Open Science: Yes You Can

Cédric H. David

2022-11-09

NASA Earth Surface and Interior Solid Earth Team Meeting, La Jolla, CA.

Jet Propulsion Laboratory
California Institute of Technology

The components of Open Science

Open science, you don't have to do it (or do you?), but you might learn some tricks regardless

"This AGU journal wants me to publish my code and data, I have no clue how"

"My team members keep building the same code to analyze the same data, I wish there was a better way"

"Can you believe how these authors analyzed data from my satellite? They have no clue!"

"Sorry, that was like 10 years ago, I don't remember this specific detail of my analysis"

"I'm afraid to update my code, it might break things"

"Can't do open science at NASA, I'll get fired!"

1. Paper

EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF SCIENCE AND TECHNOLOGY POLICY
WASHINGTON, D.C. 20502

February 22, 2013

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: John P. Holdren
Director

SUBJECT: Increasing Access to the Results of Federally Funded Scientific Research

EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF SCIENCE AND TECHNOLOGY POLICY
WASHINGTON, D.C. 20502

August 25, 2022

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: Dr. Alondra Nelson
Deputy Assistant to the President and Deputy Director for Science and Society
Performing the Duties of Director
Office of Science and Technology Policy (OSTP)

SUBJECT: Ensuring Free, Immediate, and Equitable Access to Federally Funded Research

4. Methods

5. Example
6. Bonus
7. Challenges/Limits
8. "How to" at JPL?
9. Recommendations
10. Reads

2. Data

3. Software

This document has been reviewed and determined not to contain export controlled technical data.

Open Paper

- Publish your paper
  - Good: Drop your submitted manuscript in a NASA repository or an open archive (e.g. Earth and Space Science Open Archive ESSOAr)
  - Better: Choose journals that offer an option for open access fees and pay the fee
  - Best: Choose one of the many journals that only offers open access because NASA/JPL won't pay for optional open access fees

TL, DR

Where?
Science Advances
Scientific Reports
AGU Advances
AGU Earth and Space Science
EGU journals

License?
Selected by journal

Open Data

- Publish your data
  - Good: pick a license (https://creativecommons.org/choose/) and drag/drop your files to Zenodo (or FigShare)
  - Better: include description of the data
  - Best: adopt popular file formats and metadata conventions in your field (e.g. netCDF Climate and Forecast Conventions)

Free of charge, version controlled, and you get DOIs!

TL, DR

Where?
https://zenodo.org

License?
CC-BY

Open Software

- Publish your code
  - Good: pick a license (https://choosealicense.com/) and drag/drop your files to GitHub (or BitBucket)
  - Better: include some description (can be comments within code) and basic installation instructions
  - Best: Describe software dependencies (Linux: apt-get, Mac: brew, Windows: Chocolatey; Python: pip), use continuous integration (e.g. Travis), release an image (e.g. Docker)

Free of charge, version controlled, DOIs with Zenodo!

TL, DR

Where?
https://github.com

License?
BSD 3-Clause

Open Methods

- Publish your methods
  - Good: short file with copy/paste of how your program was run
  - Better: executable scripts
  - Best: Jupyter Notebook

It'll help remember how you actually did this!

TL, DR

Where?
With software

License?
Same as software

Example

Emery et al. (2020)

Paper

https://doi.org/10.1175/JHM-D-19-0084.1

Data

This version:

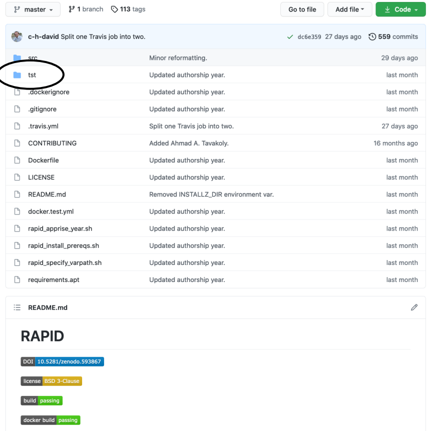
https://doi.org/10.5281/zenodo.3688691

All versions:

https://doi.org/10.5281/zenodo.3688690

Methods

Software



This version:

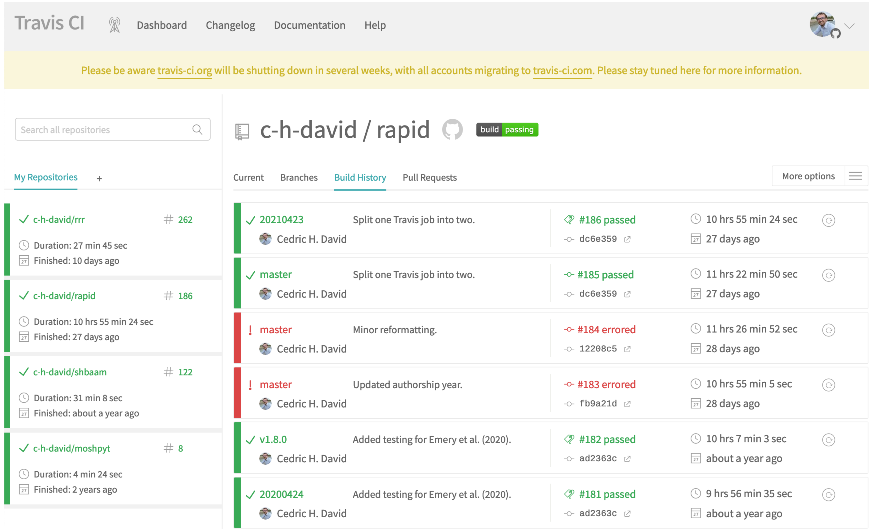
https://doi.org/10.5281/zenodo.3765342

All versions:

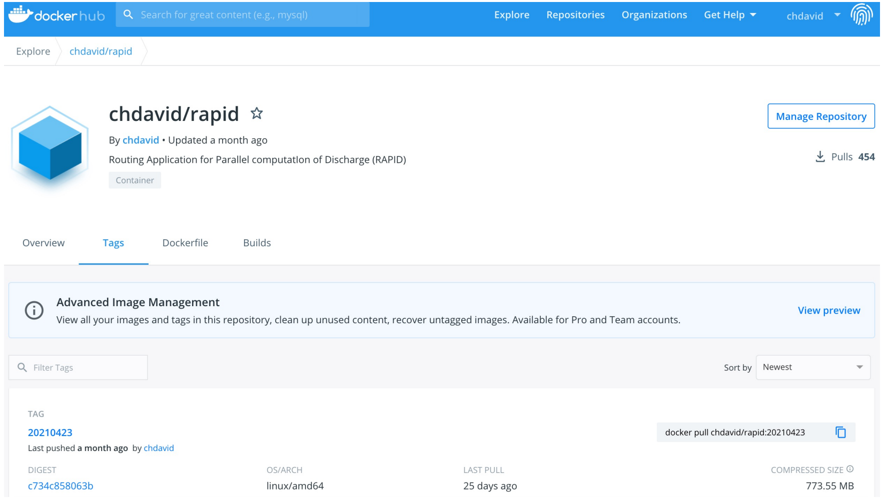
https://doi.org/10.5281/zenodo.593867

Bonus stuff you'll learn to love about open science

Continuous Integration (Travis CI)



Containerization (Docker)



<u>Peace of mind</u>: every update (even minor is fully tested), you can know if you broke anything
<u>Less email traffic</u>: "I can't install your code!", "what's wrong with my file?"
<u>Faster research</u>: next team member (yourself included) can grab the previous study and build on it

Community Challenges and the "Limits of Sharing"

- Technical training → we need to know/teach how to do this
- Self-perceived inadequacy → "not good enough"
- Documentation → takes time and effort
- Acknowledgment, evaluation, recognition of digital scholarship → "carrot" (citations, metrics, annual reviews, awards)
- Sustainable sharing → open science is <u>not</u> free support
- Keep up → rapidly-evolving tools for open science require time/effort

Suggested reads


commentary

Of carrots and sticks

Jens Kattge, Sandra Díaz and Christian Wirth

Journals and funders increasingly require public archiving of the data that support publications. We argue that this mandate is necessary, but not sufficient: more incentives for data sharing are needed.


WORLD VIEW A personal take on events

Publish your computer code: it is good enough

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

NOBODY IS ENTITLED TO DEMAND TECHNICAL SUPPORT FOR FREELY PROVIDED CODE: IF THE FEEDBACK IS UNHELPFUL, IGNORE IT.


AGU PUBLICATIONS

Earth and Space Science

RESEARCH ARTICLE 10.1002/2015EA000142

A decade of RAPID—Reflections on the development of an open source geoscience code

Cédric H. David[1,2], James S. Famiglietti[1,2,3], Zong-Liang Yang[4], Florence Habets[5], and David R. Maidment[6]

https://doi.org/10.1038/ngeo2280

https://doi.org/10.1038/467753a

https://doi.org/10.1002/2015EA000142

Jet Propulsion Laboratory
California Institute of Technology

jpl.nasa.gov