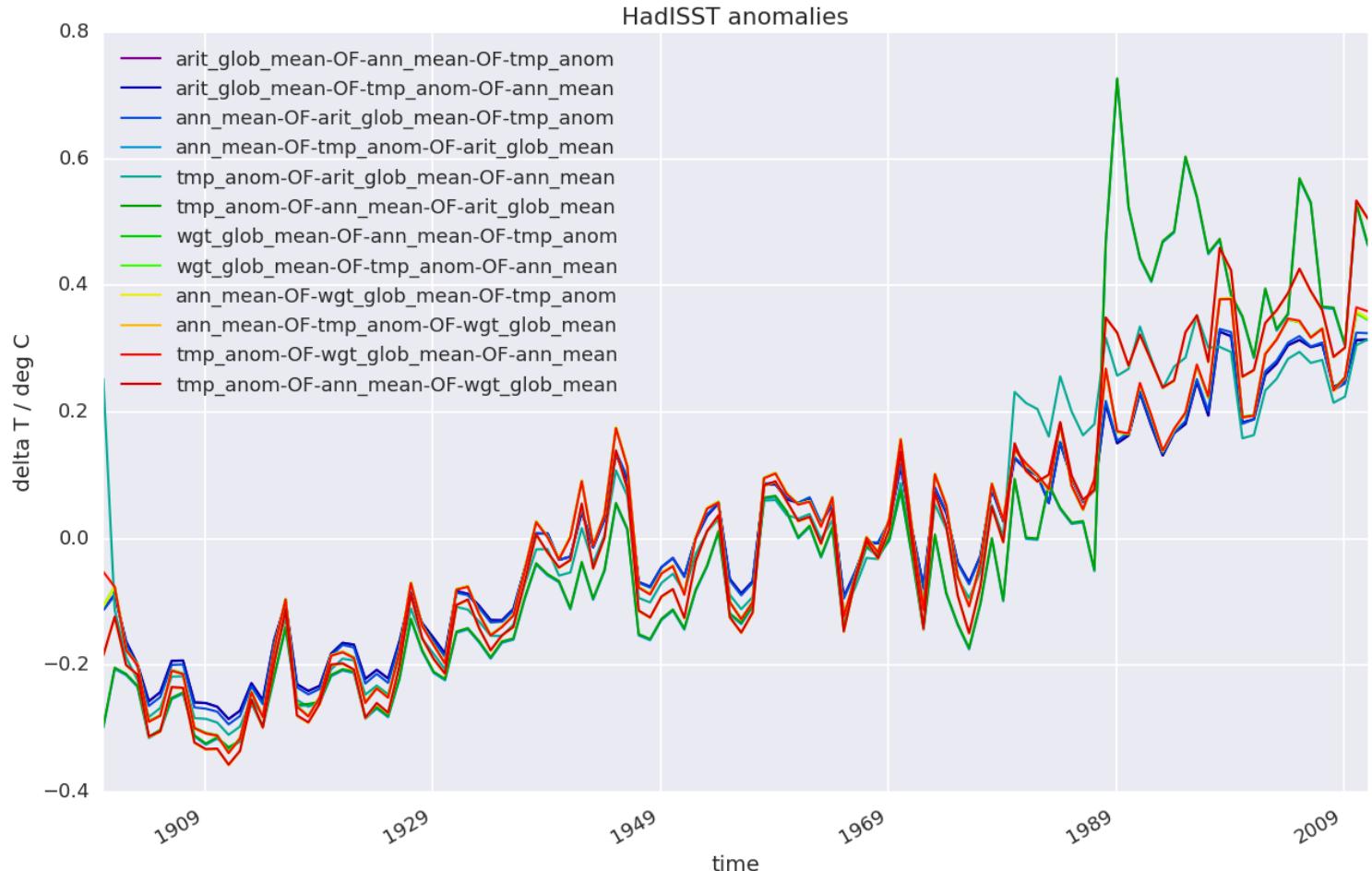
# Example — A Simple Time Series?



Annual-mean HadISST anomalies.



This notebook shows variants of this simple plot.



This notebook shows *variants* of this simple plot.

And we didn't even account for the different lengths of calendar months.



© Fabien Perissinotto

# Building Repeatable Work Flows

1. Provide a data set containing necessary to re-plot and compare the data presented in the analysis.
2. Provide fully from the original data to the final presentation.
3. Provide an overview of all the and used in the analysis and of their exact versions.
4. Provide a pointer to the full used in the analysis.
5. Provide a full of the development of the analysis. ← *That's more of a bonus.*

# Building Repeatable Work Flows

1. *all the numbers*

2. *documented steps*

3. *tools & libraries*

4. *raw data*

5. *time line*

# Building Repeatable Work Flows

1. *all the numbers* ← already required by many journals
2. *documented steps* ← already required by many journals
3. *tools & libraries*
4. *raw data*
5. *time line*

# Building Repeatable Work Flows

1. *all the numbers* ← already required by many journals
2. *documented steps* ← already required by many journals
3. *tools & libraries*
4. *raw data* ← already required by *some* journals
5. *time line*

# Building Repeatable Work Flows

1. *all the numbers* ← already required by many journals
2. *documented steps* ← already required by many journals
3. *tools & libraries* ← be prepared for those!
4. *raw data* ← already required by *some* journals
5. *time line* ← be prepared for those!



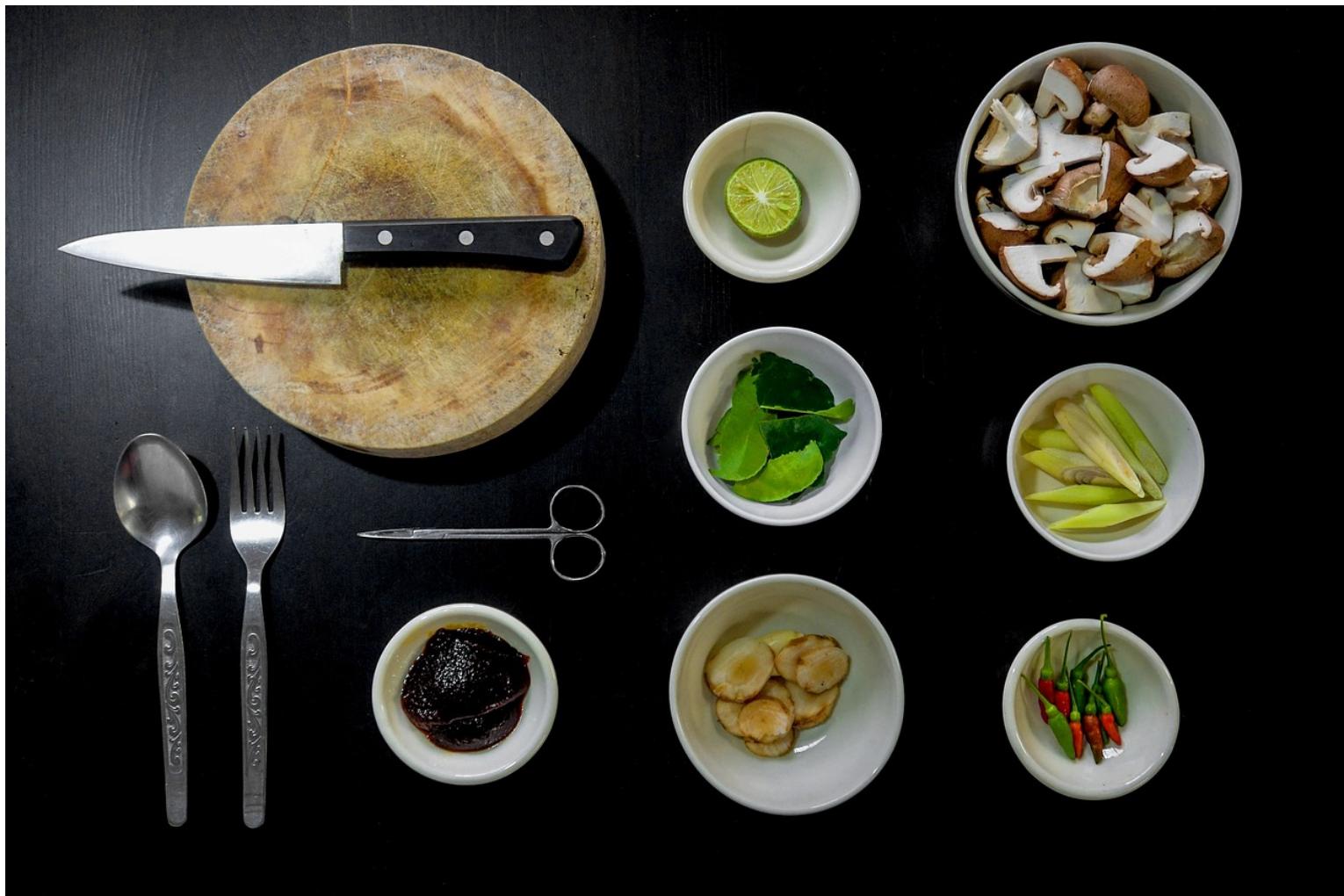
CCO-licensed



cc-by-sa/2.0 - Stepping stones across Afon... by David Purchase - [geograph.org.uk/p/5134739](http://geograph.org.uk/p/5134739)



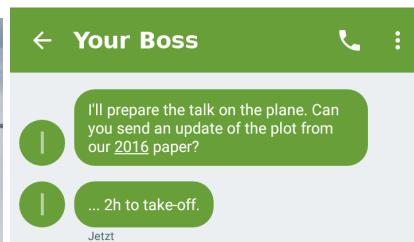
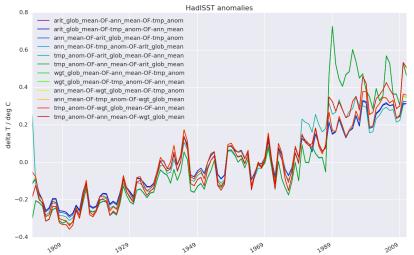
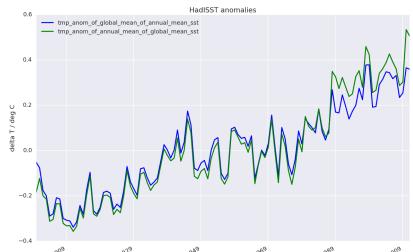
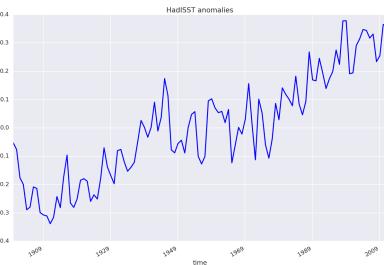
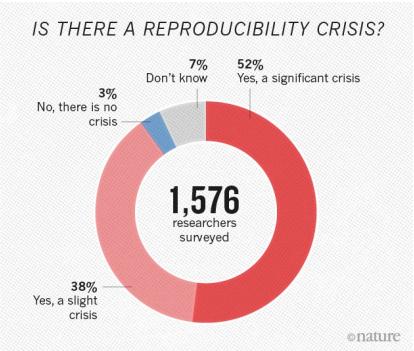
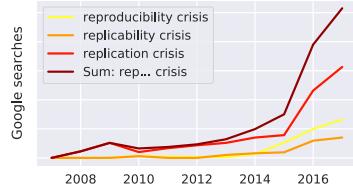
CCo-licensed

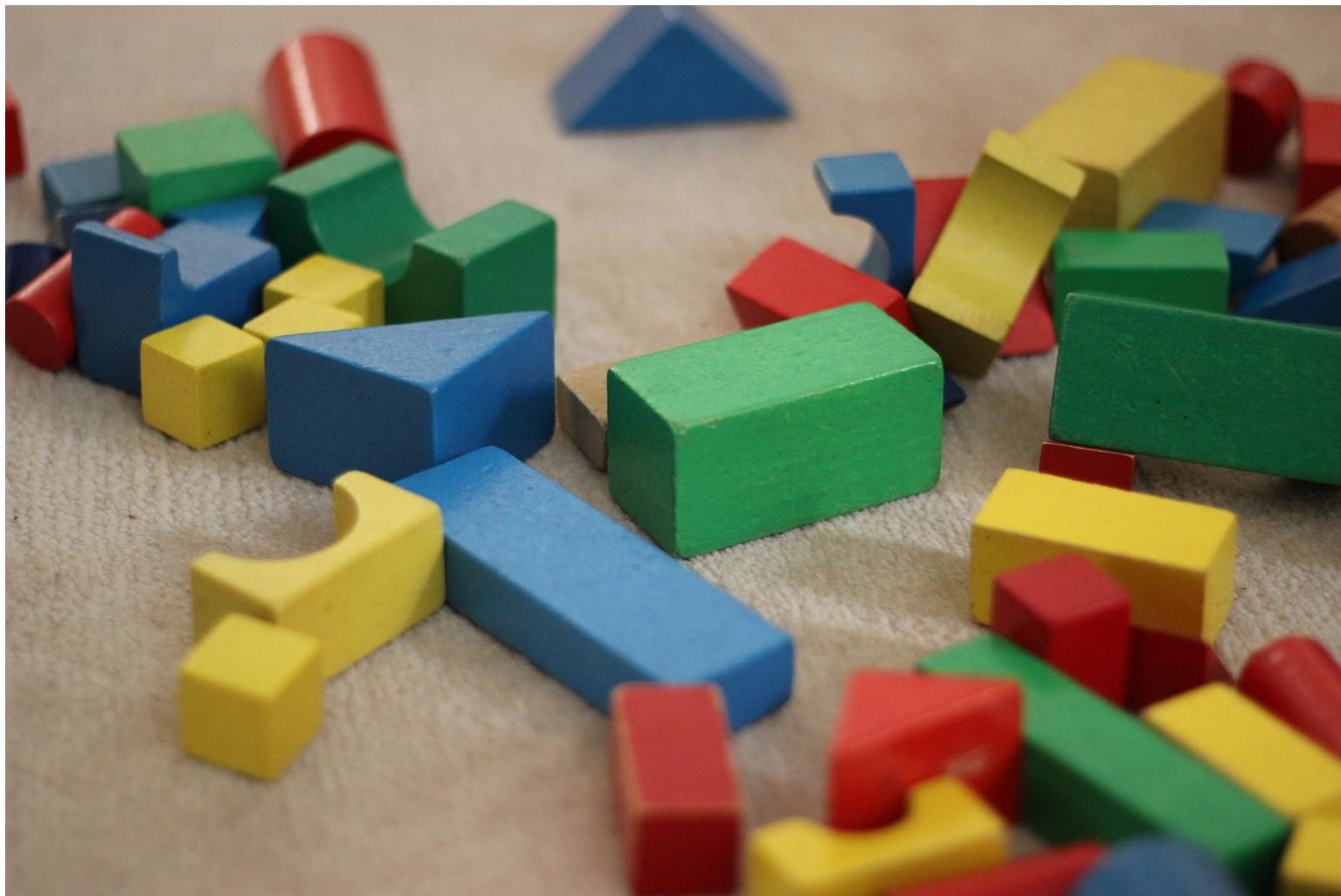


CCo-licensed



CCo-licenced





CCo-licensed



**nature** International weekly journal of science

[nature news home](#) [news archive](#) [specials](#) [opinion](#) [features](#) [news blog](#) [nature news home](#)

[comments on this story](#)

Published online 13 October 2010 | *Nature* 467, 753 (2010) | doi:10.1038/467753a  
Column: World View

Stories by subject

- [Lab life](#)

This article elsewhere

- [Blogs linking to this article](#)
- [Add to Digg](#)
- [Add to Facebook](#)
- [Add to Newsvine](#)

 Nick Barnes

Freely provided working code — whatever its quality — improves programming and enables others to engage with your research, says Nick Barnes.

I am a professional software engineer and I want to share a trade



## Open Science Isn't Always Open to All Scientists

BY CHRISTIE BAHAI, LEWIS J. BARTLETT, KEVIN R. BURGIO, AURIEL M. V. FOURNIER, CARL N. KEISER, TIMOTHÉE POISOT, KAITLIN STACK WHITNEY

Current efforts to make research more accessible and transparent can reinforce inequality within STEM professions.

**EOS**

## It's Time to Shift Emphasis Away from Code Sharing

Building well-documented, citable frameworks for Earth data analysis will encourage scientific replicability by addressing the underlying issues that inhibit code sharing.



**PLOS COMPUTATIONAL BIOLOGY**

BROWSE PUBLISH ABOUT

### Ten Simple Rules for Reproducible Computational Research

Geir Kjetil Sandve  Anton Nekrutenko, James Taylor, Elvind Hovig  
Published: October 24, 2013 • <https://doi.org/10.1371/journal.pcbi.1003285>

Article	Authors	Metrics	Comments	Media Coverage
Rule 1: For Every Result, Keep Track of How It Was Produced	Rule 2: Avoid Manual Data Manipulation Steps	Rule 3: Archive the Exact Versions of All External Programs Used	Rule 4: Version Control All Custom Scripts	Citation: Sandve GK, Nekrutenko A, Taylor J, Hovig E (2013) Ten Simple Rules for Reproducible Computational Research. PLOS Comput Biol 9(10): e1003285. <a href="https://doi.org/10.1371/journal.pcbi.1003285">https://doi.org/10.1371/journal.pcbi.1003285</a>
Editor: Philip E. Bourne, University of California San Diego, United States of America	Published: October 24, 2013	Copyright: © 2013 Sandve et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.		

- <https://www.nature.com/news/2010/101013/full/467753a.html>  
<https://www.americanscientist.org/article/open-science-isnt-always-open-to-all-scientists>  
<https://eos.org/opinions/its-time-to-shift-emphasis-away-from-code-sharing>  
<https://doi.org/10.1371/journal.pcbi.1003285>

# Reading List

- “Open Science Isn’t Always Open to All Scientists”: <https://www.americanscientist.org/article/open-science-isnt-always-open-to-all-scientists>
- “It’s Time to Shift Emphasis Away from Code Sharing”: <https://eos.org/opinions/its-time-to-shift-emphasis-away-from-code-sharing>
- “1,500 scientists lift the lid on reproducibility”: <https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970>
- “Publish your computer code: it is good enough”: <https://www.nature.com/news/2010/101013/full/467753a.html>
- “Open code for open science?” : <http://www.nature.com/geo/journal/v7/n11/full/ngeo2283.html>
- “Why bitwise reproducibility matters”: <https://khinsen.wordpress.com/2015/01/07/why-bitwise-reproducibility-matters/>
- “Which mistakes do we actually make in scientific code?”:  
<http://blog.khinsen.net/posts/2017/05/04/which-mistakes-do-we-actually-make-in-scientific-code/>
- “A Minimum Standard for Publishing Computational Results in the Weather and Climate Sciences”:  
<http://journals.ametsoc.org/doi/full/10.1175/BAMS-D-15-00010.1>
- “Good Scientific Practice at MPI-M”: <http://www.mpimet.mpg.de/en/science/publications/good-scientific-practice.html>
- “Nature - Code share”: <https://www.nature.com/news/code-share-1.16232>
- “Ten Simple Rules for Reproducible Computational Research”: <http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003285>
- “Best Practices for Scientific Computing”: <https://arxiv.org/abs/1210.0530>
- “Most computational hydrology is not reproducible, so is it really science?”: <http://onlinelibrary.wiley.com/doi/10.1002/2016WR019285/full>
  - first comment: <http://onlinelibrary.wiley.com/doi/10.1002/2016WR020190/full>
  - first reply: <http://onlinelibrary.wiley.com/doi/10.1002/2017WR020480/full>
  - second comment: <http://onlinelibrary.wiley.com/doi/10.1002/2016WR020208/full>
  - second reply: <http://onlinelibrary.wiley.com/doi/10.1002/2017WR020476/full>