

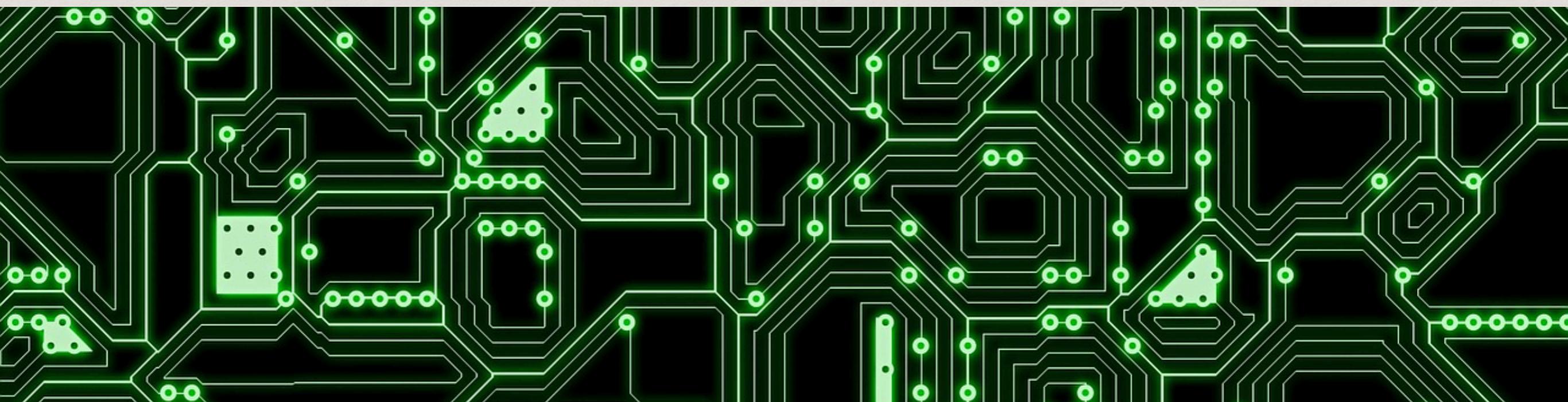


Data project management

How to keep your
data project
organized...

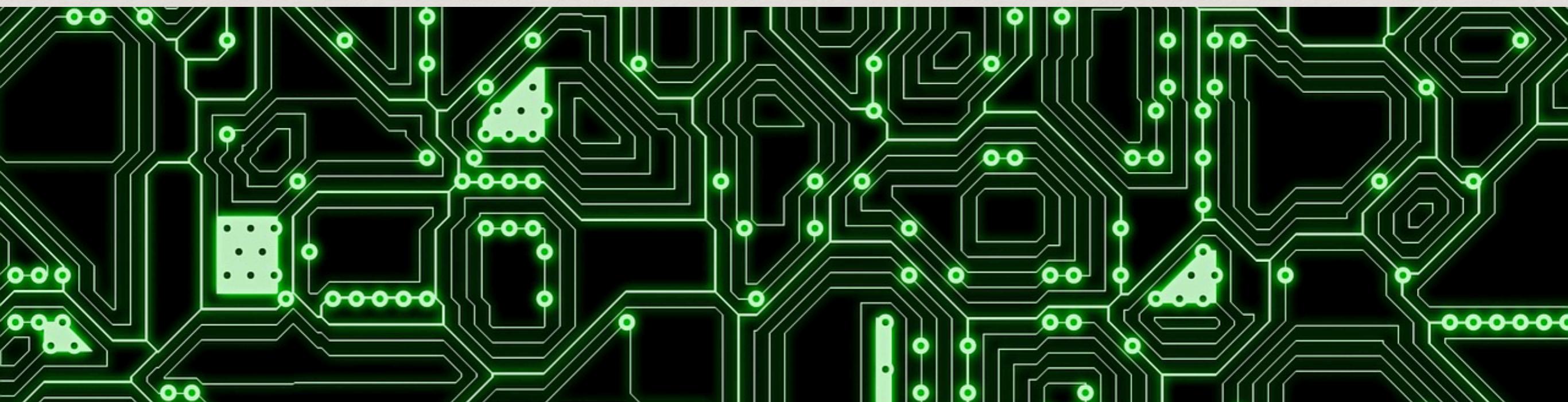


You've done so much hard work collecting
data. Now what?

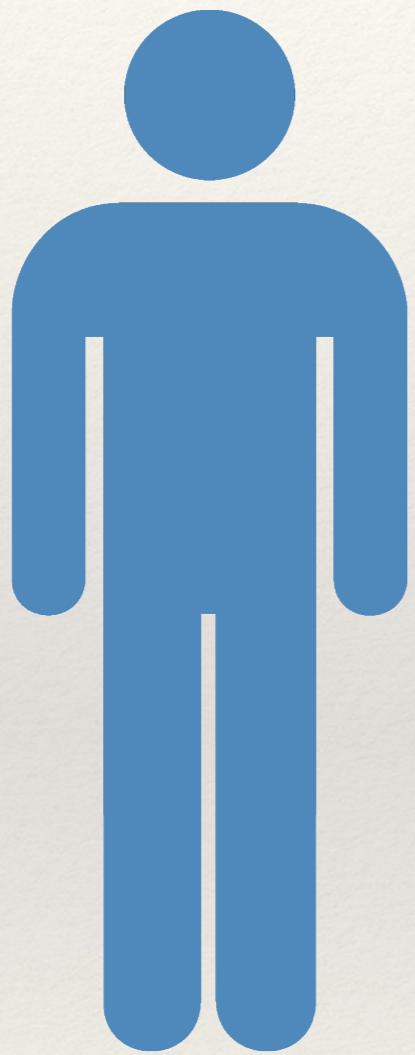




Now let's talk about how to manage your data project on your laptop.



Why? Work & collaborate more effectively



You

Your future self is your first & most important collaborator



Today



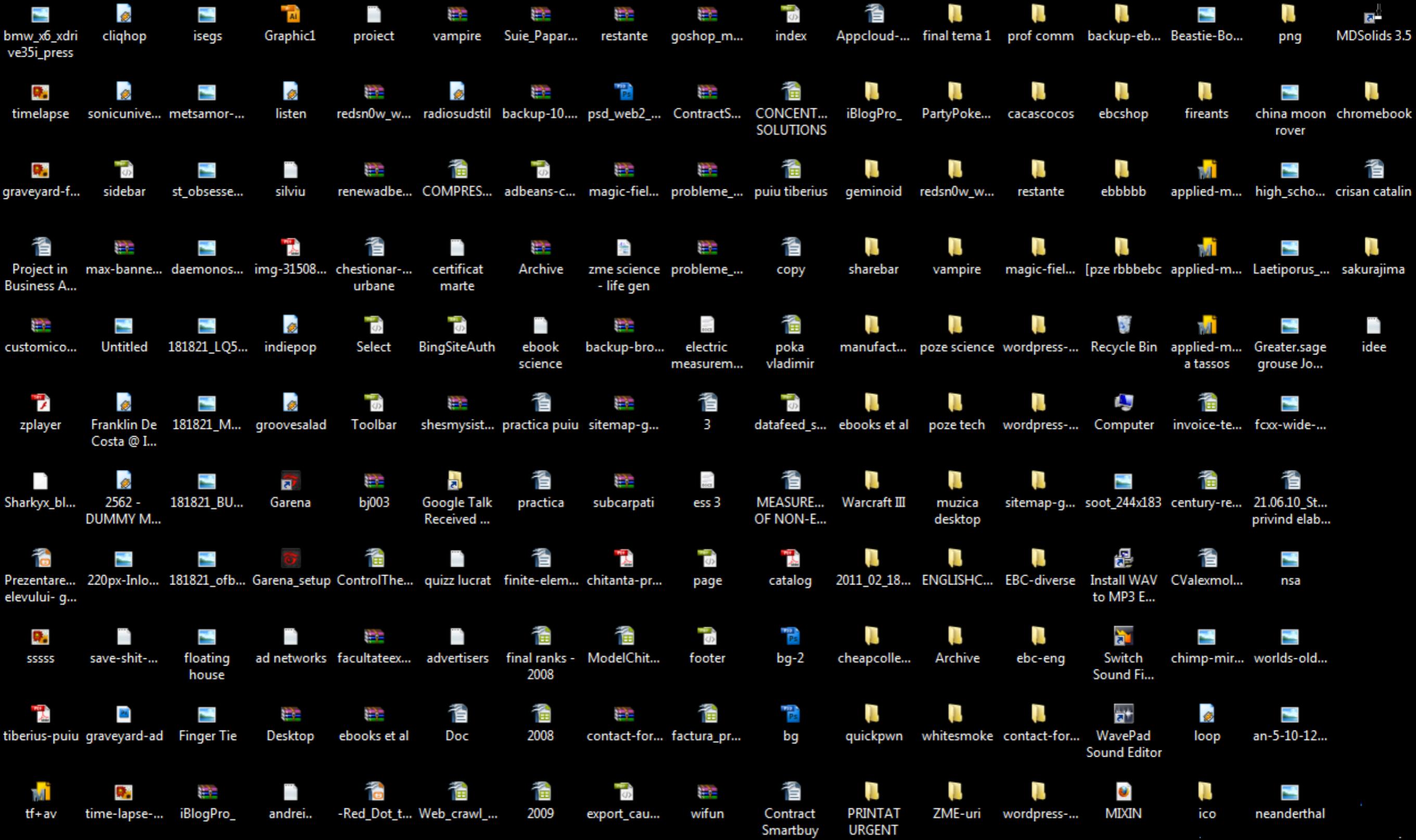
Week/month/year later

Key points:

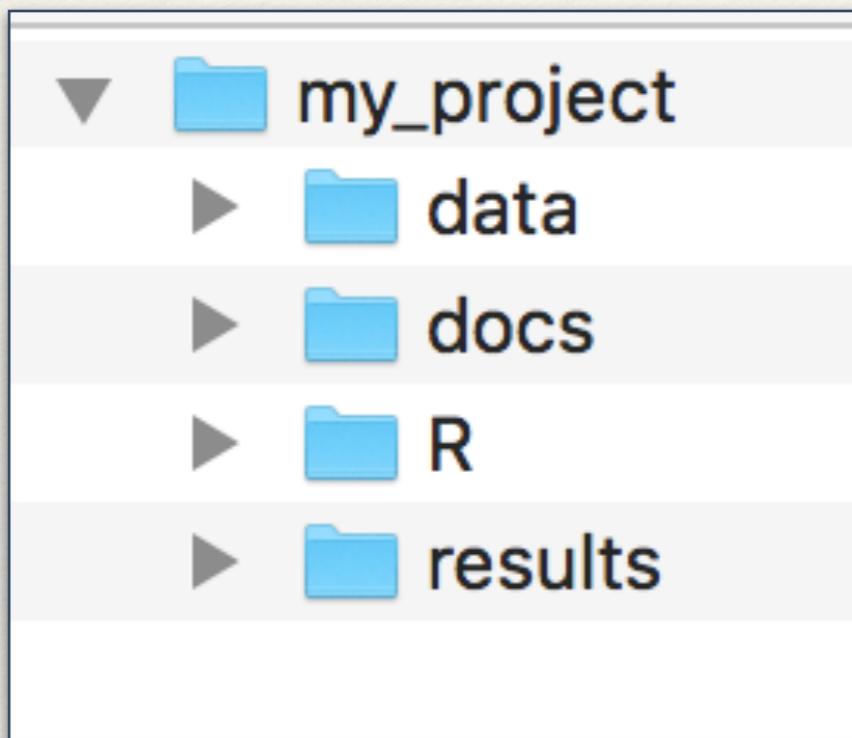
- A. Organize your project folder
- B. Protect your raw data
- C. Name your files effectively
- D. Track your project's changes
- E. Backup your project

Organize your project
folder

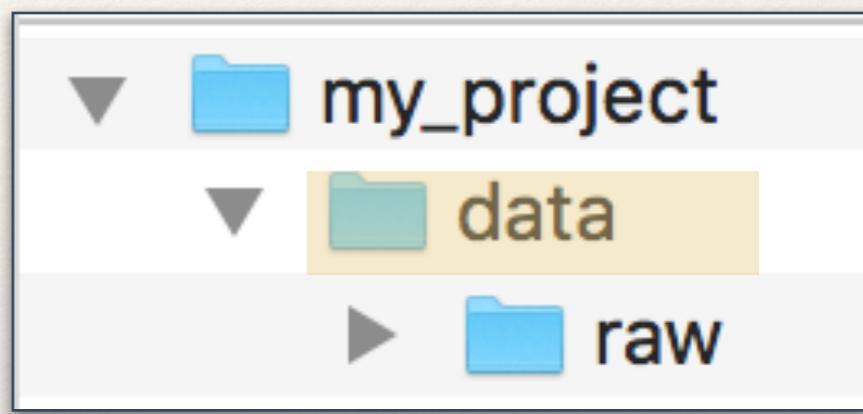
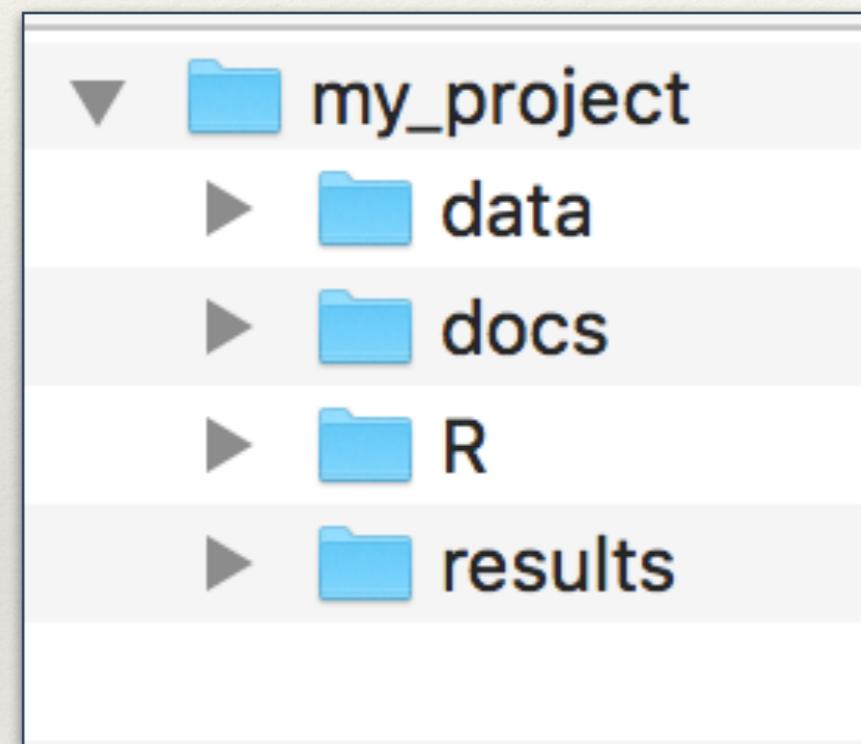




Project folder



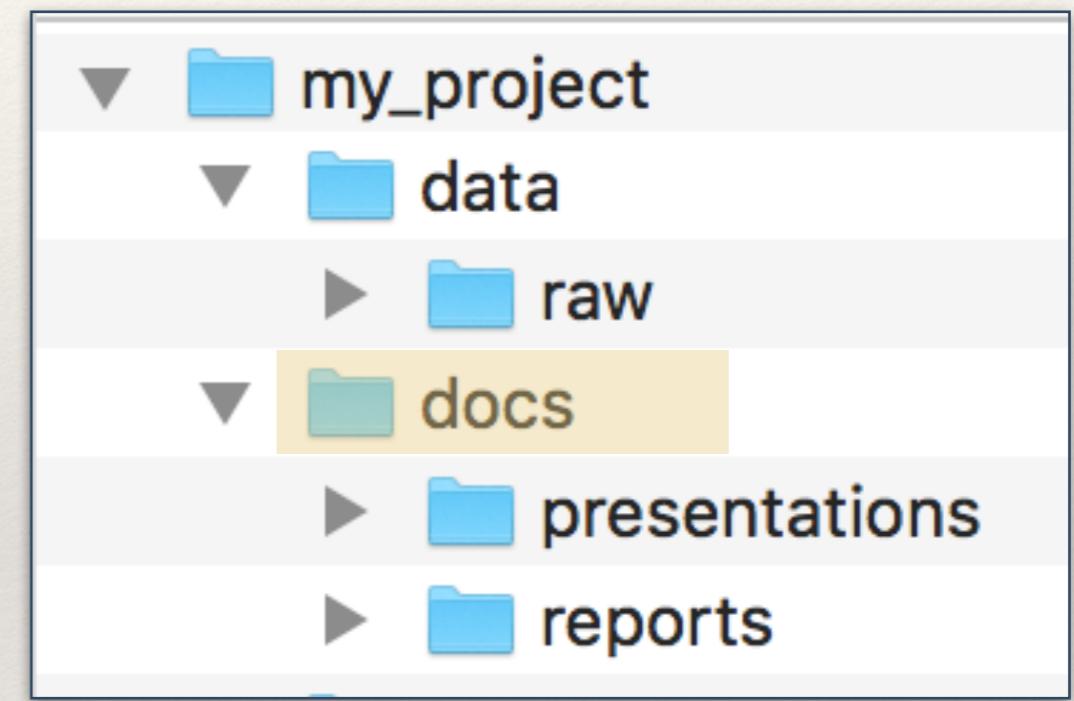
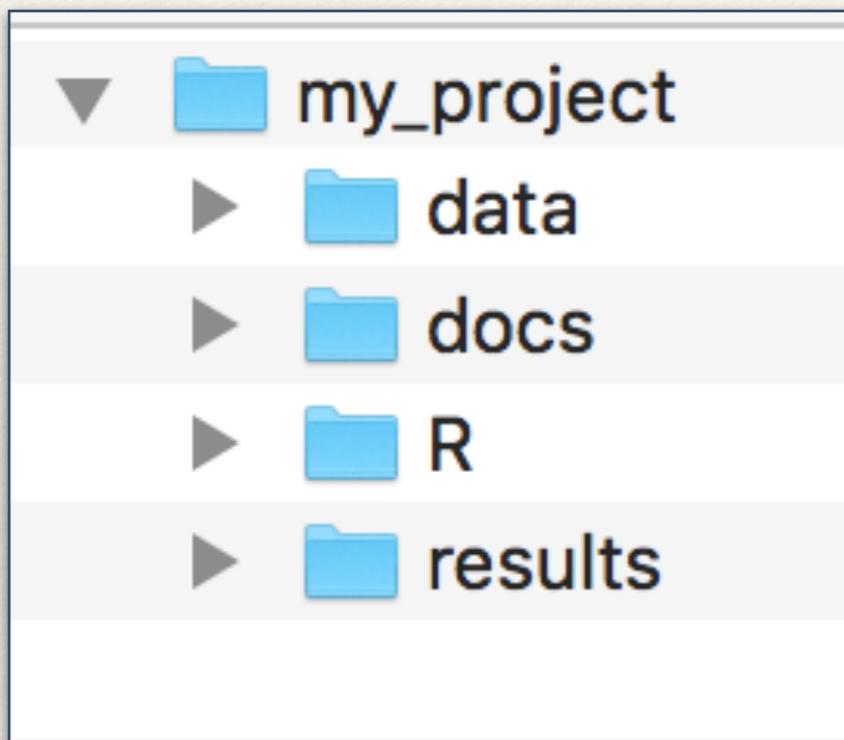
Project folder



- ❖ AnimalsSampled_Export_May16_0438.csv
❖ Site_Export_May19_2353.csv

Project folder

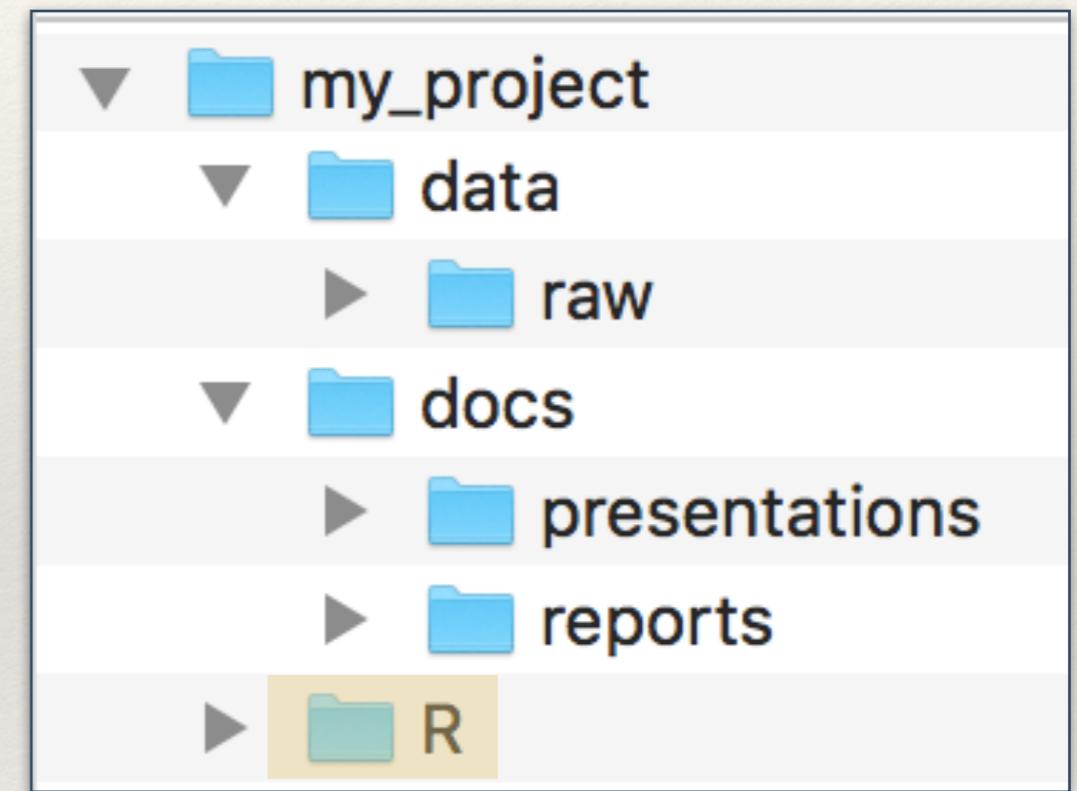
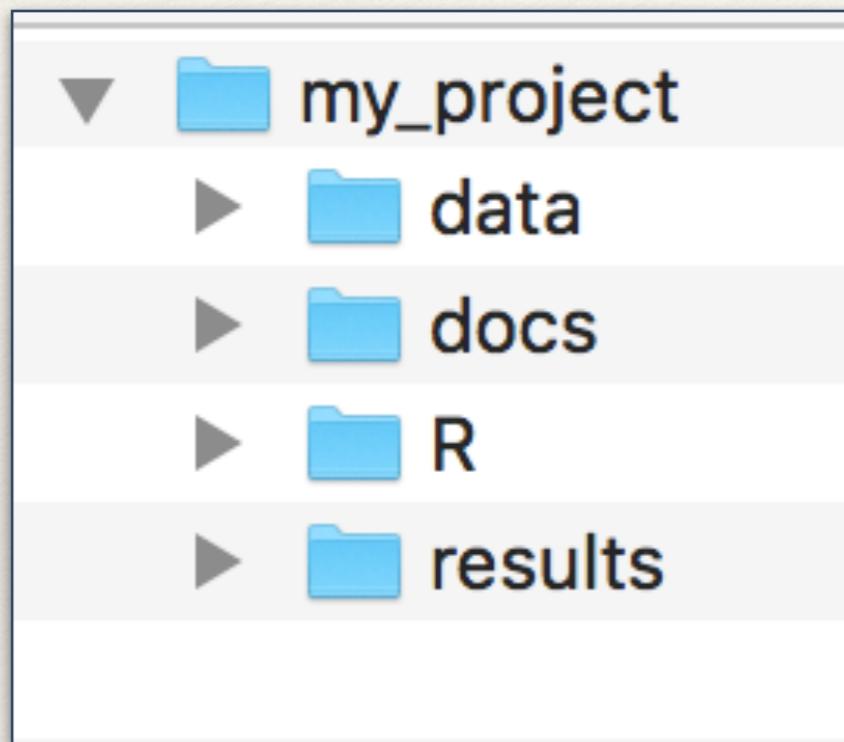
docs/documents folder



- 2019_conference-name.docx
- 2019_conference-name.pptx
- 2019_predict-country-report.docx

Project folder

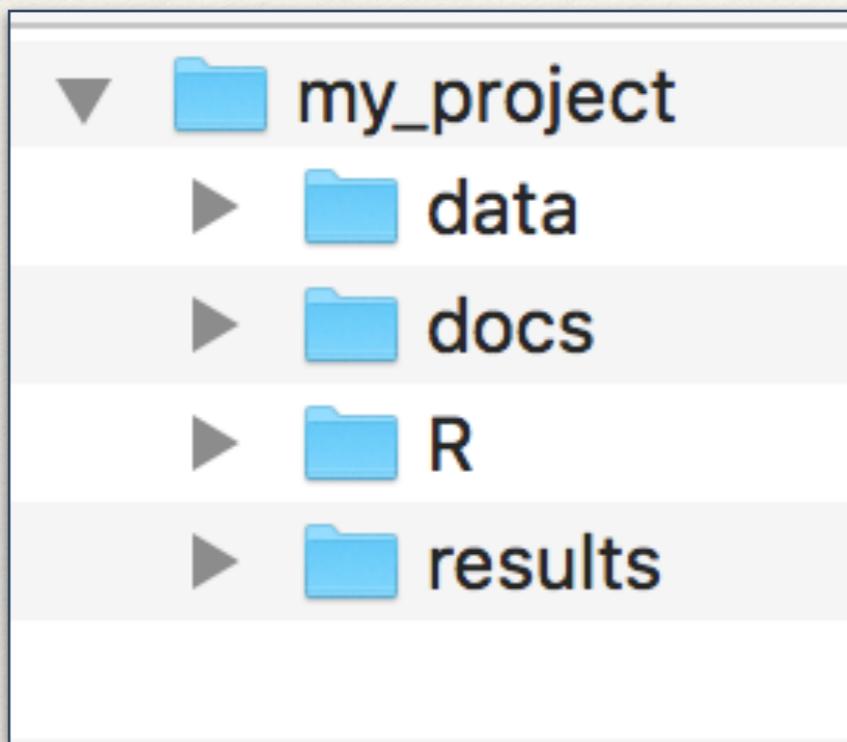
R folder



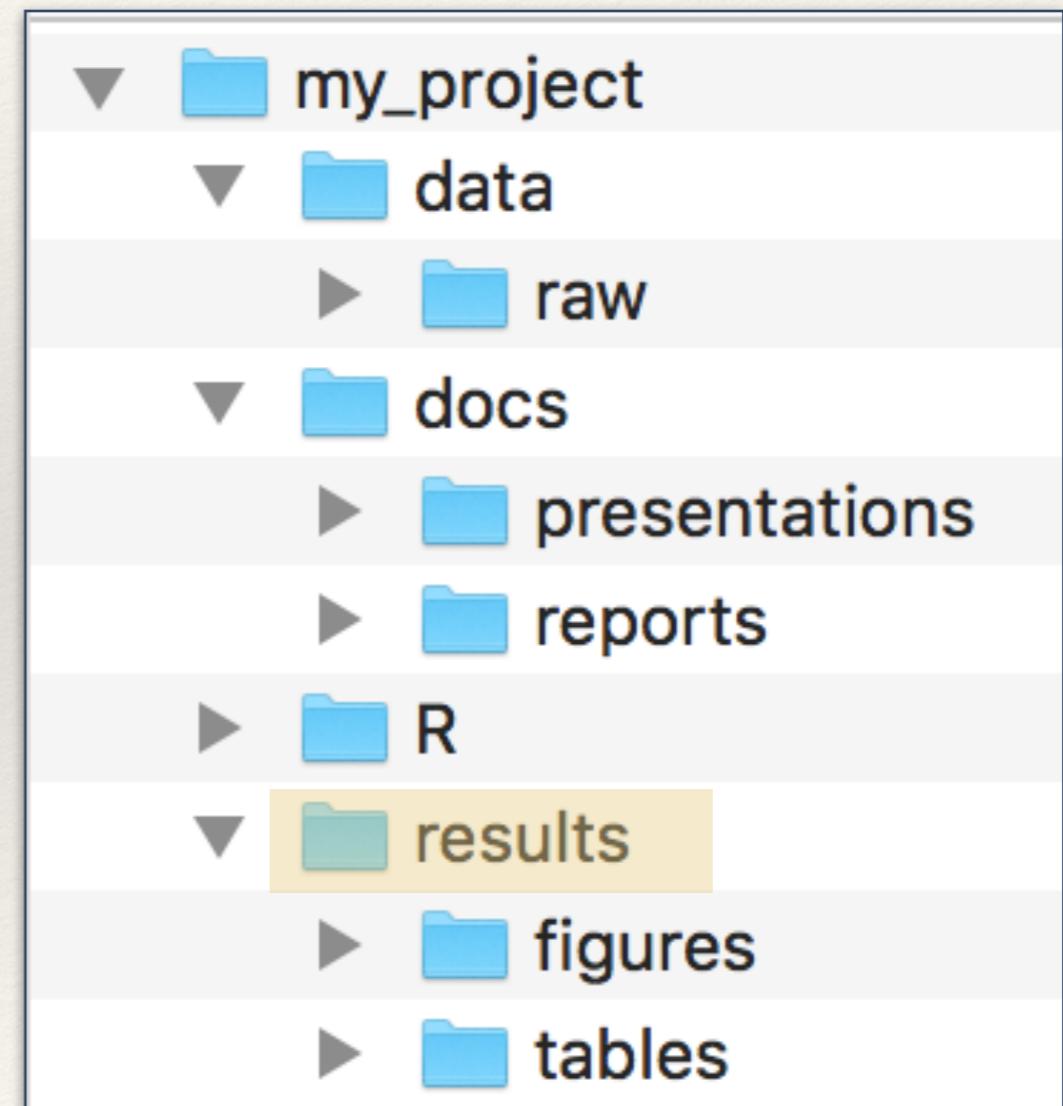
- ❖ sc01_explore-animals.R
- ❖ sc02_analyse-data.R
- ❖ sc03_country-report.R

Project folder

results folder

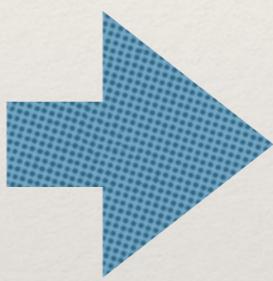


- ❖ fig01_animals-sampled.jpg
- ❖ fig02_specimens.jpg
- ❖ map01_country-map.jpg



Project folder

▼	2019_predict-conference
►	data
►	documents
►	lessons
►	R
►	results

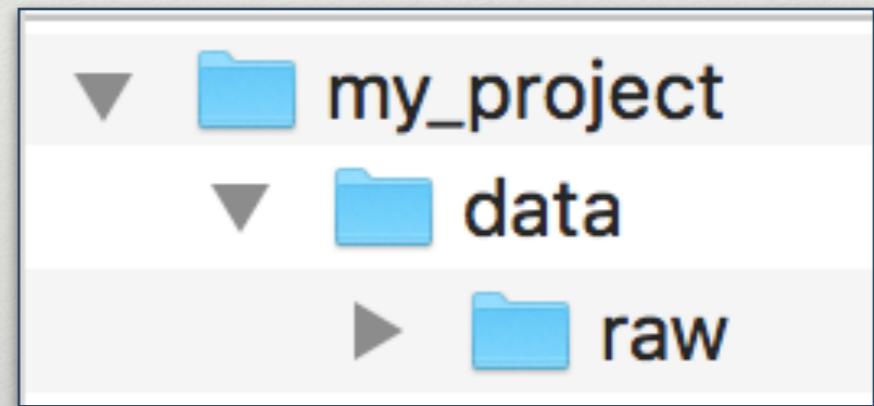


▼	2019_predict-conference
►	data
►	GIS
►	output
►	raw
►	sample
►	documents
►	eidith-bubble-forms
►	presentations
►	reports
►	lessons
►	gis
►	R
►	misc_create-sample-subset.R
►	misc_get-country-polygons.R
►	results
►	figures
►	maps

Protect your raw data

- Do NOT edit raw data directly
- Copy and work with it so the original data is not modified

data folder



Name your files effectively

Untitled 138.docx
Untitled 241.doc
Untitled 138 copy.docx
Untitled 138 copy 2.docx
Untitled 139.docx
Untitled 40 MOM ADDRESS.jpg
Untitled 242.doc
Untitled 243.doc
Untitled 243 IMPORTANT.doc
Untitled 41....



PROTIP: NEVER LOOK IN SOMEONE
ELSE'S DOCUMENTS FOLDER.

"FINAL".doc



FINAL.doc!



FINAL_rev.2.doc



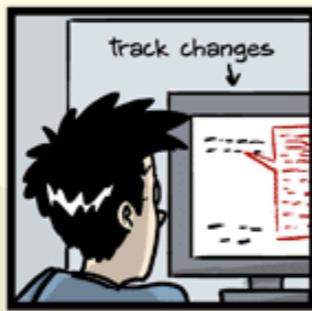
FINAL_rev.6.COMMENTS.doc



FINAL_rev.8.comments5.
CORRECTIONS.doc



JORGE CHAM © 2012



FINAL_rev.18.comments7.
corrections9.MORE.30.doc



FINAL_rev.22.comments49.
corrections.10.#@\$%WHYDID
ICOMETOGRAD SCHOOL????.doc

Name your files effectively

- ❖ Descriptive
- ❖ Consistent
- ❖ Human readable (easy to understand file content)
- ❖ Machine readable (no special characters)
- ❖ Plays well with default ordering

Name your files effectively

- ❖ ?
 - ❖ myabstract.docx
 - ❖ filenames with spaces.xlsx
 - ❖ figure 1.png
 - ❖ fig 2.png



Name your files effectively

- ❖ NO
 - ❖ myabstract.docx
 - ❖ filenames with spaces.xlsx
 - ❖ figure 1.png
 - ❖ fig 2.png
- ❖ YES
 - ❖ 2019-06-07_abstract-for-conf.docx
 - ❖ filenames_without_spaces.xlsx
 - ❖ fig01_animals-tested.png
 - ❖ fig02_specimen-types.png



LETTERS

Edited by Jennifer Sills

Editorial expression of concern

Back up your project

Minimum:

- ❖ **One local backup :**
 - external flash drive (USB),
external hard drive
- ❖ **One cloud/internet backup:**
 - Dropbox, Drive, Box

IN THE 3 June issue, *Science* published the Report “Environmentally relevant concentrations of microplastic particles influence larval fish ecology” by Oona M. Lönnstedt and Peter Eklöv (1). The authors have notified *Science* of the theft of the computer on which the raw data for the paper were stored. These data were not backed up on any other device nor deposited in an appropriate repository. *Science* is publishing this Editorial Expression of Concern to alert our readers to the fact that no further data can be made available, beyond those already presented in the paper and its supplement, to enable readers to understand, assess, reproduce, or extend the conclusions of the paper.

Jeremy Berg

Editor in Chief

REFERENCE

1. O.M. Lonnstedt, P. Eklov, *Science* **352**, 1213 (2016).

Published online 1 December
10.1126/science.aah6990

Backup your project

- ❖ Back up (almost) everything created by a human being (you!) as soon as it is created.
- ❖ Keep changes between backups small.
- ❖ Create, maintain, and use a checklist for saving and sharing changes to the project.
- ❖ Store each project in a folder that is mirrored off your own working machine.
- ❖ Copy the entire project whenever a significant change has been made.

References

- ❖ Wilson G, Bryan J, Cranston K, Kitzes J, Nederbragt L, et al. (2017) Good enough practices in scientific computing. PLOS Computational Biology 13(6): e1005510. <https://doi.org/10.1371/journal.pcbi.1005510>
- ❖ Naming files by Jenny Bryan
- ❖ Reproducible research: <http://guides.lib.berkeley.edu/c.php?g=652220&p=4575532>