Model Data and Publishers: The View from AMS

Mike Friedman
Sr. Manager for Publishing Operations
American Meteorological Society

AMS
American Meteorological Society

What is the current situation for publishers?

- We've had questions and confusion from authors about software and model data for years

- Publishers' guidance around this has either been nonexistent or all over the place

- Clear guidance would be a big boost toward FAIR compliance (what is appropriate without being too much)?

What does Enabling FAIR Data project say?

- Commitment statement and author guidelines refer to "all core research outputs" that include "data, software, appropriate samples and sample descriptions"

- FAQ 2.9 and 2.13 mention the most important elements to save for large volume model output are :

  - the code (and version used) including any unique configurations,
  - input parameters,
  - run files,
  - description of the overall run environment and parameter space tested.

  (and if too much ask for help)

Where is AMS?

- Member of COPDESS, participant in Enabling FAIR Data

- Original data policy in 2013, updated in 2019 ("Full, Open, and Timely Access to Data"), more detailed guidance is imminent, policy does not explicitly include software/model data

- In process of developing parallel software-specific policy (would love to be able to point to this project's output)

- Bottom line is for authors to provide **a transparent process** for making the data and other output behind the research available to anyone upon request, if possible

What about other publishers?

- AGU has the most detailed guidance, specifically giving guidelines for studies with modeling data

- Other publishers and societies have very general guidelines that mention software, code, algorithms, protocols (Springer Nature, Wiley, Taylor+Francis, Science)

- or even more general guidelines that don't mention software at all (Royal Meteorological Society)

What are some specific difficulties?

- The code may be specific to a particular hardware or operating system configuration that is not easily replicated

- The code used for a particular study may not be executable on currently available systems even a short time after the research project has been completed

- Sections of the program may include proprietary code that the researcher is not licensed to share

- Even if best practice software standards are followed in the documentation of the code, its use may require levels of expertise that are not widely available in other research groups

- Model and/or output may simply be too large for easy archiving and sharing

Where does that leave us?

- Anxiously awaiting results of this project!

- A clear rubric to guide authors in figuring out what is appropriate to archive and share would be a huge step forward

  - Reduce confusion and frustration

  - Single resource for researchers, authors, and publishers to refer to that provides a clear path to compliance

- Need specific examples to share with authors

Thank you

mfriedman@ametsoc.org

American Meteorological Society