Life & Health Sciences/Engineering & Physical Sciences Librarian Training: Discussion

7.18.16

1. Welcome and introduction: Rick Jaffe

2. Reproducibility discussion: Anna Sackmann

* Lack of consensus on the definition of reproducibility
* Proposal to divide into three types of reproducibility:
  + Methods reproducibility: procedures or methods are exactly repeated
  + Results reproducibility: (also described as replicability) obtaining the same results from an experiment whose procedures are closely matched.
  + Inferential reproducibility: drawing qualitatively similar conclusions from a replication of the study or a reanalysis of the study.

Goodman, S. N., Fanelli, D., & Ioannidis, J. P. A. (2016). What does research reproducibility mean? *Science Translational Medicine*, *8*(341), 341ps12–341ps12. Retrieved from http://stm.sciencemag.org/content/8/341/341ps12.abstract

3. Research Data in the Sciences: Funder and publisher data requirements: Elliott Smith

(see slides in Discussion 1 folder: RDM\_2016-07-18\_ES.pptx

4. Discussion

Q1: How can I make my data and metadata available to the public? Does the university have a set policy on this?

Comments:

* Would GitHub be sufficient? Maybe, but it depends on your audience. It’s a great tool for sharing, but it probably wouldn’t fit your funder requirements.
* What about DASH? Yes. UC guarantee of long-term preservation
* Publish it as a supplemental material as your article.
* What about a time-based visualization - what kind of repository would it go into then?
* Researchers will sometimes upload an entire zipped file into a repository, but not everyone will want the full zipped file - they want a video or a screenshot. One option is to deposit an archival copy in an institutional repository and an accessible copy in a proprietary place like Github, Figshare, or Vimeo (specifically for videos)
* Researchers will sometimes ask if it’s acceptable to make their research available on their website - the university does not have a set policy on this, but this type of solution will not be sufficient for most funders.

Q2:  Do I need to make my code available to others? I don’t want to have to explain how to use it to everyone – can I just post it and let them cope?

Comments:

* some funder requirements are explicit about requiring code and a read.me file
* at least include a read.me file. especially if you’re talking about reproducibility
* should your code be self-documenting? yes! but you need enough documentation that you’ll be able to run this again in 10 years.
* Include enough documentation that will allow others to run the code on their own; otherwise, they may be returning to you with questions for years to come.

Q3: I created a visualization that is currently saved as a .gml and portions of it as screenshots (as .jpgs). Where do I store my maps and original data files? How I do link them? How do I describe how the underlying data was collected?

Comments:

* what does your final object or container look like?
  + search string in web of science
  + settings and parameters - fill out all image regarding the info
* challenge - take a screen capture of all the steps (what are the tools and what are the standards I need to help me do this?)
* Take these steps for yourself, if not for other people.
* Sometimes there are technical solutions, like Jupyter notebooks, developed here at UCB