QR codes are coming!

Software in the Spectrum of Open Science

Rebecca Ringuette – HDRL/GSFC/NASA

Oral presentation given at the 2023 Jack Eddy conference in Golden, CO.

What is Open Science?

Word cloud created using multiple definitions of Open Science:

- UNESCO
- NASA
- NSF
- US White House
- Europe
- Japan
- Africa
- Vietnam

Making scientific knowledge and collaborations open.

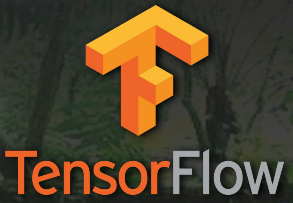
Taming the Jungle: Archiving Software FAIRly

- Efforts are ramping up to understand how to archive software FAIRly.
- Beginning with some basics:
  - **Findable:** Has a DOI, linked to other objects (e.g. publications!), searchable.
  - **Accessible:** DOI leads to descriptive landing page even if restricted software.
  - **Interoperable:** Described with standard vocabularies.
  - **Reusable:** Clear and (preferably open) license.
- Goals:
  - Increase our efficiency by making software easier to find, access and use.
  - Increase recognition via software citations and better contributor ordering.
  - **Make a curation process that is EASY!**
  - Allow for a spectrum of capabilities.

Wilkinson *et al.* (2016). https://doi.org/10.1038/sdata.2016.18

Software in the *Wild*

- General use software



- Mission Software

- Software associated with publications
- Modeling and simulation software

Open Science for General Use Software

- The Python in Heliophysics Community
  - Mission statement supports open source best practices, reusability, communication, collaboration, interoperability and community standards.
  - Membership levels will accommodate a spectrum of capabilities .
    - Entry level: Self-reporting on standards, included on the PyHC website.
    - Verified: Satisfactorily completes a code review.
    - Gold: Aligns to highest standards of PyHC.
- Collaboration between PyHC and pyOpenSci to create a code review method for Heliophysics.

New Code Review Method

Open Science for Missions

https://docs.google.com/presentation/d/1dSlMZwVBZwCVabzK4nulz_C95PhjOXA597M8KavHidw/edit?usp=sharing

https://doi.org/10.5281/zenodo.8415584

- Missions need ideas and guidance on how to ***be more competitive*** in the open science tasks in proposals.
- NASA SPD-41a requirements for missions' software are higher than others:
  - Have a persistent identifier (e.g. DOI).
  - Code of conduct, contribution guidelines.
  - Open source licenses (e.g. Apache 2.0, BSD).
  - Outputs aligned with community standards.
  - **Developed in the open!**
- Not all missions can do this, how to compensate with other ideas?
  - Task ideas related to publications, data, software practices.
  - Also for **mission culture and collaboration** with the community!
- The community deserves the chance to say how they want to be judged!

JOIN THE CONVERSATION!

Open Science for Software Associated with Publications

FULLY CLOSED → **FULLY OPEN**

No data, software, or instructions cited or available.

Includes plotted data, plotting scripts, mostly incomplete instructions.

Partially complete citation list for data and software, missing components, instructions lacking critical pieces.

What is
DIFFICULT?
easy?
critical?

Get your voice heard.
Join the experiment!

Complete citation list for all data and software used, mostly complete instructions.

Intermediate datasets included, all analysis scripts included, detailed instructions.

Completely executable!

DOI: 10.3389/fspas.2022.977781

Image Idea: NASA TOPS, slide 19 of https://doi.org/10.5281/zenodo.5621673

https://www.doi.org/10.17605/OSF.IO/V4DRT

Magnetopause Open Validation Experiment (MOVE)

Open Science for Modeling and Simulation Software

- Goals for software intended for limited use → publication scripts.
- Goals for software intended for wide use
  - Data readers, plotting routines, interpolator codes, even modeling codes.
- Build upon requirements for Mission software:
  - Have a persistent identifier (e.g. DOI).
  - Code of conduct, contribution guidelines.
  - (preferably) Open source licenses (e.g. Apache 2.0, BSD, MIT).
  - Outputs (and inputs) aligned with community standards (API vs file formats).
  - **Developed in the open?**
  - **Installable by the community?**
- **What is useful?**

https://www.youtube.com/watch?v=dgB8Ry_cRwE&list=PLChfyH8TVDGmWyGePxToZ4ZcNiBkFKubI&index=4

Discussion on Software Standards

https://tinyurl.com/JackEddy2023Software

- Miro board etiquette:
  - We are about to use a Miro board to facilitate input from the entire group. **Please create an account**. (should be only a few clicks).
  - Please do not edit or delete comments or other inputs from others.
- Discussion goal: Get an initial pulse on what is important.
- Discussion questions:
  - Software supporting publications: analysis scripts, plotting routines, some modeling codes
  - Model/simulation codes intended for wide use: interpolator software, file readers, plotting routines, even modeling/simulation codes themselves.
  - What is ***easy***, critical, and difficult when attempting to use someone else's code?
  - What is ***easy***, critical, and difficult when making your own code usable by others?