**Coding in Public : Workshop guide**

Suggested workshop structure

1. Introduction
   1. What is reproducibility?

What is the crisis in reproducibility, what might be causing it and how are people going about trying to resolve it.

Poll questions about whether or not participants have heard about this crisis, whether they have ever tried to reproduce their own or someone else’s work, what they think the problems are, whether they think the proposed solutions will help, etc.

* 1. Why should I care about reproducibility?

Impacts of the crisis on faith in researcher, on funding and publication, and on individual researcher reputations.

Poll questions about what they think of these impacts, which they think are the most pressing, etc.

* 1. Special focus on the challenges and benefits of reproducibility in social and data sciences

Discussion around how the crisis affects various disciplines, how the different disciplines have responded, how the proposed solutions may or may not work well for some disciplines.

Poll questions about how the proposed solutions may be more or less successful in some disciplines than others, whether solutions need to be targeted or whether disciplines need to adapt, etc.

* 1. Q&A Segment: Opportunities to ask questions, suggest anything relevant that has not been covered, raise topics that may want to be addressed in detail later, etc .
  2. Brief break out groups to allow participants to introduce themselves, share their ideas on the topic and do a bit of icebreaker stuff. This is intended to make participants more comfortable rather than to really generate much serious discussion.

1. Building habits around documentation
   1. Documenting your thinking

Introduction to first steps of Get Things Done method, mind-mapping, notebooks and other idea-capture tools, etc.

Poll questions about whether participants have heard of this method, what tools they have tried before and whether they liked them, etc.

* 1. Documenting your processes and communication

Introduction to the next steps of Get Things Done method, processes for reviewing captured ideas, scheduling time for projects, establishing formal communication plans, building habits around meetings and note taking, action items, email follow ups, etc. Specific attention paid to how this documentation helps in potential disputes later.

Polls questions around experience with scheduling, communication plans, following up on spoken action items, etc. Opportunities to bring up experiences people have had with documentation and how it has helped resolve disputes.

* 1. Documenting your data

Discussion of raw data, meta data, research-ready data, DOI and more. Also covers repositories, security and sharing.

Poll questions around use of different kinds of data, storage, referencing and sharing.

* 1. Q&A Segment: Opportunities to ask questions, suggest anything relevant that has not been covered, raise topics that may want to be addressed in detail later, etc .
  2. Brief breakout in which groups discuss the tools they have used or want to try, the processes they have used or want to try, their experiences with task distribution and communication, their experiences with accessing/citing shared data or sharing data, etc.

1. Collaboration
   1. Dividing work and documenting contribution

Discussion of ‘urgent’ versus ‘important’ when prioritising tasks. Discussion around recognising strengths and weaknesses, being ok with not being perfect at everything, how to measure contribution and effort, how to apportion tasks, how to attribute credit fairly.

Poll questions around what tasks participants prefer/dislike, how much time a variety of tasks would take them, how uncomfortable they feel with admitting they can’t do something on their own, etc.

* 1. Online tools and version control

Discussion around online repositories, collaborative work spaces, co-writing sites and more. Discussion around version control and commenting.

Poll questions about whether participants have used the various tools and sites mentioned, whether they usually write useful comments (if applicable) or not, etc.

* 1. Command line and writing good code

Introduction of command line and principles of writing good code, including comments, functions, variable names and more.

Poll questions around experience with these tools, which type of code looks more readable, etc.

* 1. Deployment and sharing

Introduction of jupyter notebooks and .Rmd deployments in online sites.

* 1. Q&A Segment: Opportunities to ask questions, suggest anything relevant that has not been covered, raise topics that may want to be addressed in detail later, etc .
  2. Brief break out group discussion in which participants run through some exercises around joining a github repository, uploading documents, editing documents, making useful commit comments and more.
  3. If time permits, another set of exercises around creating and posting code deployments in either jupyter notebooks, .Rmd or both.