

Carney Summer Program

Wrapping Up a Computational Research Project

Jason Ritt

jason_ritt@brown.edu

Scientific Director of Quantitative Neuroscience



ROBERT J. & NANCY D. CARNEY
INSTITUTE FOR BRAIN SCIENCE

BROWN UNIVERSITY

<https://github.com/brownridd/cfsc2023>

What does it mean to wrap up?

Generally you will want to share your work with at least one of these groups of people:

- Lab members
- Outside collaborators
- Other scientists in your field
- Future you

They all have the issue that they will not have internal memory of what you did, so in order to build upon your work, they may first need to laboriously excavate it (or give up and start over).

Wrapping up a computational project means leaving it in a state that

- is findable and accessible
- concisely and accurately describes what you did
- can be continued and built upon with little overhead

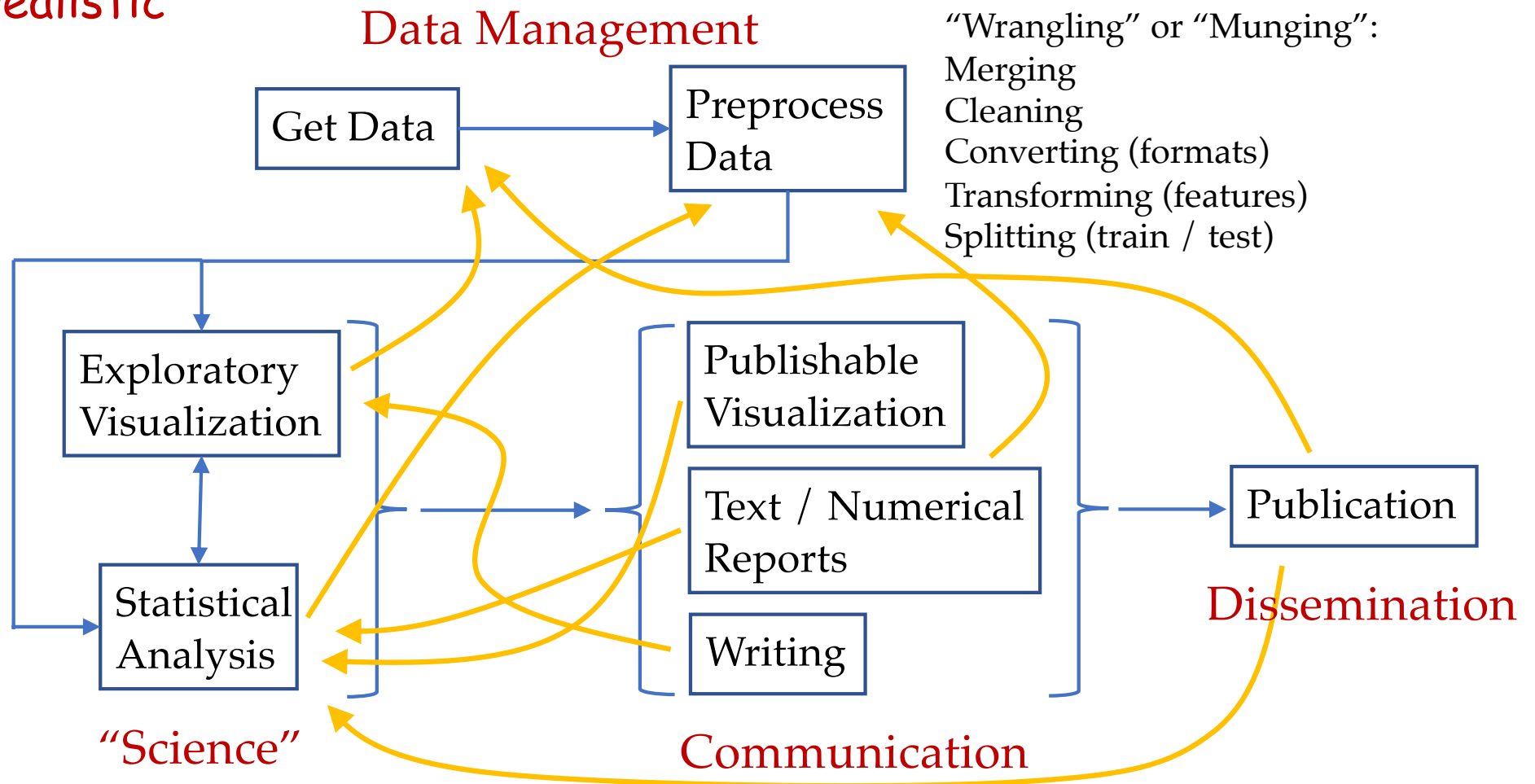
```
# DEAR FUTURE SELF,  
#  
# YOU'RE LOOKING AT THIS FILE BECAUSE  
# THE PARSE FUNCTION FINALLY BROKE.  
#  
# IT'S NOT FIXABLE. YOU HAVE TO REWRITE IT.  
# SINCERELY, PAST SELF  
  
| DEAR PAST SELF, IT'S KINDA  
| CREEPY HOW YOU DO THAT.  
  
# ALSO, IT'S PROBABLY AT LEAST  
# 2013. DID YOU EVER TAKE  
# THAT TRIP TO ICELAND?  
  
STOP JUDGING ME!
```



<https://xkcd.com/1421/>

An ~~ideal~~ scientific analysis workflow (1000 ft view)

realistic



An ~~ideal~~ scientific analysis workflow (1000 ft view)

realistic

We'd like to have the “computational fluency” to

Efficiently *implement* each step, and its connections to others

Track what we actually did (memory is unreliable!)

Validate what we did, and know how to *fix* anything wrong

Reproducibly and meaningfully *communicate* what we did

Redo parts of the flow without having to redo it all

Reuse our work in future projects

Find solutions when confronted with the unknown

Almost all of these steps could use *automation*

“Statistics”

Communication

Dissemination

Checklist for wrapping up a **computational** research project

These are prompts for reflection; not every question applies to every project.

- ☐ Is there a publication that describes *and links* to all research outputs?
- ☐ Do you have an internal summary of your work including “post review”?

- ☐ Is all essential code archived in a *findable* and *accessible* place?
- ☐ Is there a README or high level introduction?
- ☐ Is data (completely or partially) archived in a *findable* and *accessible* place?
- ☐ Is there inline documentation in the code?

- ☐ Are there instructions on how to install and run the code?
- ☐ Are all software dependencies provided in machine readable form?
- ☐ Does the code “just work”? Is it machine independent (including paths!)?

- ☐ Is there a declaration of who “owns” and is responsible for the research output?
- ☐ Is there a description of what others can do with the research output?