How to
Organize
Your Files.
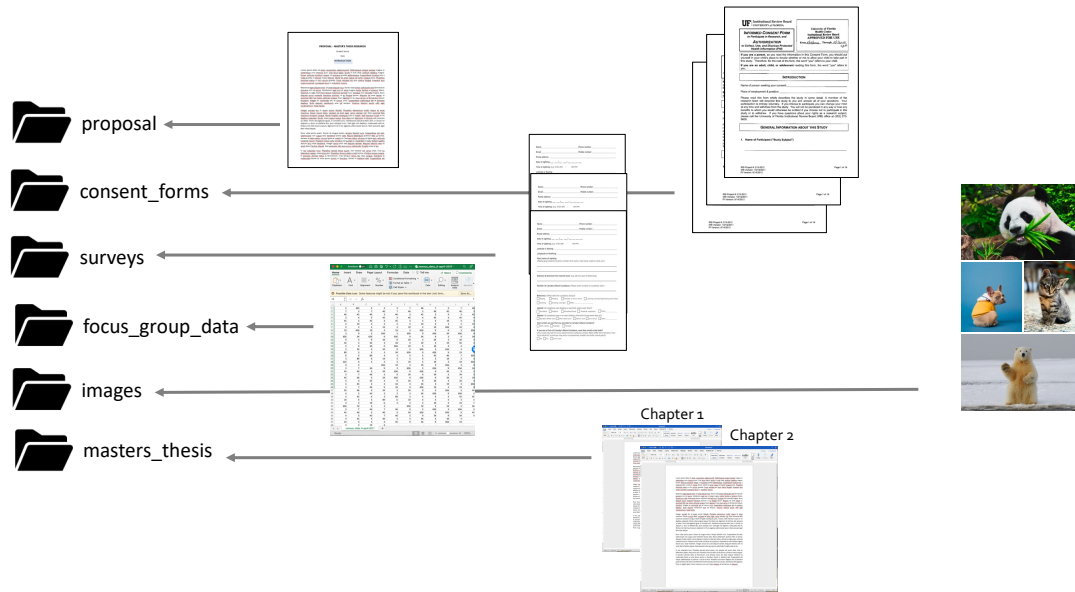
LAS 6292:
Data Collection & Management
Emilio M. Bruna
University of Florida

1



This is ~~you~~ all of us.

2

We usually organize our stuff by 'file type'

proposal
consent_forms
surveys
focus_group_data
images
masters_thesis

Chapter 1
Chapter 2

3

---

There is no "best" system to organize your files, but the one we will use is definitely better.
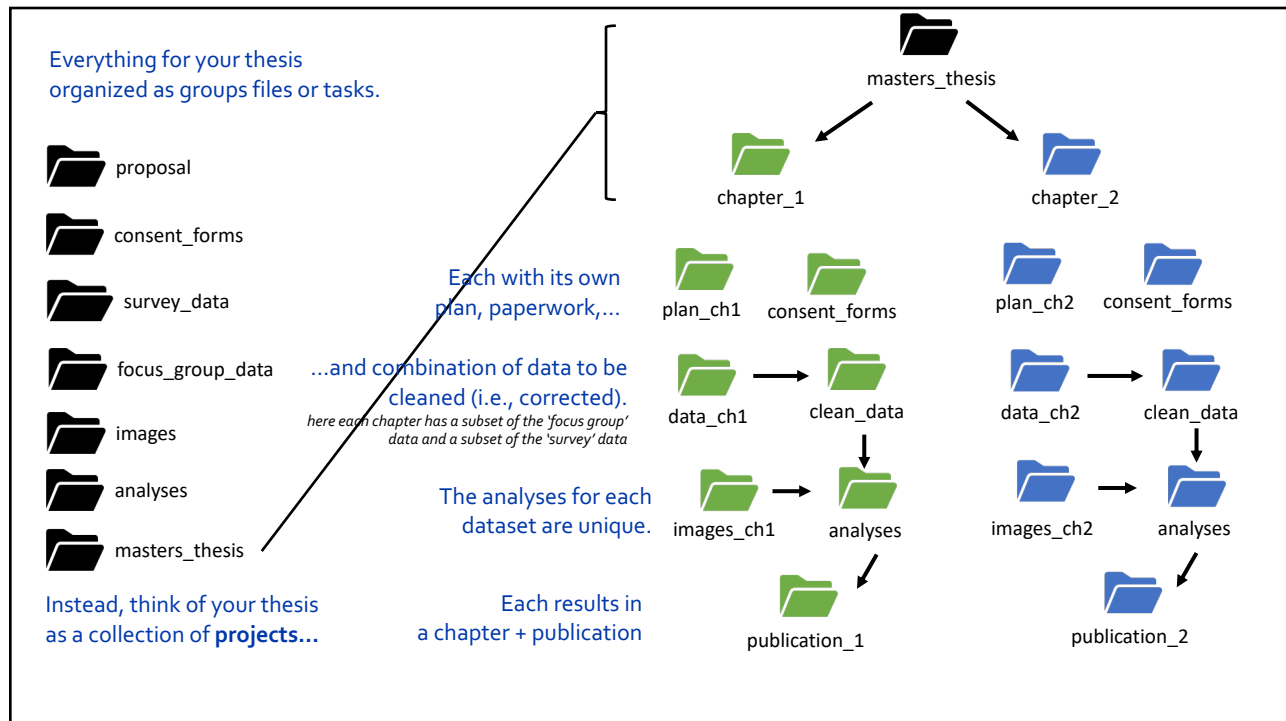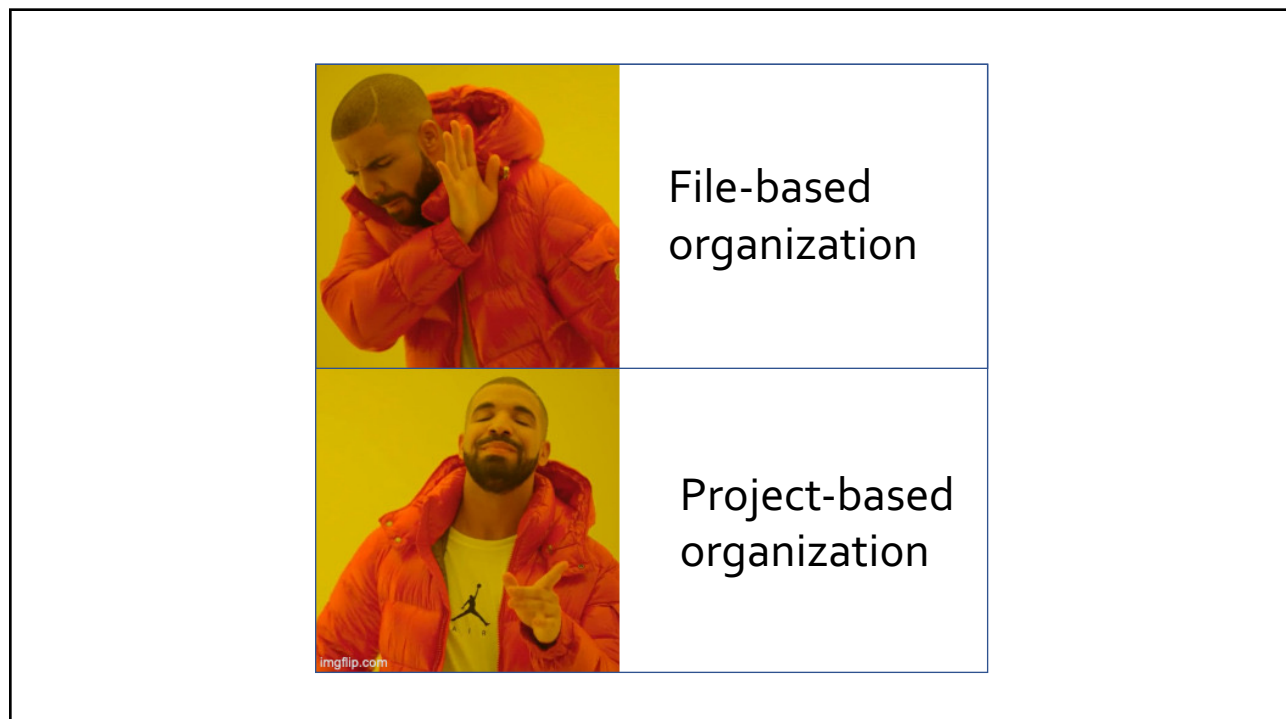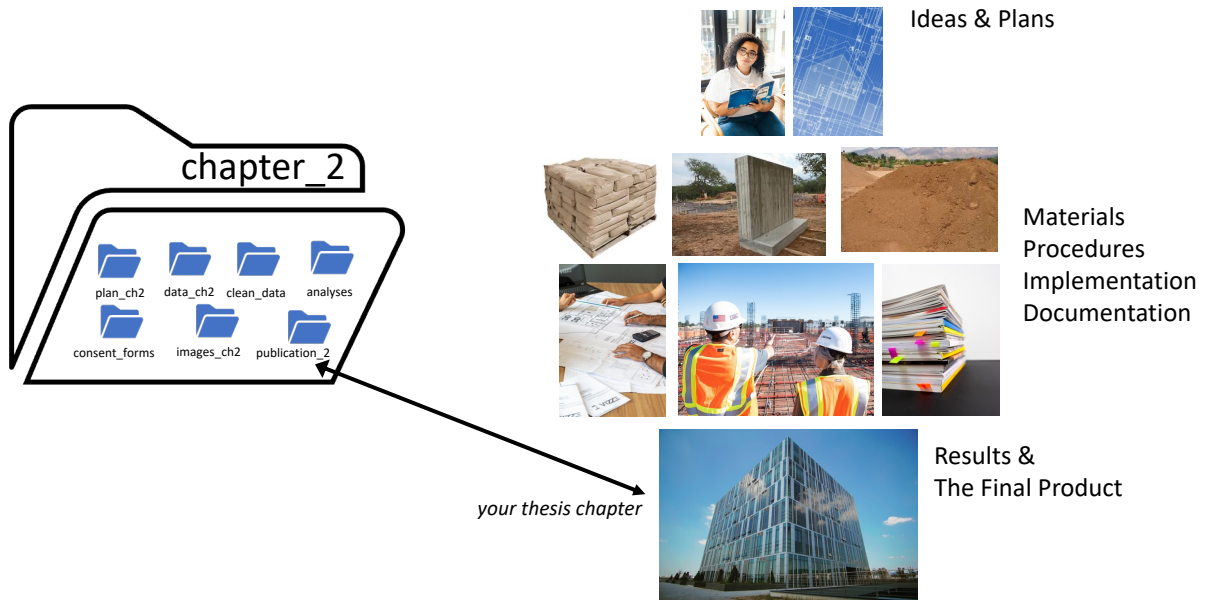
It is More Logical.

It is More Efficient.

4

"More Efficient" means less of this….

5



…and more of this.

6

Everything for your thesis organized as groups files or tasks.

- proposal
- consent_forms
- survey_data
- focus_group_data
- images
- analyses
- masters_thesis

Instead, think of your thesis as a collection of **projects**...

Each with its own plan, paperwork,...

...and combination of data to be cleaned (i.e., corrected).
*here each chapter has a subset of the 'focus group' data and a subset of the 'survey' data*

The analyses for each dataset are unique.

Each results in a chapter + publication

masters_thesis → chapter_1, chapter_2

chapter_1: plan_ch1, consent_forms, data_ch1 → clean_data → analyses → publication_1, images_ch1

chapter_2: plan_ch2, consent_forms, data_ch2 → clean_data → analyses → publication_2, images_ch2
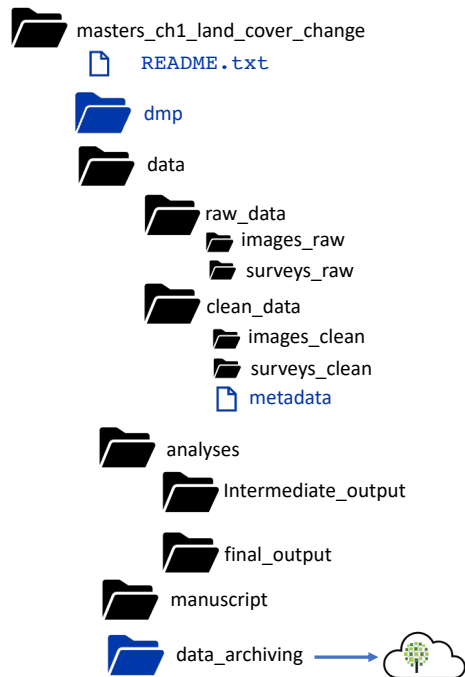
7



File-based organization

Project-based organization

8

## Project-based puts everything for a project – from start to finish - in a single, self-contained, and easy to share unit



Ideas & Plans

chapter_2

plan_ch2   data_ch2   clean_data   analyses

consent_forms   images_ch2   publication_2

Materials
Procedures
Implementation
Documentation

*your thesis chapter*

Results &
The Final Product

9

---

masters_ch1_land_cover_change
  README.txt
  dmp
  data
    raw_data
      images_raw
      surveys_raw
    clean_data
      images_clean
      surveys_clean
      metadata
  analyses
    Intermediate_output
    final_output
  manuscript
  data_archiving ➔ 🌳

## Additional Items to Include
(to be discussed in class)

10

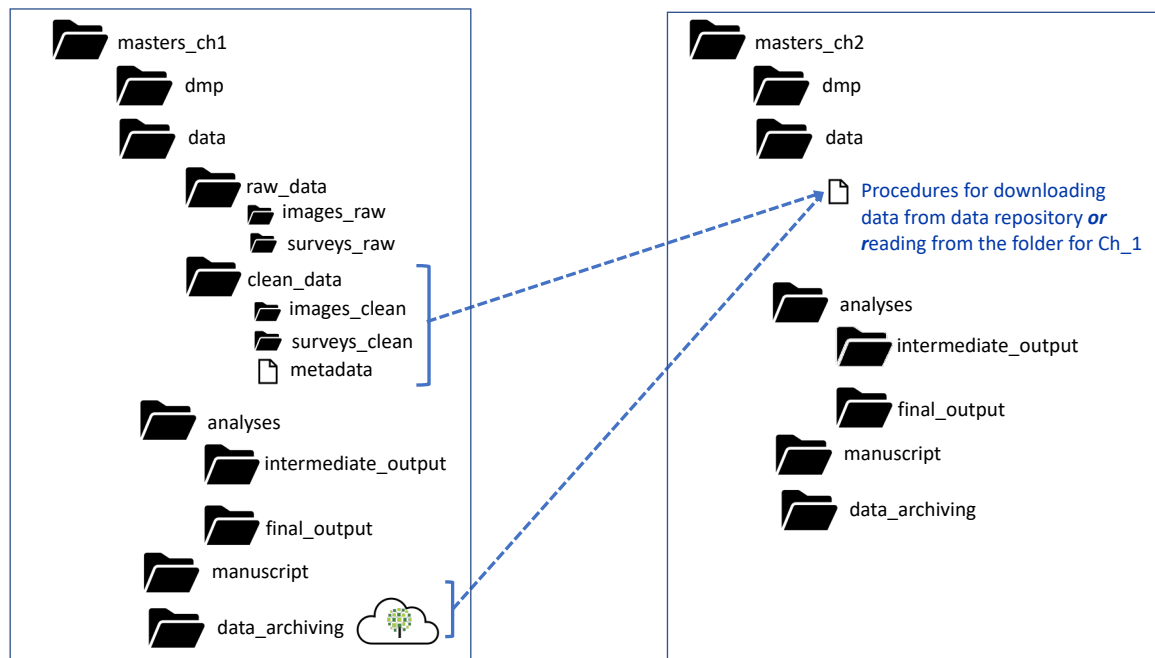BTW, the organization of physical files can (should) also be Project-Based

11



But what if some Data or Results from
Project 1 are being used in Project 2???

12

## No Problem.

- masters_ch1
  - dmp
  - data
    - raw_data
      - images_raw
      - surveys_raw
    - clean_data
      - images_clean
      - surveys_clean
      - metadata
  - analyses
    - intermediate_output
    - final_output
  - manuscript
  - data_archiving

- masters_ch2
  - dmp
  - data
    - Procedures for downloading data from data repository *or* *r*eading from the folder for Ch_1
  - analyses
    - intermediate_output
    - final_output
  - manuscript
  - data_archiving

13

# OK, let's get to work.



14

Fin.

Photos: Unsplash or EM Bruna
Music: www.bensound.com

15