

Use Jupyter Book for your supplemental material!

Traditional supplemental material

- ✗ Have to access the journal article first
- ✗ Need to download files
- ✗ Invisible by direct online searching
- ✗ Hard to find relevant data & info [1]
- ✗ **Low Reproducibility**

jupyter{book}-based supplemental material

- ✓ Independent and citable DOI
- ✓ Online HTML; easy to share
- ✓ Can be indexed online (better SEO)
- ✓ All info interconnected in one place
- ✓ Higher presentation quality
- ✓ Multimedia compatible
- ✓ Executable on the cloud (optional)
- ✓ **Better Reproducibility**

What is Jupyter Book?

Jupyter Book (jupyterbook.org) is a software tool that compiles Jupyter Notebooks and MyST markdown files (myst.tools) into fully structured written work in various printable and digital formats, such as PDF, ePub, and HTML.



Check out this QR code for an example showing real supplementary material built with Jupyter Book. The accompanying paper is published in *The Cryosphere*. [2]

- [1] Greenbaum et al., 2017, [10.1186/s13059-017-1205-3](https://doi.org/10.1186/s13059-017-1205-3)
[2] Zheng, 2022, [10.5194/tc-16-1431-2022](https://doi.org/10.5194/tc-16-1431-2022)
[3] According to Stodden et al. (2018, [10.1073/pnas.1708290115](https://doi.org/10.1073/pnas.1708290115)), only 26% of the articles in *Science* are fully reproducible.
[4] According to Gabelica et al. (2022, [10.1016/j.jclinepi.2022.05.019](https://doi.org/10.1016/j.jclinepi.2022.05.019)), the chance to get the requested data is likely < 10%.

How does the new method work?

If the source files are hosted on GitHub, we can use GitHub Actions to make Jupyter Book automatically build and publish the written work to GitHub Pages. With a proper data source setup (e.g., Zenodo), we can further use a cloud computing environment (e.g., mybinder.org) to execute and reproduce all the results presented in the written work.



What are the current practices?

The current practices for preparing supplemental information can be classified into five categories based on how much effort is needed to reproduce the presented work. The results resemble a pyramid shape with $\sim 1/3$ of papers sticking with “data and code available upon request.” (also [3])

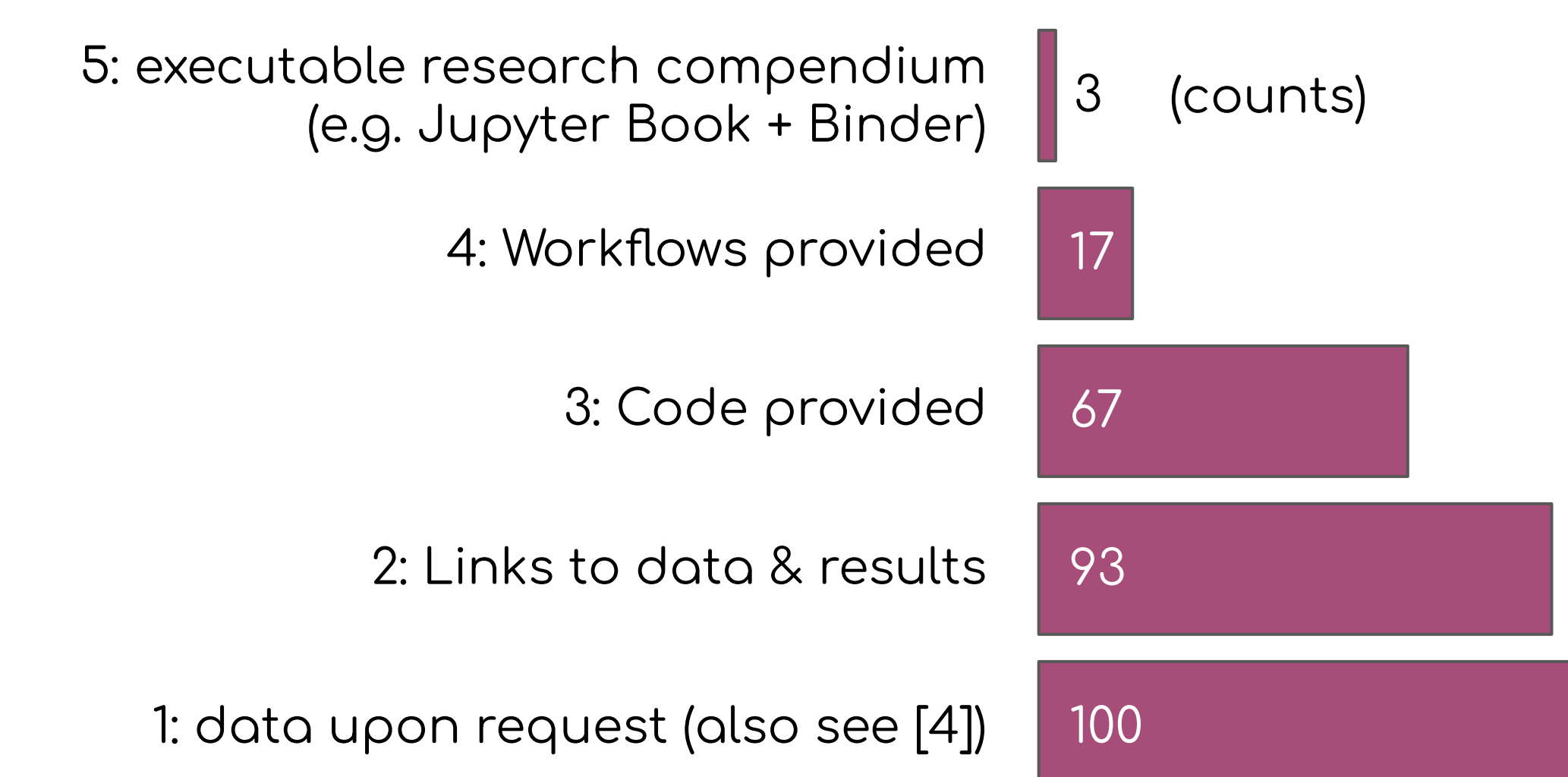


Figure 1. “Level of reproducibility” for all articles published in *The Cryosphere* between June 2021 and May 2022 (N=280).

Jupyter Book-based Supplemental Material: a FAIR Practice to Connect Research Articles with Scientific Data

Whyjay Zheng¹, Fernando Pérez¹, Chris Holdgraf^{1,2}, Erik Sundell³, Matthew Siegfried⁴, Tasha Snow⁴, Shane Grigsby⁵, Facundo Sapienza¹, Jonathan Taylor⁶, Executable Books Community

¹UC Berkeley Consulting

²2i2c.org

³Sundell Open Source

⁴Colorado School of Mines

⁵NASA Goddard

⁶Stanford University

This work is part of the Jupyter meets the Earth project, supported by the NSF EarthCube program (awards 1928406 & 1928374).



#AGU22
ED12C-0369