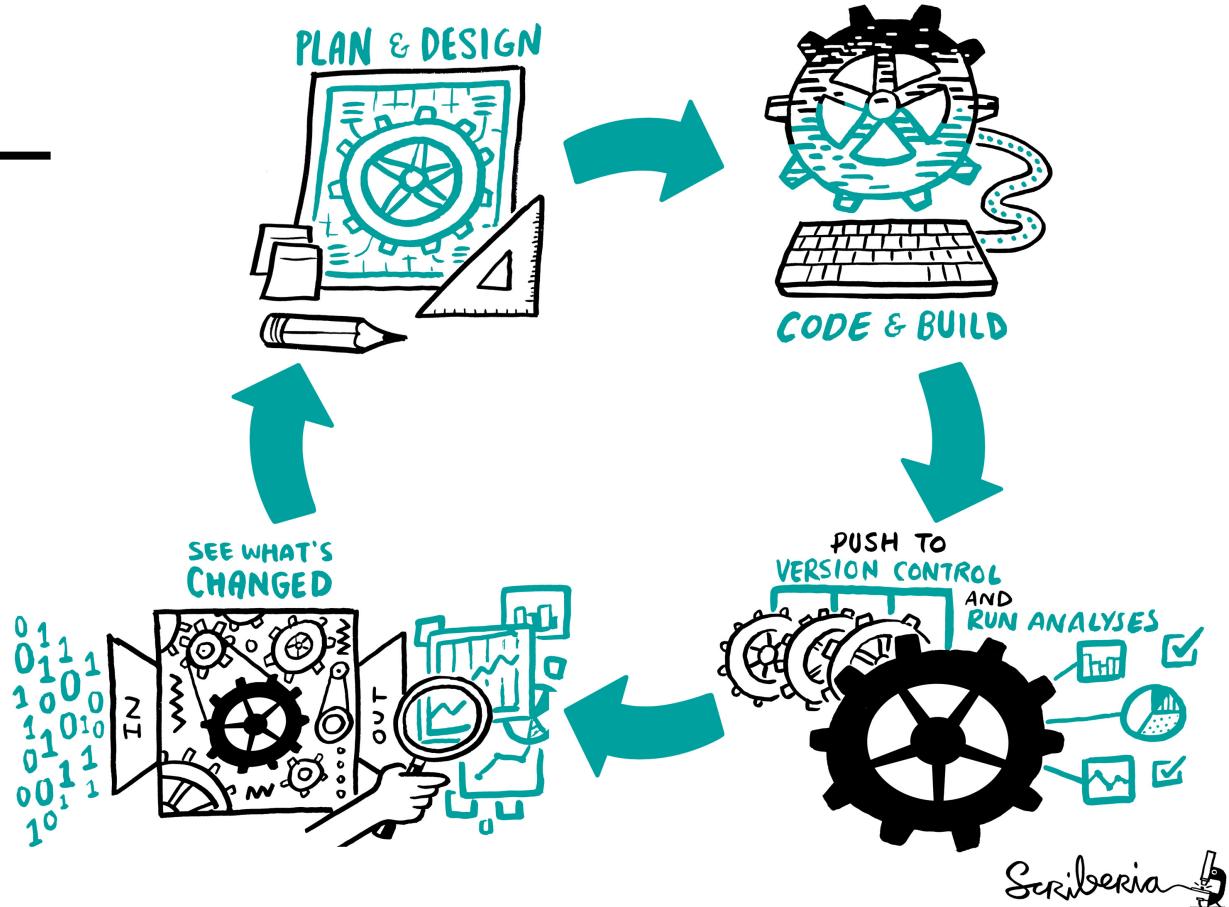


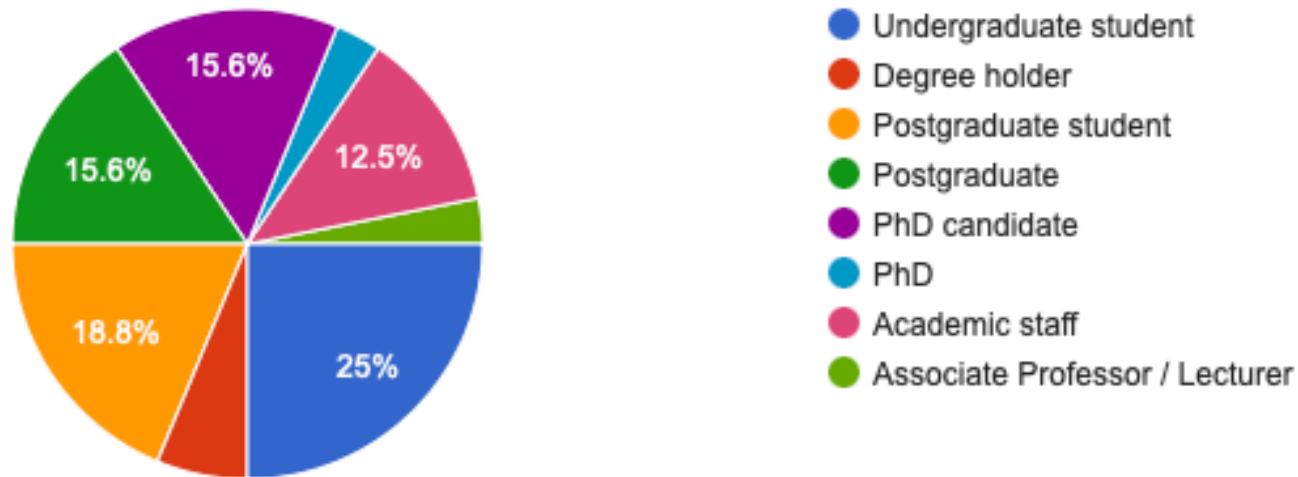
Information and data architecture in Ecology: Folder structure, metadata and project management through GitHub



Academic level

32 responses

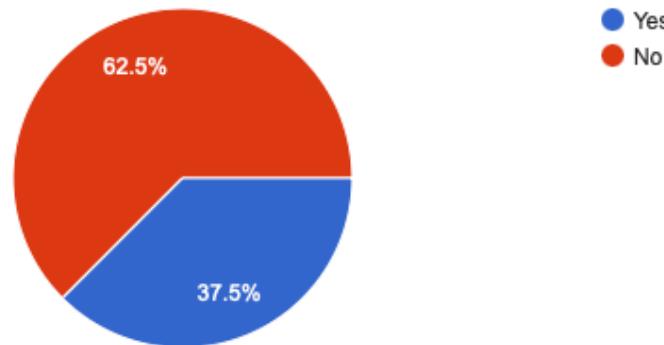
 Copy



Have you been taught how to organise your project (i.e. thesis) into collaborative and reproducible folders?

 Copy

32 responses



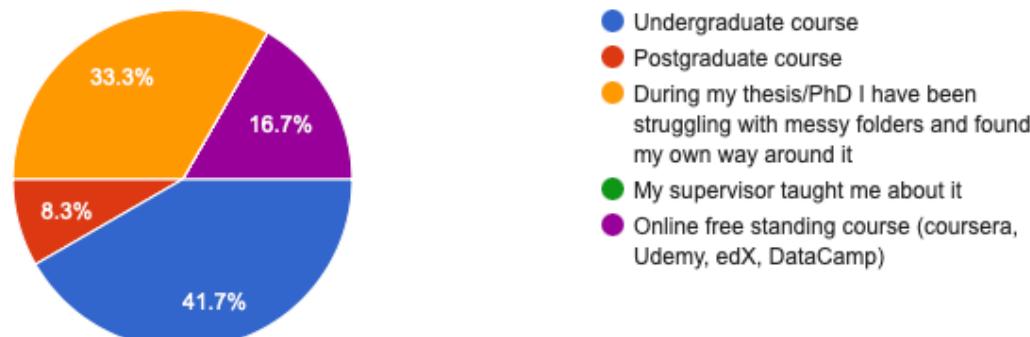
Yes

No

If you answered yes in the previous question, how did you learn about folder management?

 Copy

12 responses



Undergraduate course

Postgraduate course

During my thesis/PhD I have been
struggling with messy folders and found
my own way around it

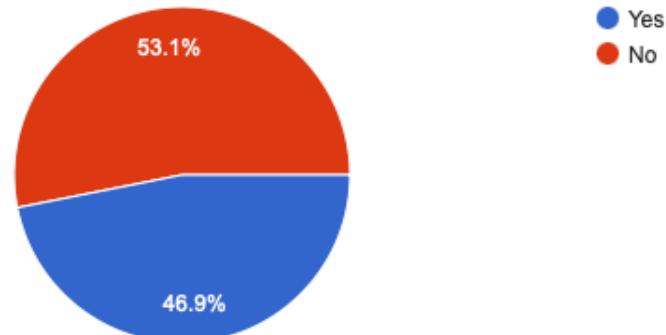
My supervisor taught me about it

Online free standing course (coursera,
Udemy, edX, DataCamp)

Have you been taught how to name files and folders in your PC?

 Copy

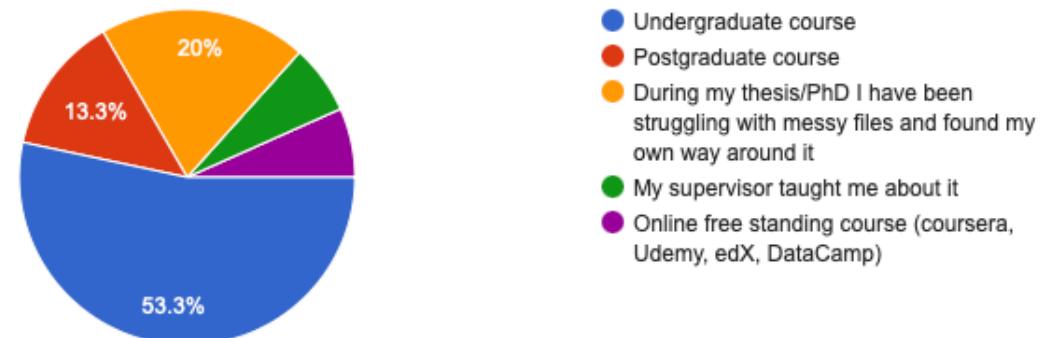
32 responses



If you answered yes in the previous question, how did you learn about file and folder onomastics?

 Copy

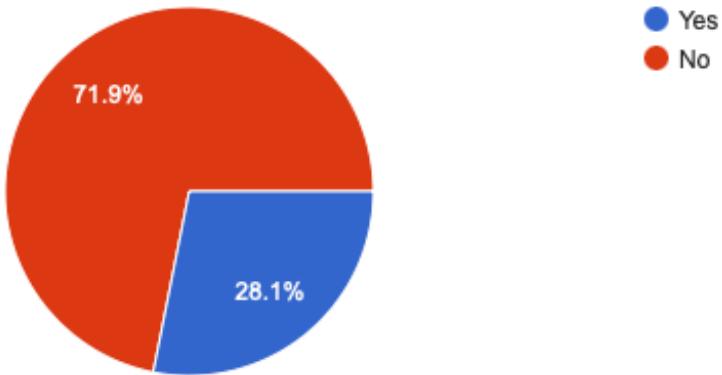
15 responses



Are you familiar with version control systems (git)?

 Copy

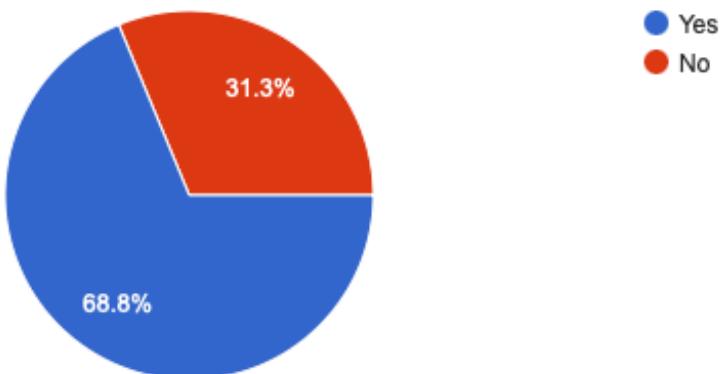
32 responses



Are you aware of GitHubs' existence?

 Copy

32 responses



Background

BSc in Environmental science

(Πανεπιστήμιο Αιγαίου) 2013-2018

- Toxicity effects of triclosan on *D. magna* and *L. minor*
(Danmarks Tekniske Universitet) Winter 2018
-

MSc in Biology: Ecology and Conservation

(Uppsala Universitet) 2019-2021

- Effects of coexistence between closely related orchids
on floral evolution of *G. conopsea* and *G. densiflora*
(Uppsala Universitet) summer 2020
 - Height-diameter allometry for a dominant palm to
improve understanding of carbon and forest dynamics
in forests of Puerto Rico
(Universidad de Puerto Rico & Uppsala Universitet) winter 2022
-

PhD candidate

(Universidad de Cadiz) spring 2022 - present

- Quantification of asymmetric dependancies in R
(Philipps-University Marburg) coming February 2024





Ophrys speculum

“God dammit Pas, I don't understand this. When I decided to do ecology, I thought I would be in the jungle or in a natural ecosystem collecting data or observing pandas. I want action not programming and statistics. Why to be here? Aren't pandas cool animals?”

Y.S., Uppsala 2019



```
a.R ✘ algaida_long ✘ algaida_species ✘ Untitled1* ✘ ALGAIADA_species_composition.R ✘ MESO ✘
+ Source on Save | Run | Source | Help | Addins |
```

```
1 library(nkpif)
2
3 read.csv("stats.csv")
4
```

Console Background Jobs

R 4.3.0 · ~/Documents/GitHub/SpatialNet/01_Algaida_natural_manipulation_experiment/

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

```
> library(nkpif)
Error in library(nkpif) : there is no package called 'nkpif'
> read.csv("stats.csv")
Error in file(file, "rt") : cannot open the connection
In addition: Warning message:
In file(file, "rt") :
  cannot open file 'stats.csv': No such file or directory
>
```

Files Plots Packages Help Viewer Presentation

R: Apply a function to each element of a vector | Find in Topic

map {purrr}

R Documentation

Apply a function to each element of a vector

Description

The map functions transform their input by applying a function to each element of a list or atomic vector and returning an object of the same length as the input.

- `map()` always returns a list. See the [`modify\(.f\)`](#) family for versions that return an object of the same type as the input.
- `map_lgl()`, `map_int()`, `map_db1()` and `map_chr()` return an atomic vector of the indicated type (or die trying). For these functions, `.f` must return a length-1 vector of the appropriate type.
- `map_vec()` simplifies to the common type of the output. It works with most types of simple vectors like Date, POSIXct, factors, etc.
- `walk()` calls `.f` for its side-effect and returns the input `.x`.

Usage

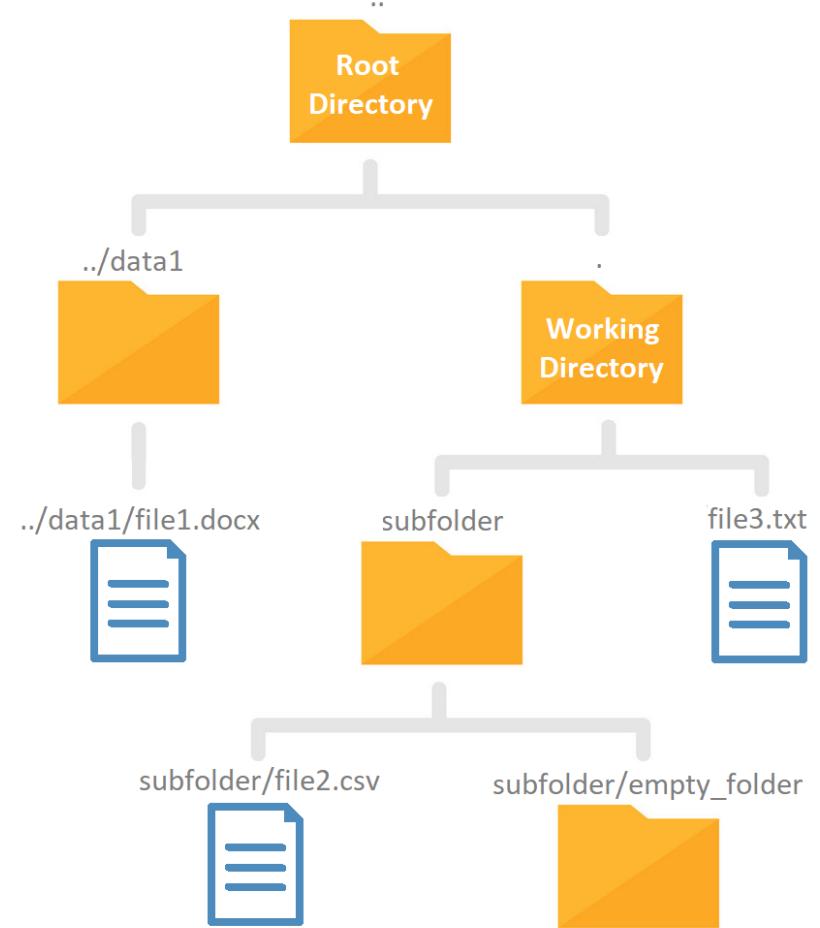
Still wondering:

Have we been taught:

- What is a working directory?
- Where is our working directory?
- What is reproducibility?
- Do we need more soft-skill training?

I studied environmental science, and my colleagues have a background in biology, ecology or life sciences:

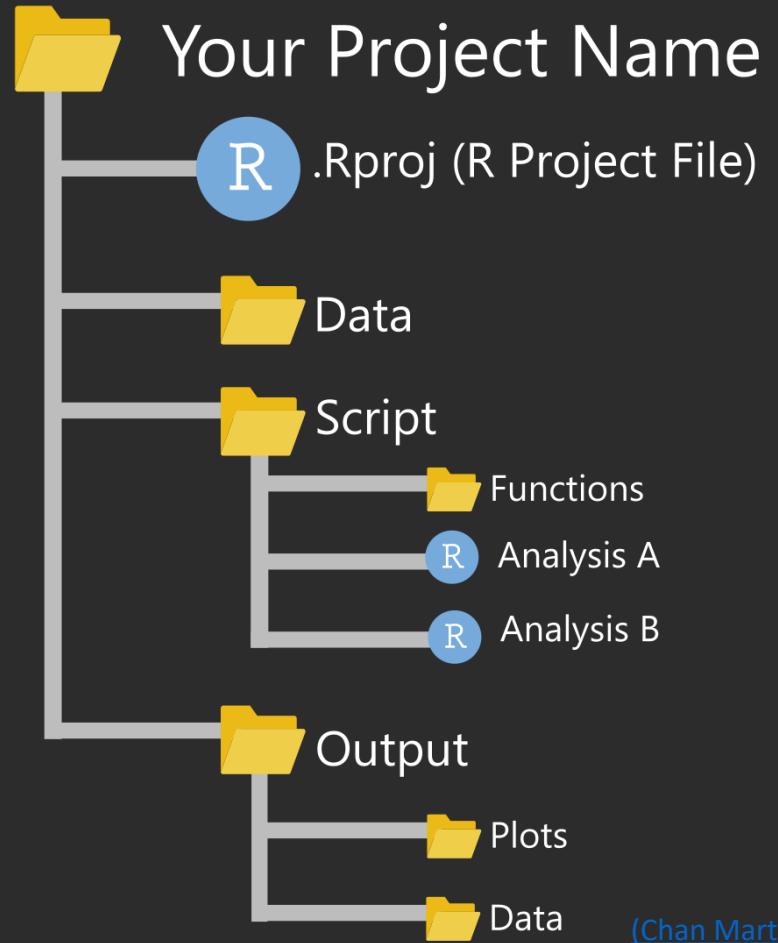
- Who is responsible to introduce us to concepts such as working directories, project management, file naming principles, open science?



Overview

- File systems
- Folder structure
- Fine naming principles
- Metadata
- Version Control Systems
- GitHub

A basic R project set up



Disclaimer

Working directory | *getwd()*, *setwd()*

.csv file → comma separated value

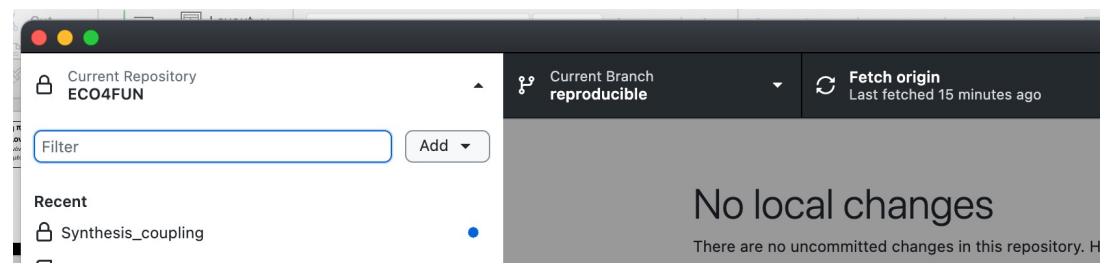
aov() model → Analyses Of Variance (aov)

lm() → linear model

Simplified googling “how to ...”

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

```
> library(nkpif)
Error in library(nkpif) : there is no package called 'nkpif'
> read.csv("stats.csv")
Error in file(file, "rt") : cannot open the connection
In addition: Warning message:
In file(file, "rt") :
  cannot open file 'stats.csv': No such file or directory
> getwd()
[1] "/Users/paschalis/Documents/GitHub/SpatialNet/01_Algaida_natural_manipulation_e
xperiment"
> |
```



Here is the most basic approach: Go to your unwanted blank page in your Word document, click as close to the bottom of the page as you can, and press your Backspace key until the page is removed. If that fails, try this: Click the "View" tab from the Ribbon and select (check) the Navigation Pane in the "Show" section. Feb 5, 2021





File systems

This ... > Documenta... > Search Documents

Quick access

OneDrive

This PC

3D Objects

Desktop

Documents

Downloads

Music

Pictures

Videos

Local Disk (C:)



Photos



Recipes



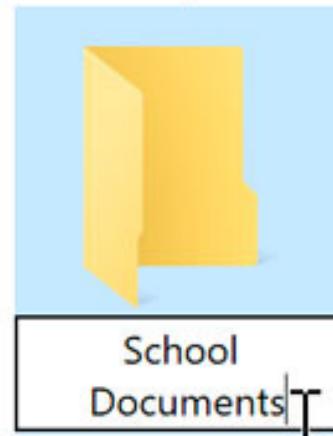
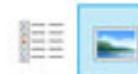
Work



Budget



Cover Letter

School
Documents

		< > ⌂ WorOptReproSciEcoEnvSci		☰ ▾	☶ ☰ ▾	↑	🔗	⋯ ▾	🔍	
		WorOptReproSciEcoEnvSci								
		Name	▲	Date Modified	Size	Kind				
⌄	poster			20 October 2023, 13:33	--	Folder				
		UOA_lecture_poster_distribute.jpg	✉	18 October 2023, 11:41	1.5 MB	JPEG image				
		UOA_lecture_poster_to_print.png		20 October 2023, 13:33	2.4 MB	PNG image				
⌄	presentation			Today, 14:16	--	Folder				
		20231026_UOA_data_presentation.pptx		Today, 14:16	93.4 MB	PowerP...(.pptx)				
⌄	presentation_handout			23 October 2023, 21:15	--	Folder				
		2023_presentation_handout_PCH.docx		23 October 2023, 20:58	1.6 MB	Micros...(.docx)				
		README.md		3 October 2023, 17:13	70 bytes	Markdo...ument				
⌄	raw_data			19 October 2023, 14:12	--	Folder				
		participating_experiments.xlsx		19 October 2023, 14:12	115 KB	Micros...k (.xlsx)				
		README.md		3 October 2023, 17:46	285 bytes	Markdo...ument				
		README.md		Today, 13:09	10 KB	Markdo...ument				
⌄	scripts			Today, 13:35	--	Folder				
		.participatin_sites_map.R		Today, 13:35	1 KB	Rez Source				
		README.md		3 October 2023, 17:13	142 bytes	Markdo...ument				
		WorOptReproSciEcoEnvSci.Rproj		Today, 13:23	205 bytes	R Project				
OneDrive > WorOptReproSciEcoEnvSci										
16 items, 113.32 GB available										

```
paschalis ~ % brew help [COMMAND]
man brew
https://docs.brew.sh
paschalis@192 ~ % brew install tree

---> Downloading https://ghcr.io/v2/homebrew/core/tree/manifests/2.1.1
#####
---> Fetching tree
---> Downloading https://ghcr.io/v2/homebrew/core/tree/blobs/sha256:f4da9aeecc7912
#####
---> Pouring tree--2.1.1.arm64_sonoma.bottle.tar.gz
---> /opt/homebrew/Cellar/tree/2.1.1: 8 files, 177.2KB
---> Running 'brew cleanup tree' ...
Disable this behaviour by setting HOMEBREW_NO_INSTALL_CLEANUP.
Hide these hints with HOMEBREW_NO_ENV_HINTS (see 'man brew').
paschalis@192 ~ % tree /Users/paschalis/Library/CloudStorage/OneDrive-uca.es/WorOptReproSciEcoEnvSci
/Users/paschalis/Library/CloudStorage/OneDrive-uca.es/WorOptReproSciEcoEnvSci
├── README.md
├── WorOptReproSciEcoEnvSci.Rproj
├── poster
│   ├── UOA_lecture_poster_distribute.jpg
│   └── UOA_lecture_poster_to_print.png
├── presentation
│   └── 20231026_UOA_data_presentation.pptx
├── presentation_handout
│   └── 2023_presentation_handout_PCH.docx
└── README.md

6 directories, 11 files
paschalis@192 ~ %
```

WorOptReproSciEcoEnvSci

Name	Date Modified
poster	20 October
UOA_lecture_poster_distribute.jpg	18 October
UOA_lecture_poster_to_print.png	20 October
presentation	Today, 14:1
20231026_UOA_data_presentation.pptx	Today, 14:1
presentation_handout	23 October
2023_presentation_handout_PCH.docx	23 October
README.md	3 October
raw_data	19 October
participating_experiments.xlsx	19 October
README.md	3 October
README.md	Today, 13:0
scripts	Today, 13:3
participatin_sites_map.R	Today, 13:3
README.md	3 October
WorOptReproSciEcoEnvSci.Rproj	Today, 13:2

OneDrive > WorOptReproSciEcoEnvSci

16 items, 113.32 GB available

brew help [COMMAND]
man brew
https://docs.brew.sh
paschalis@192 ~ % brew install tree

==> Downloading https://ghcr.io/v2/homebrew/core/tree/manifests/2.1.1
100.0%
==> Fetching tree
==> Downloading https://ghcr.io/v2/homebrew/core/tree/blobs/sha256:f4da9aec7912
100.0%
==> Pouring tree--2.1.1.arm64_sonoma.bottle.tar.gz
/opt/homebrew/Cellar/tree/2.1.1: 8 files, 177.2KB
==> Running `brew cleanup tree`...
Disable this behaviour by setting HOMEBREW_NO_INSTALL_CLEANUP.
Hide these hints with HOMEBREW_NO_ENV_HINTS (see `man brew`).
paschalis@192 ~ % tree /Users/paschalis/Library/CloudStorage/OneDrive-uca.es/WorOptReproSciEcoEnvSci

```
/Users/paschalis/Library/CloudStorage/OneDrive-uca.es/WorOptReproSciEcoEnvSci
├── README.md
├── WorOptReproSciEcoEnvSci.Rproj
└── poster
    ├── UOA_lecture_poster_distribute.jpg
    └── UOA_lecture_poster_to_print.png
└── presentation
    └── 20231026_UOA_data_presentation.pptx
└── presentation_handout
    ├── 2023_presentation_handout_PCH.docx
    └── README.md
└── raw_data
    ├── README.md
    └── participating_experiments.xlsx
└── scripts
    ├── README.md
    └── participatin_sites_map.R
```

6 directories, 11 files

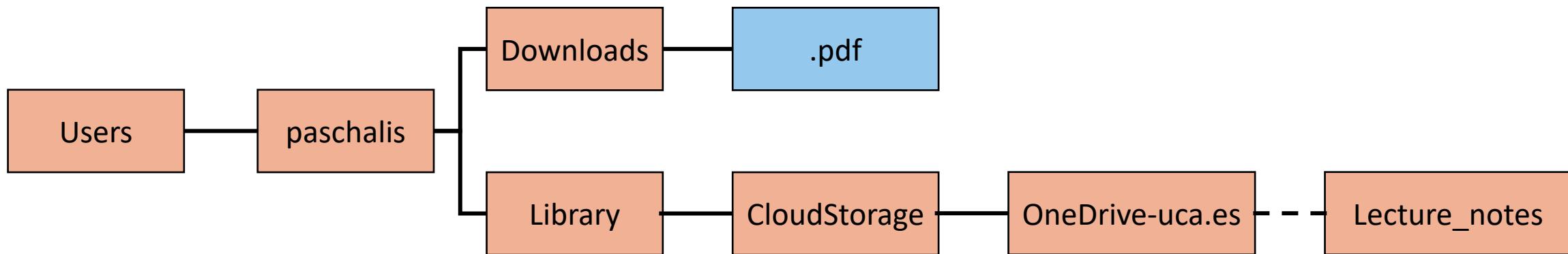
Folder = directory

What common people call a ‘computer folder’ is a directory

File paths

Have you ever wondered where your computer stores data?

- /Users/paschalis/Downloads/VolunteersSelection.pdf
- /Users/paschalis/Library/CloudStorage/OneDrive-uca.es/WorOptReproSciEcoEnvSci/lecture_notes



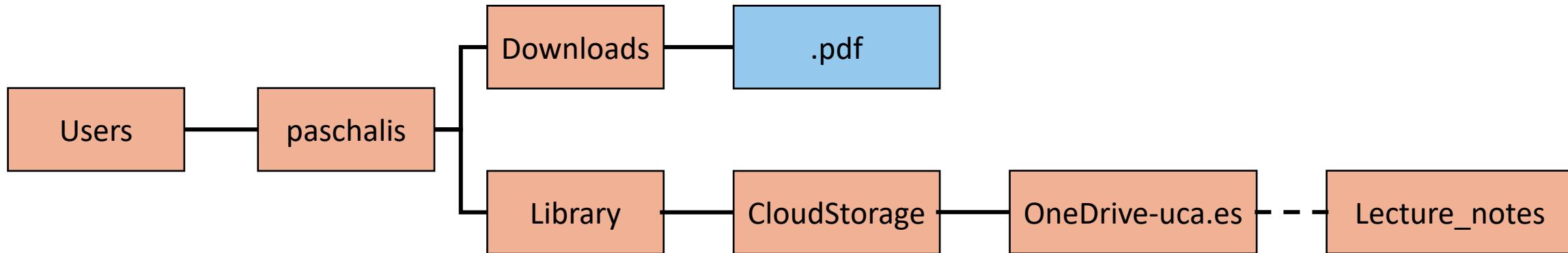
**Different file structure in different operating systems,
backslash in windows (\) and forward slash in UNIX (/)**

File paths

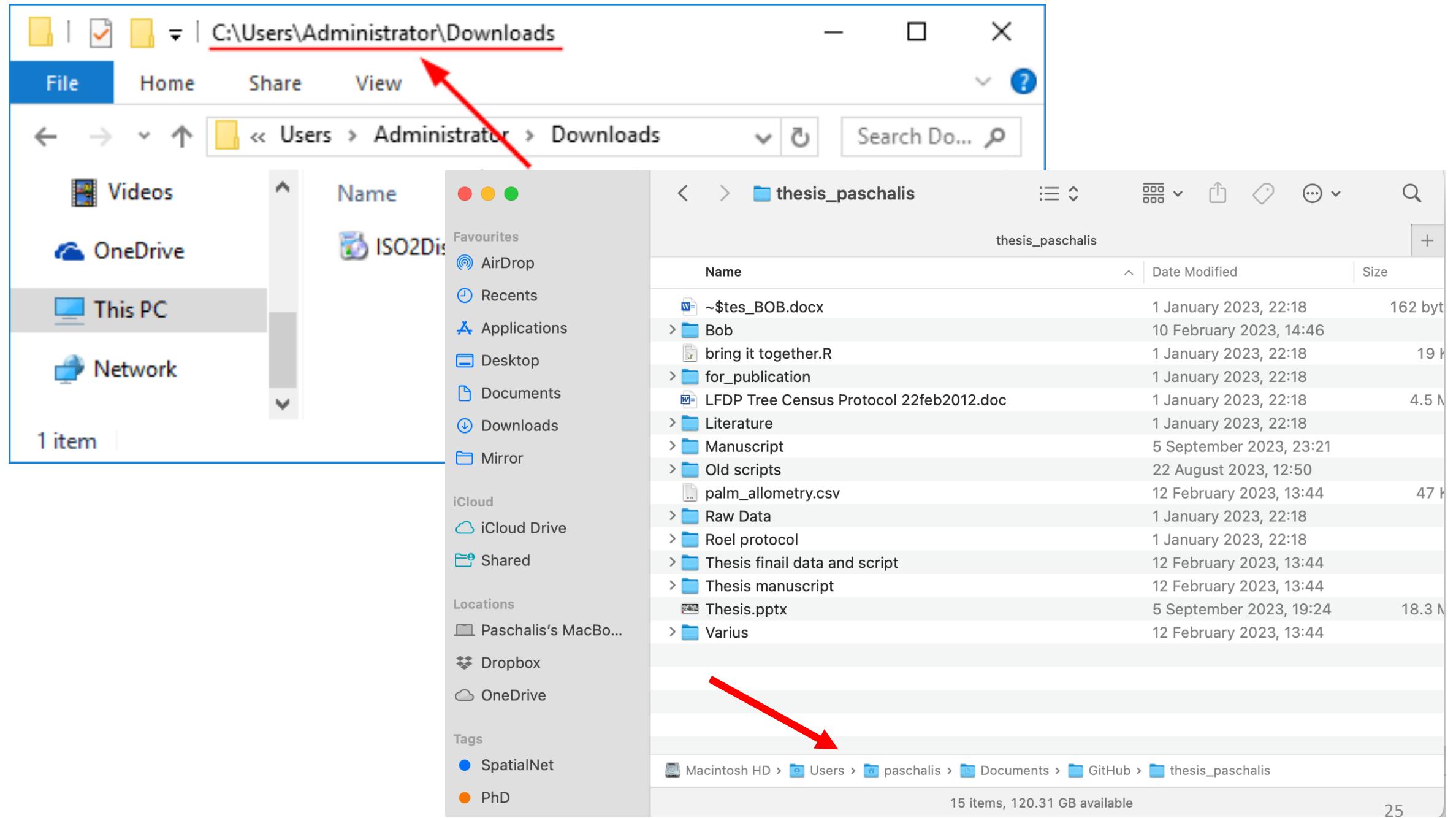
```
> getwd()
```

```
[1] "/Users/pas161/Documents/GitHub/SpatialNet/code_raul"
```

```
>
```



**Different file structure in different operating systems,
backslash in windows (\) and forward slash in UNIX (/)**



Take-home message:

An absolute path is the location of the file we look at in the hard drive:

/Users/paschalis/Library/CloudStorage/OneDrive-uca.es/WorOptReproSciEcoEnvSci/lecture_notes

- Each slash, represents a different folder
- One day someone simplified directories and named them ‘folders’
 - Directory = folder
- When we double click into a folder our computer executes this path to reach the file
- The absolute path is different for each computer:
/Users/Nico/documents/data.csv
- Different file paths for different operating systems:
 - Windows : C:\user\documents
 - Macintosh: ~/user/documents

How do we collaborate, while exchanging code scripts, between different users and different operating systems?

The screenshot shows the RStudio interface with the following details:

- Top Bar:** Shows multiple open files: Vero2.R, multiple, resha, script_for_publication_RM-2.R, Xu_code.R, and Untitled1*. The tabs for script_for_publication_RM-2.R and Untitled1* are highlighted.
- Toolbar:** Includes icons for file operations like Open, Save, and Print, along with "Source on Save" and search/filter tools.
- Code Editor:** Displays the following R code:

```
1 # Set WD:  
2 getwd()  
3 setwd("C:/Users/.../D")
```
- Console Tab:** Active tab, showing the R startup message and a command history:

```
R version 4.3.2 (2023-10-31) -- "Eye Holes"  
Copyright (C) 2023 The R Foundation for Statistical Computing  
Platform: aarch64-apple-darwin20 (64-bit)  
  
R is free software and comes with ABSOLUTELY NO WARRANTY.  
You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.  
  
Natural language support but running in an English locale  
  
R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.  
  
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.  
  
> # Set WD:  
> setwd()  
Error in setwd() : argument "dir" is missing, with no default  
> # Set WD:  
> getwd()  
[1] "/Users/pas161/Documents/GitHub/SpatialNet/code_raul"  
>
```

From absolute to relative paths

In almost every STEM science, everything starts by loading the raw data:

- `read_csv("file_path")`

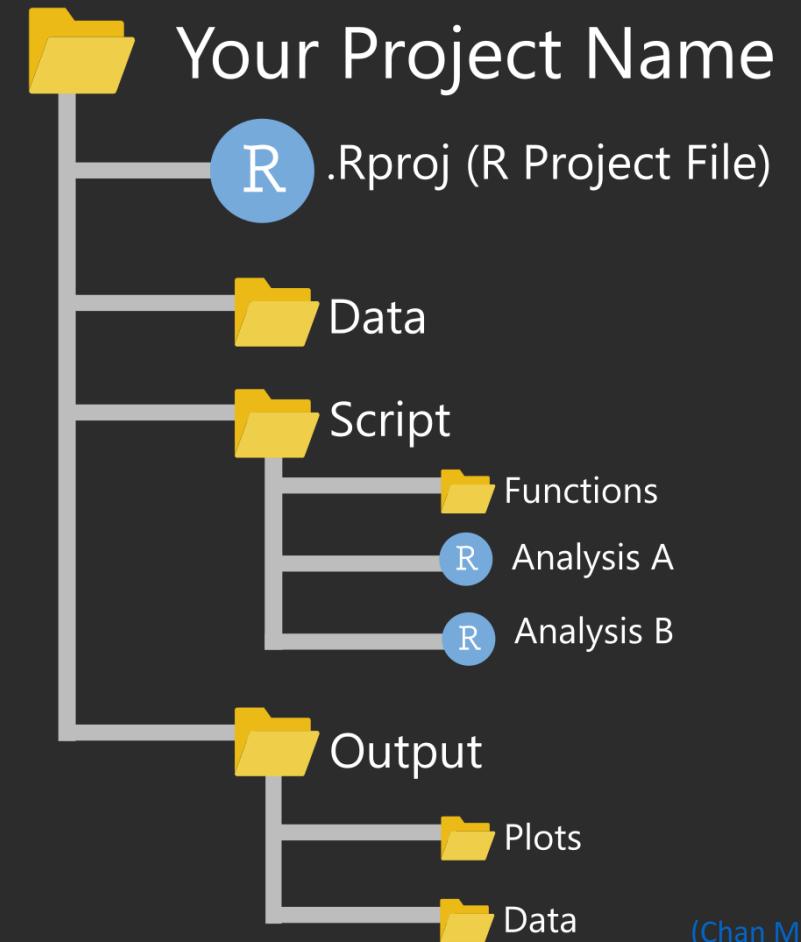
R projects simplify path names and enhance reproducibility

- Under an R-project can be organized many R-scripts:
 - `Analysis_A.r`
 - `Analysis_B.r`

Save the R-project file to the higher working directory.

- From now on, you won't need absolute but relative paths
 - Handy for those working or collaborating remotely and exchange R-scripts

A basic R project set up



(Chan Martin)

Paths and other demons

Absolute path:

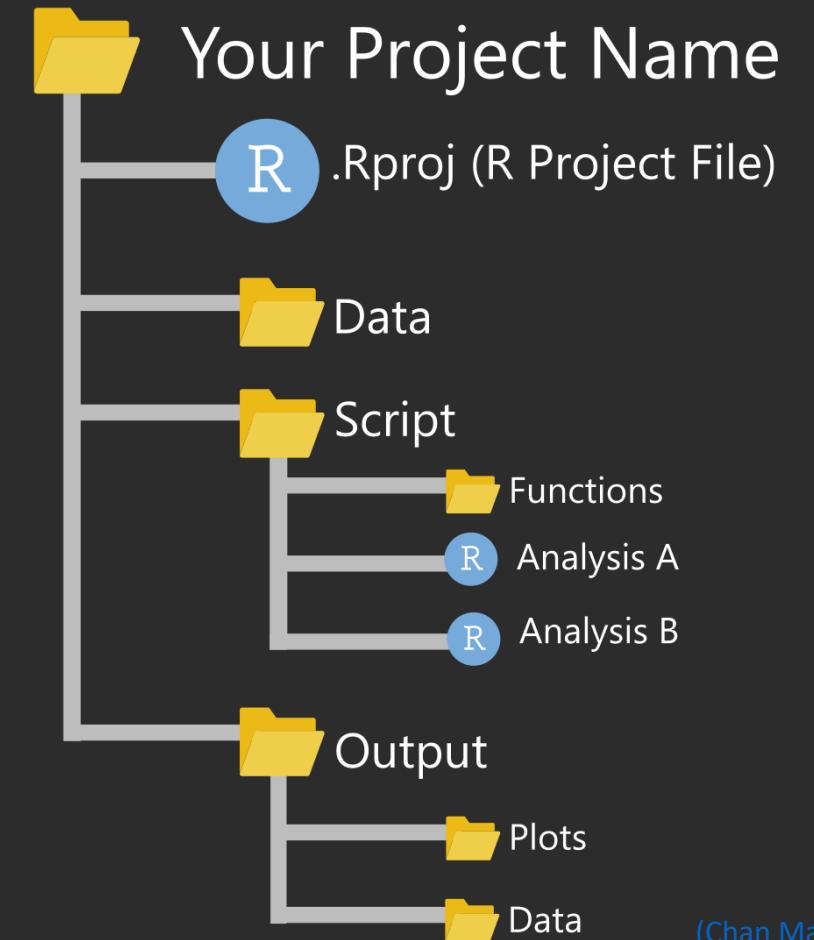
- user/documents/your_project_name/script/Analysis_A

Relative path:

- /script/Analysis_A

Relative to the location of
the .Rproj

A basic R project set up



For example

R project

The screenshot shows an RStudio interface with the following components:

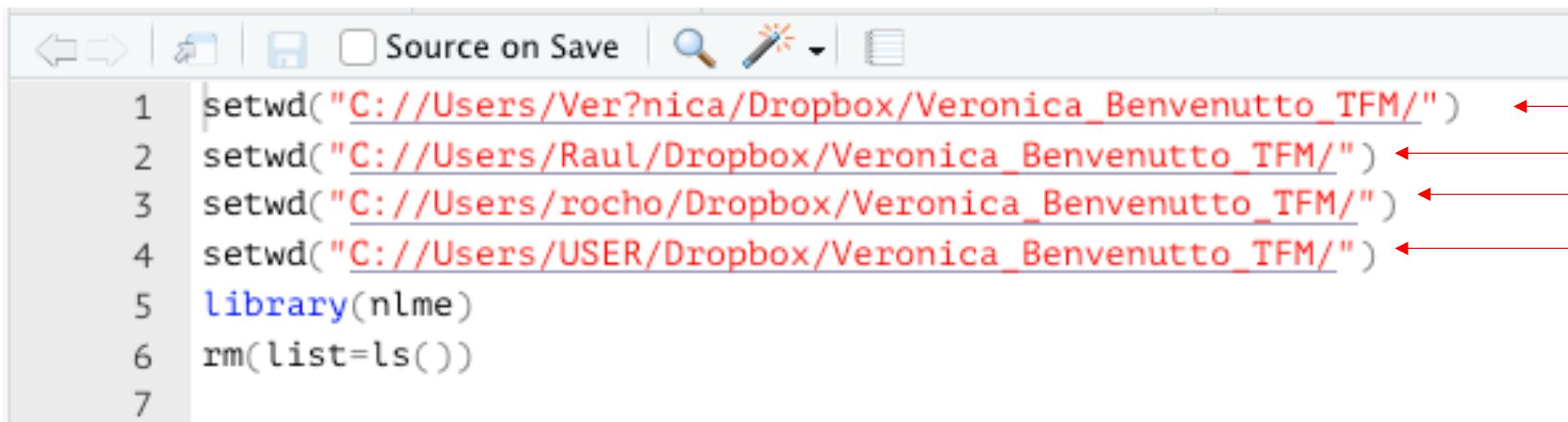
- Code Editor:** Displays the R script `participatin_sites_map.R`. Several lines of code are highlighted with red circles:
 - Line 10: `absolute_path_KTH <- "/user/theodorou/documents/WorOptReproSciEcoEnvSci/raw_data"`
 - Line 12: `absolute_path_PCH <- "~/Library/CloudStorage/OneDrive-uca.es/WorOptReproSciEcoEnvSci/raw_data"`
 - Line 15: `file_path <- "raw_data/participating_experiments.xlsx"`
- Console:** Shows the R command prompt and the resulting output. The output includes:
 - A warning message about relative paths.
 - An absolute folder path: `/Library/CloudStorage/OneDrive-uca.es/WorOptReproSciEcoEnvSci/raw_data/participating_experiments.xlsx`.
 - A relative path: `raw_data/participating_experiments.xlsx`.
- File Browser:** Shows the directory structure under the R project root `WorOptReproSciEcoEnvSci`. The contents of the `raw_data` folder are listed:

Name	Size	Modified
participating_experiments.xlsx	112.2 KB	Sep 27, 2023, 3:03 PM
README.md	285 B	Oct 3, 2023, 5:46 PM
- Project Bar:** Located at the top right, it shows the project name `WorOptReproSciEcoEnvSci` with a red circle around it.
- Bottom Navigation:** Includes tabs for Environment, History, Connections, Git, Tutorial, and a Git commit history.

Annotations:

- A yellow box highlights the text: "An absolute folder path describes the absolute location of the file relatively to the hard-drive".
- A yellow box highlights the text: "On the other hand, relative path is relative to the location of the Rproj file".

Instead of having one working directory for each computer...



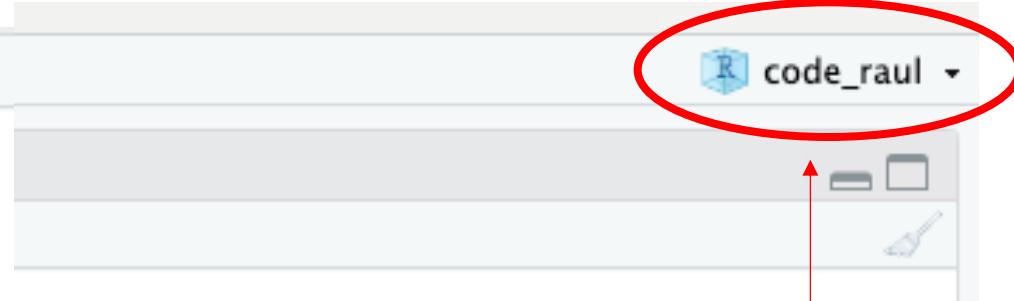
```
1 setwd("C://Users/Veronica/Dropbox/Veronica_Benvenutto_TFM/")
2 setwd("C://Users/Raul/Dropbox/Veronica_Benvenutto_TFM/")
3 setwd("C://Users/rocho/Dropbox/Veronica_Benvenutto_TFM/")
4 setwd("C://Users/USER/Dropbox/Veronica_Benvenutto_TFM/")
5 library(nlme)
6 rm(list=ls())
7
```

Veronicas computer
Raul's computer at home
Raul's computer at IVAGRO
Cloud directory

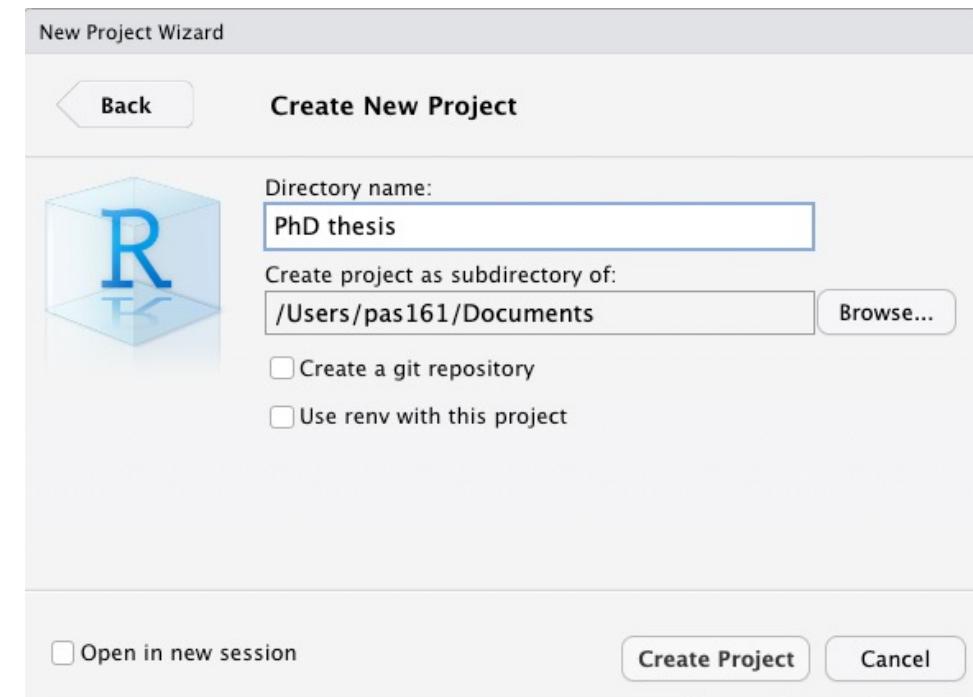
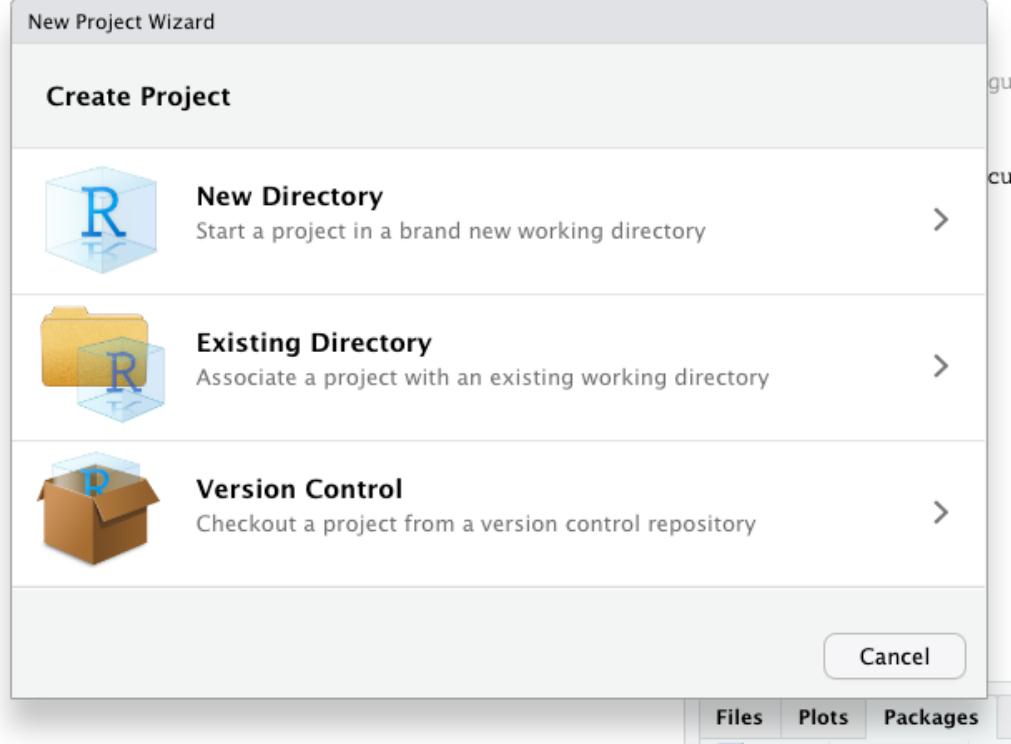
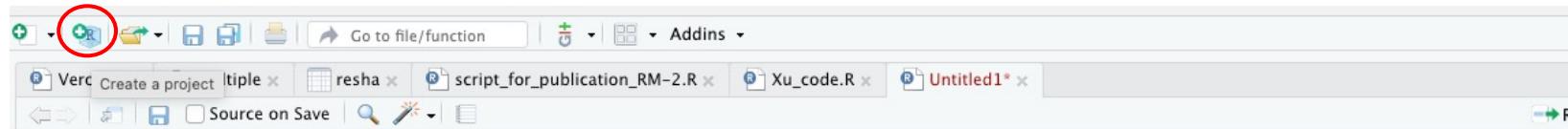
... we can use .Rproj and work with relative paths

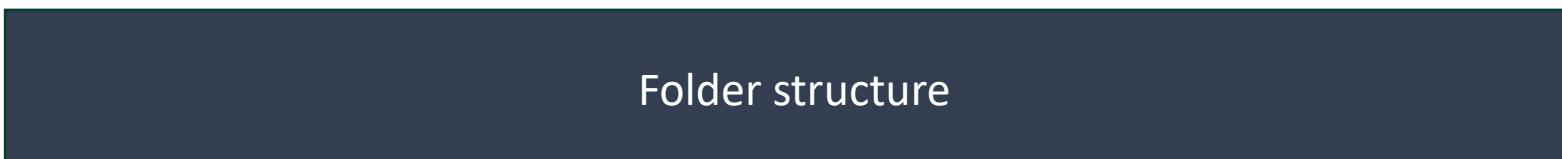


```
1 # Load packages
2 library(nlme)
3 library(tidyverse)
4
5 # Load data
6 datos <- read_csv("veronica/Activ_Enz_Regajal.csv")
7 datos <- as.data.frame(datos)
8
```

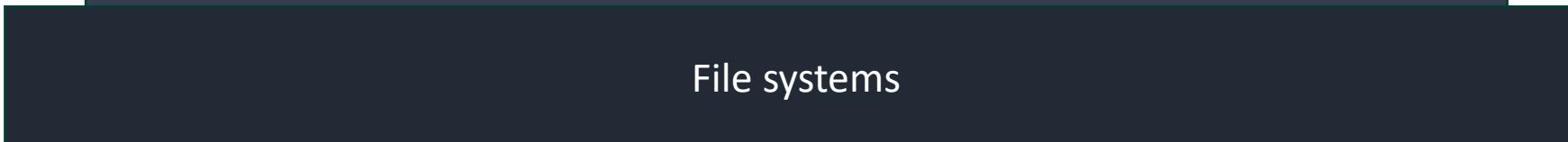


Any user on the planet can access it,
if has access to the Rproj





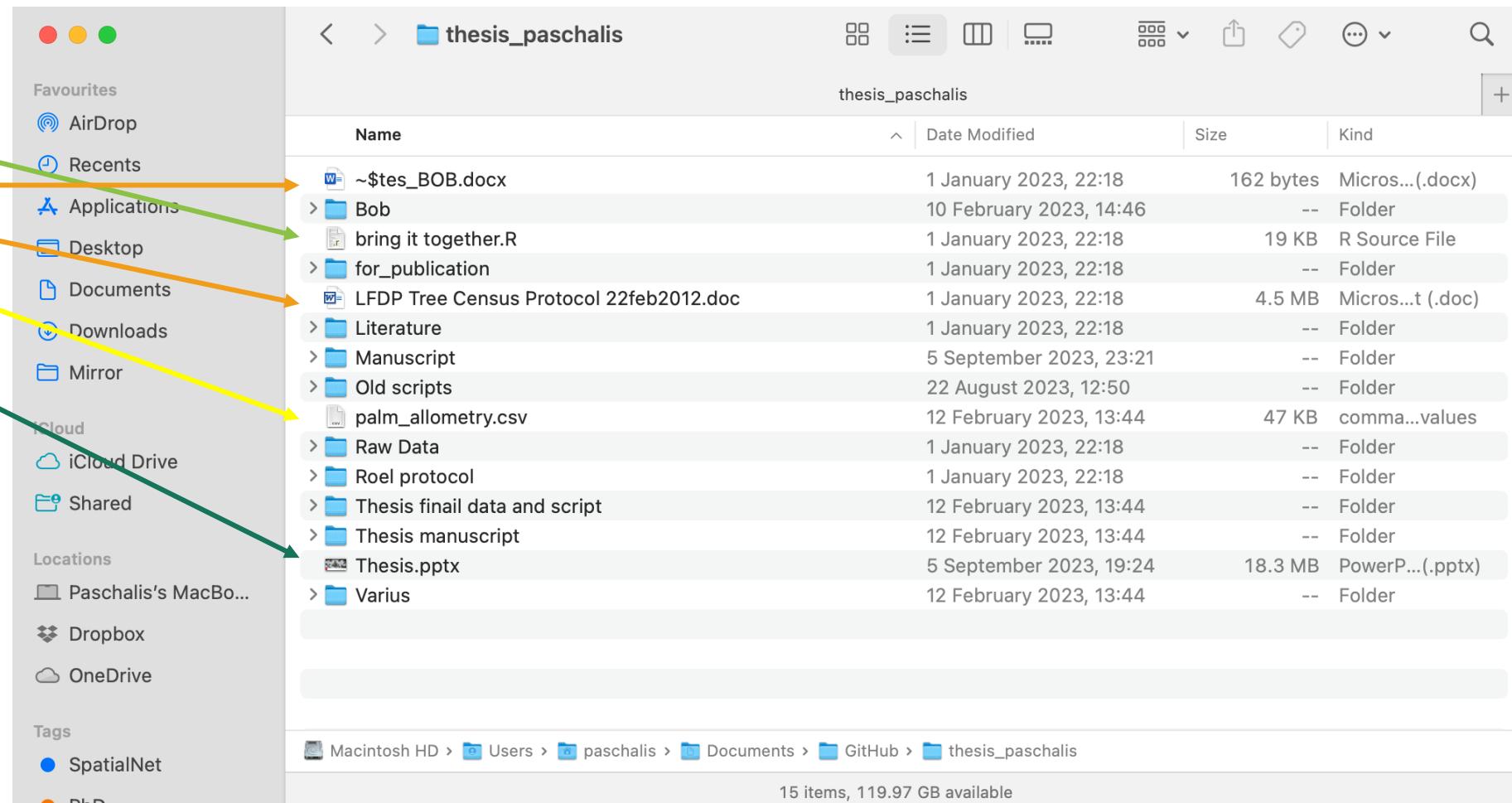
Folder structure



File systems

Ἐν ἀρχῇ ἦν τὸ Χάος (first there was chaos)

- R code scripts
- .doc files
- Spreadsheet
- Presentation
- Various folders of unknown content

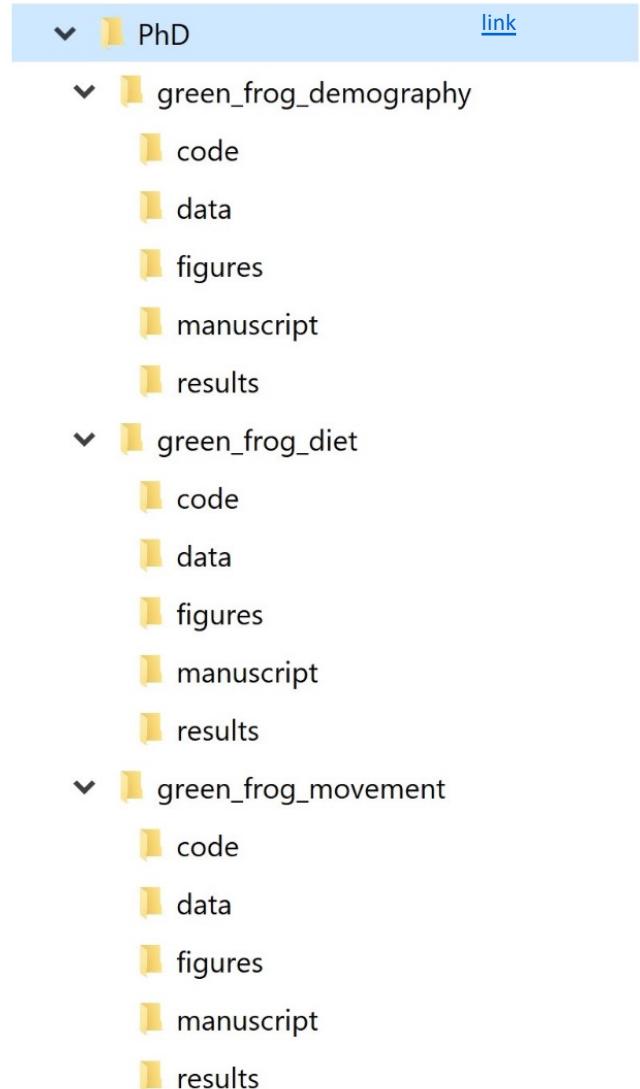


Basic principles of folder structure

- Organize directories under a main topic/project
- The name of each sub-folder (sub-directory) should indicate its content
- Organize folder based on the type of stored file:
 - */code*
 - */raw_data*
 - */clean_data*
- Store script outputs in a separate folder:
 - */output/figures*
- Always include a *readme.md* file

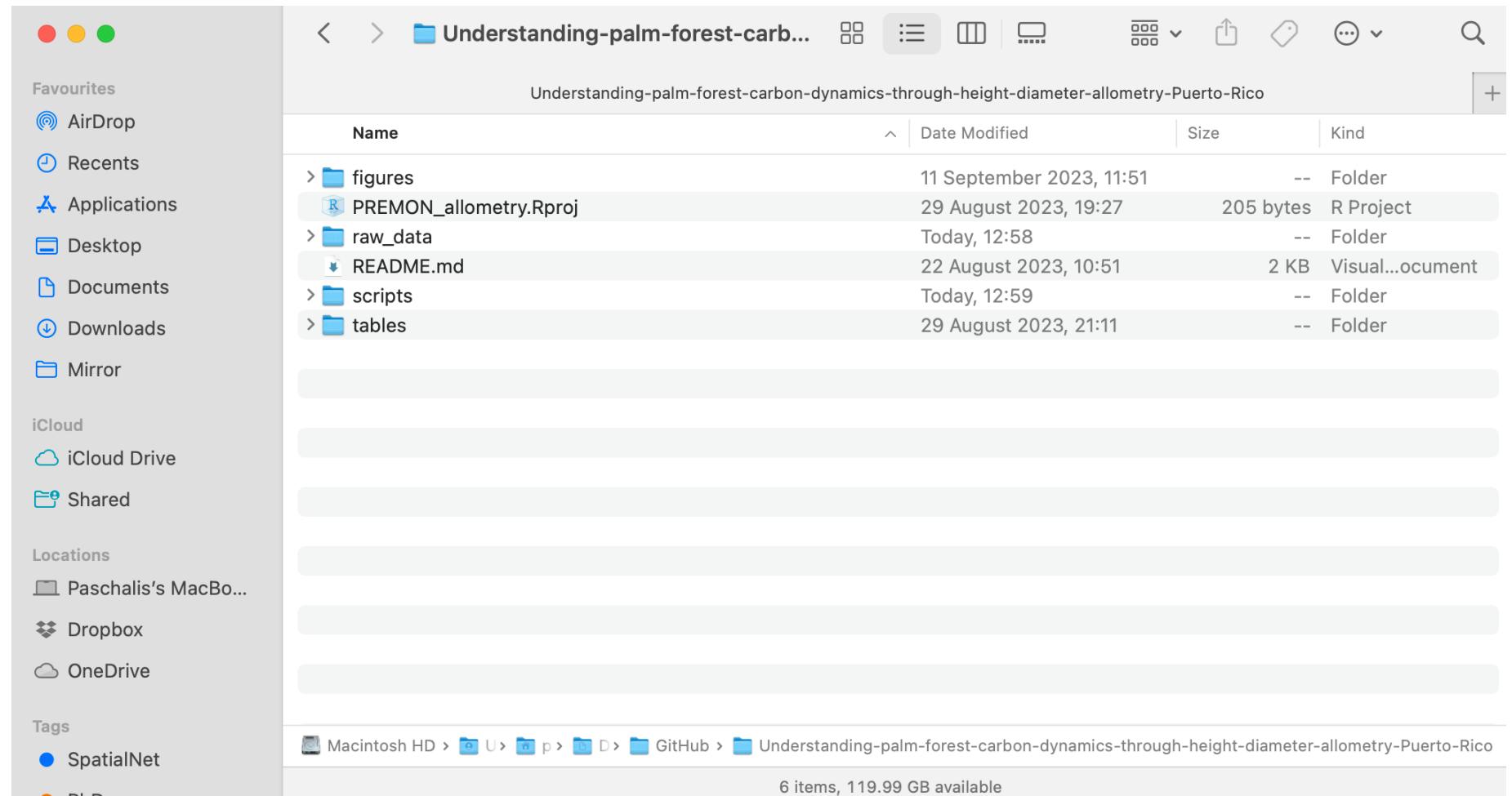
Different folders for raw and clean data

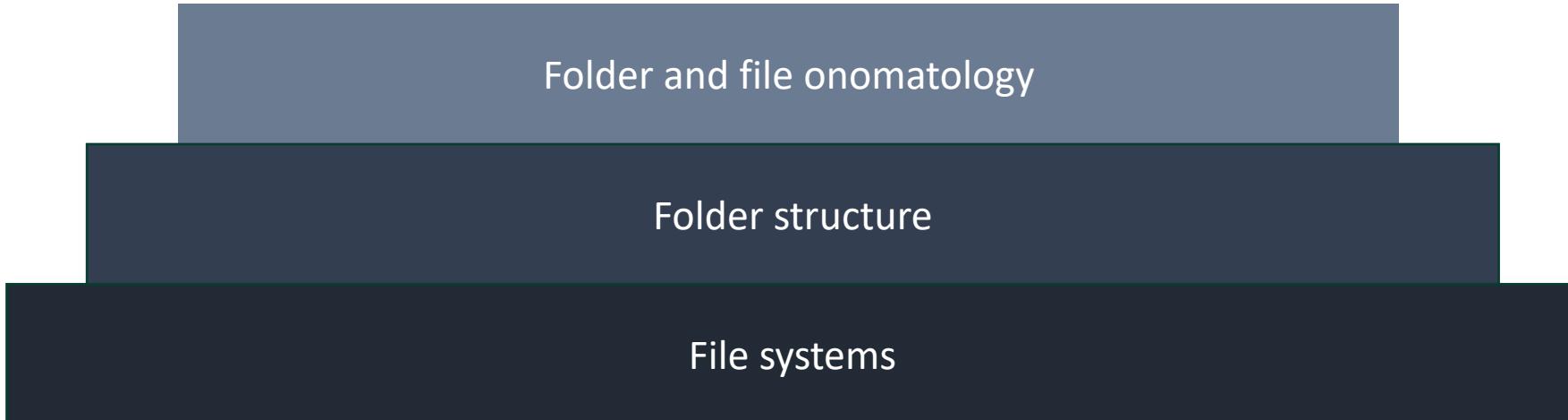
Data -> Code -> Figures and tables



... καὶ ἐγένετο ἡ τάξις (which gave birth to order)

- Clear title
- Reading a folder, we know what it contains
- .Rproj in the higher working directory
- Separate folders depending on the stored information
- */readme.md* file.





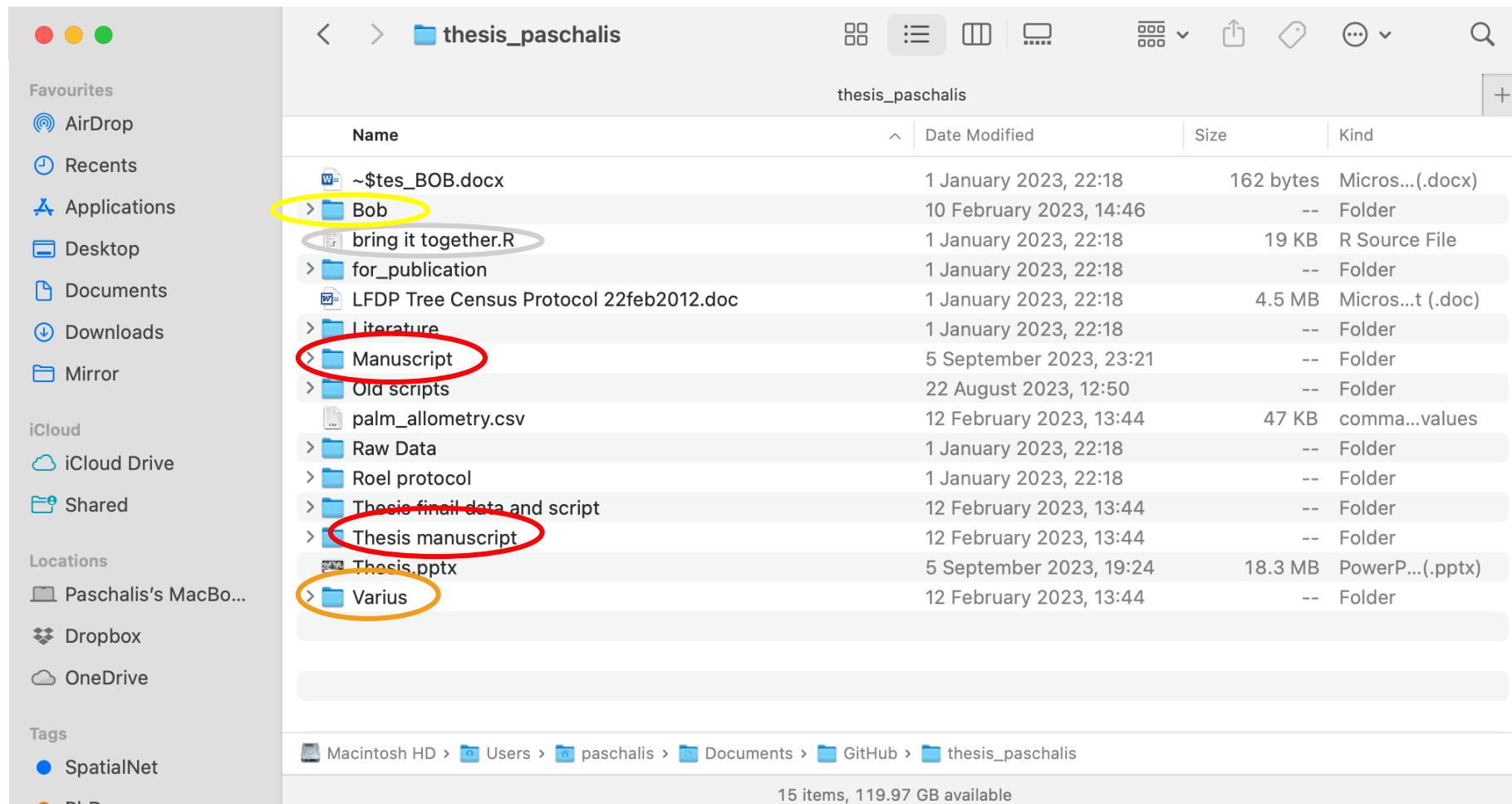
Folder and file onomatology

Folder structure

File systems

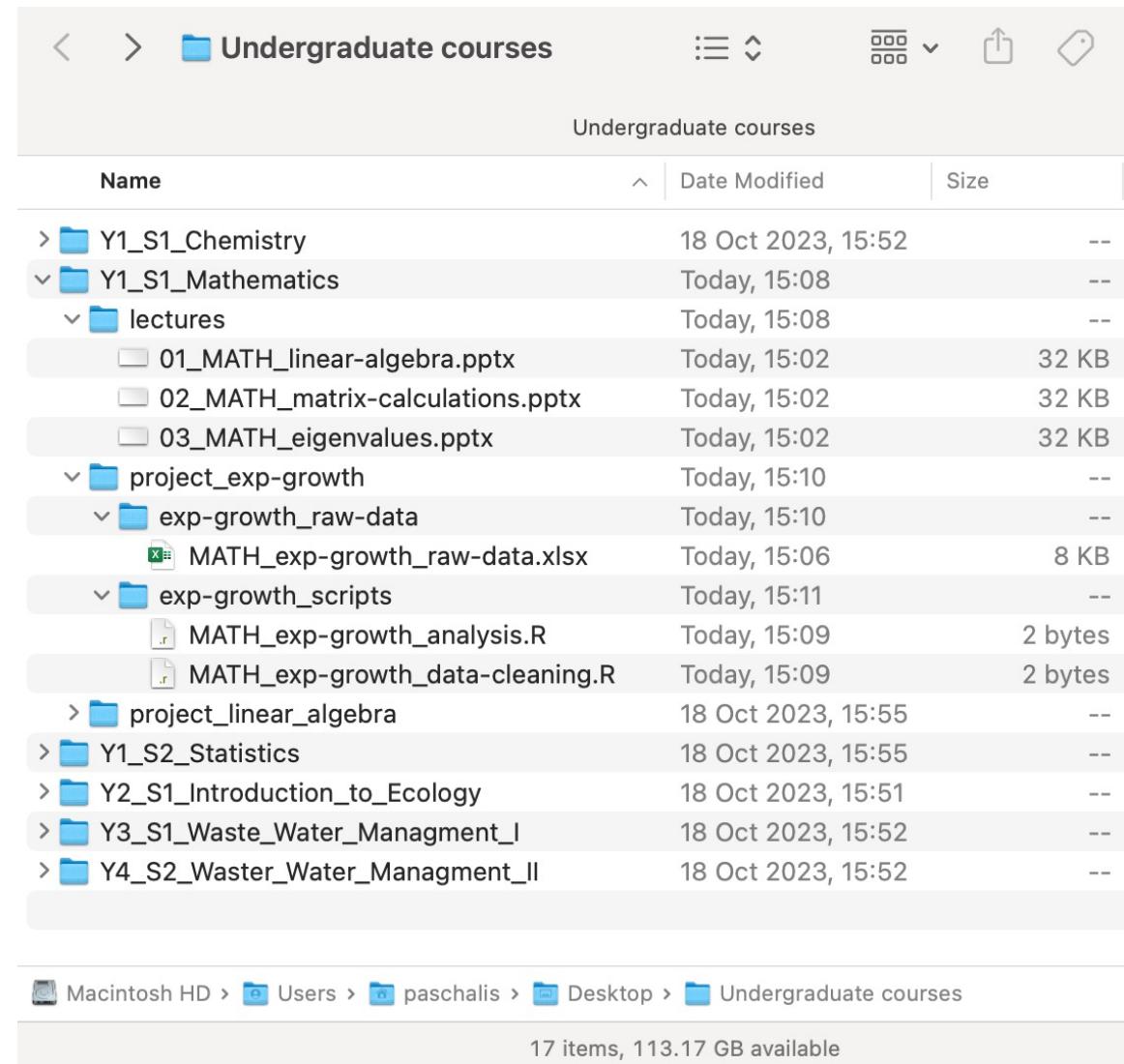
Something went wrong

- What's the difference between `/manuscript` and `/Thesis_manuscript`;
- What is stored in `/varius` ;
- Bring together what?
- Nice to meet you `/Bob...`



General folder and file naming rules

- Separate folder depending on the stored information.
 - */lectures*
 - */past_exams*
 - */data*
- Find a coding system:
 - Y1 – Year 1
 - S1 – Semester 1
- Make separate folder for script outputs
 - */output/figures*
- Use Latin characters and **NOT** special characters
- Prefer underscore to separate words (_)
- **Include a *readme.md* file!!!**

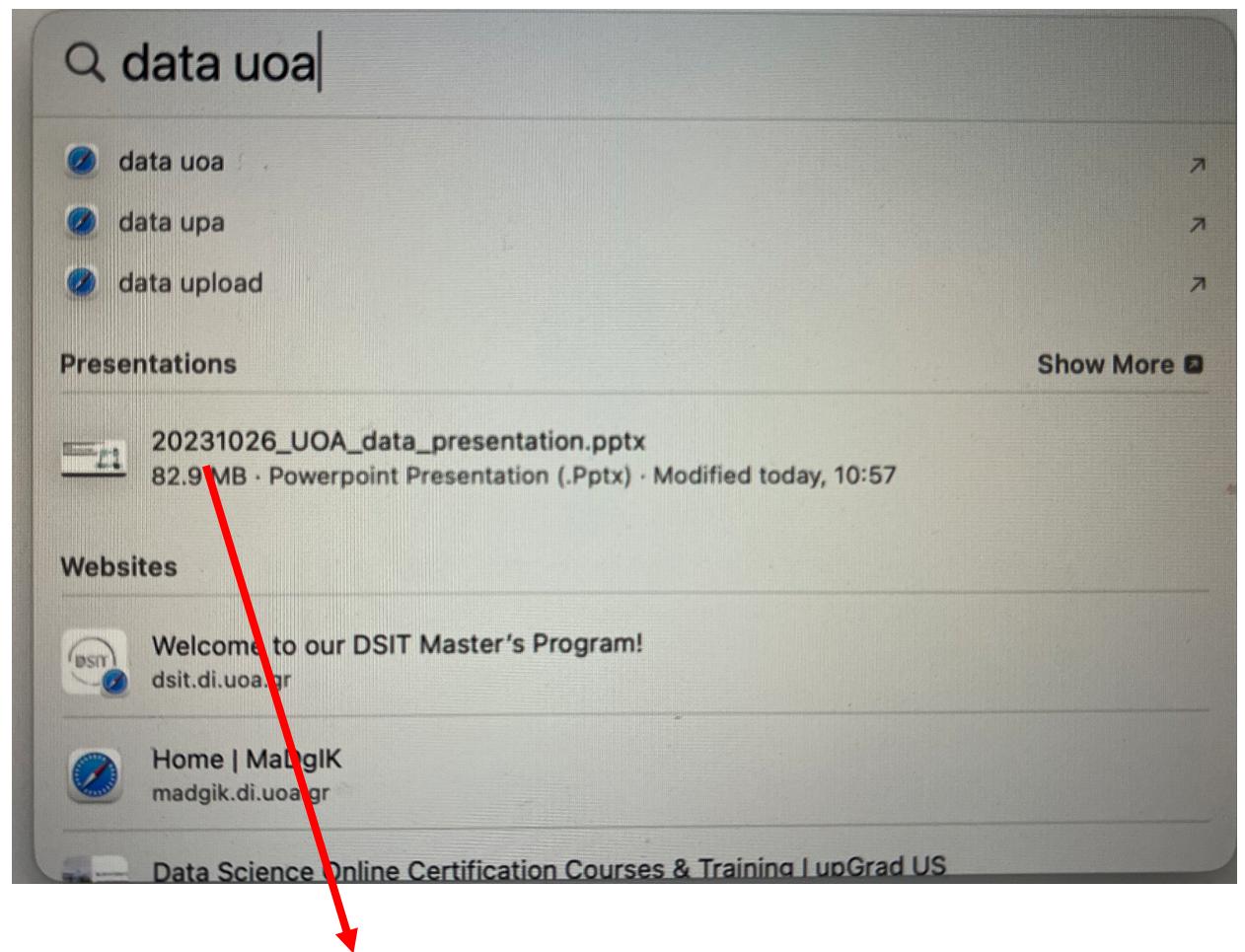
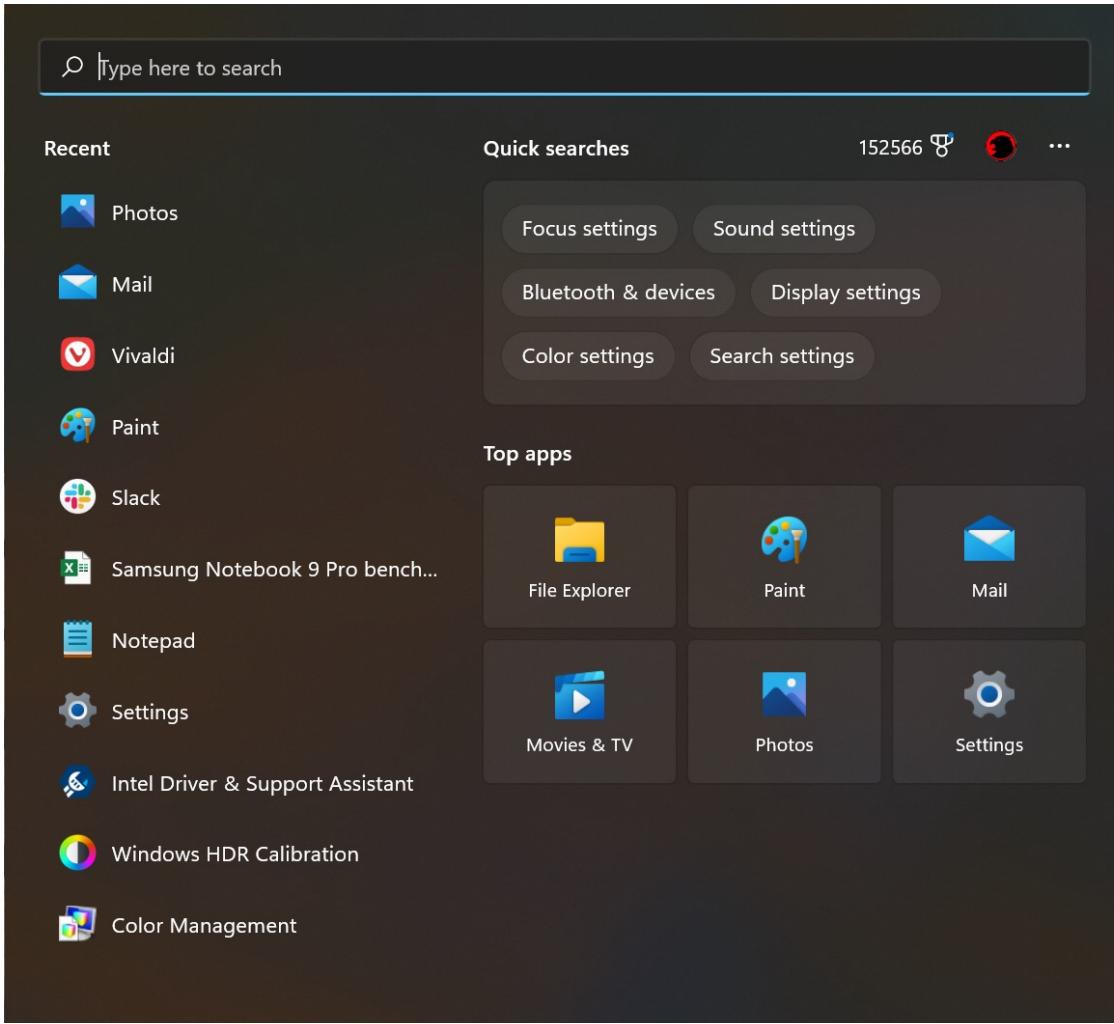


Why to prefer underscores:

	N	O	P
	TOT C	TOT N	TOT P
7	801.96	1446.74	427.01
2	899.51	1527.7	364.69
4	2154.82	2179.66	608.62
6	1024.78	1867.78	391.35
2	942.7	1874.02	422.47
4	1032.28	1808.49	537.29
3	1245.08	2388.49	493.63
9	1901.68	2101.78	503.29
3	1070.03	1417.4	518.02
4	761.91	1388.59	463.6
5	1242.04	1608.18	584.86

```
> datos$CN <- log(datos$TOT.C, 10) / log(datos$TOT.N, 10)
Error in log(datos$TOT.C, 10) :
  non-numeric argument to mathematical function
In addition: Warning message:
Unknown or uninitialized column: 'TOT.C'.
> datos$CN <- log(datos$TOT.C, 10) / log(datos$TOT.N, 10)
Error in log(datos$TOT.C, 10) :
  non-numeric argument to mathematical function
In addition: Warning message:
Unknown or uninitialized column: 'TOT.C'.
> |
```

- If file names are encoded and you have the key, any moment you can search what you want!



Date_University_name

The image shows a PDF viewer interface with two documents open side-by-side.

Top Document (Left):

- Title: **nature COMMUNICATIONS**
- Type: ARTICLE
- DOI: <https://doi.org/10.1038/s41467-020-16881-7>
- Abstract: Meta-analysis of the factors on soil micro...
- Authors: Zhenghu Zhou^{1,2}, Chuankuan Wang^{1,2}

Bottom Document (Right):

- Title: **frontiers | Frontiers in Forests and Global Change**
- Type: ORIGINAL RESEARCH
- Date: published: 12 July 2022
- DOI: doi: 10.3389/ffgc.2022.867912
- Abstract: Allometric Models to Estimate Carbon Content in Arecaceae Based on Seven Species of Neotropical Palms
- Authors: Gerardo Avalos^{1,2*}, Milena Cambronero² and Carolina Alvarez-Vergnani¹

**100
YEARS****Journal of Ecology***Journal of Ecology* 2012, **100**, 987–996

doi: 10.1111/j.1365-2745.2012.01973.x

Chance in plant communities: a new approach to its measurement using the nugget from spatial autocorrelation

Gretchen Brownstein¹, John B. Steel¹, Stefan Porter¹, Abraham Gray¹, Craig Wilson¹,
Padarn G. Wilson² and J. Bastow Wilson^{1*}

¹Botany Department, University of Otago, PO Box 56, Dunedin, New Zealand; and ²CSIRO Mathematical and Information Sciences, Clayton South, Vic. 3169, Australia

Summary

1 The unpredictable component of community structure – chance/randomness/irregularity – is



Spiros Makris

Created on June 15, 2023

Question Info

my path files are in greek characters and some programs cant find the file path way

C:\Users\spiro\OneDrive\Υπέρδιάδη\ΗΕΠΙ

my path files are in greek characters and some programs cant find the file path way

[Reply](#)[I have the same question \(0\)](#)[Subscribe](#) | [Report abuse](#)

Last updated June 18, 2023

Views 60

Applies to:

Windows / Windows 10 /



Replies (3) ✓



Sumit D - IA

Independent Advisor

Replied on June 15, 2023

[Report abuse](#)[Unlock now](#)

Hi Spiros,

I'm Sumit, here to answer your query at the Microsoft Community.

That is a folder inside Onedrive folder. If you go to <https://onedrive.com> in your browser, do you see the folder there?

Please keep us posted.

Sumit

Available 6 PM - 8 AM PST

47

You're invited to
Microsoft 365 for

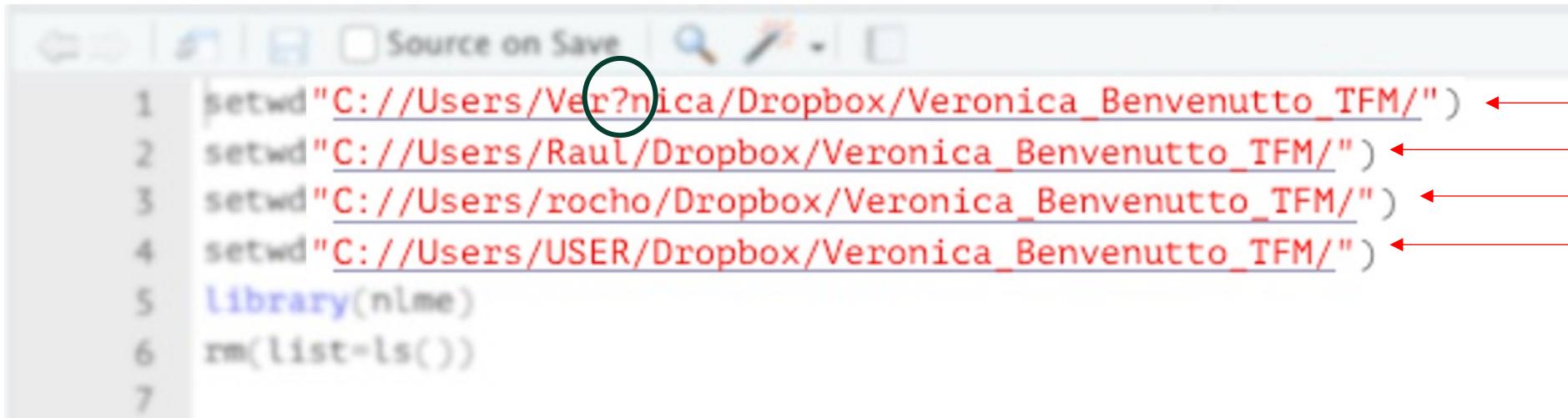
Same applies with spanish and special characters:

Ñ, é, ñ, è, ö, ò, ó, í, ï, á, à, ç, ì

The only difference is in Greek there are ONLY ‘special’ characters:

Αα, Ββ, Γγ, Δδ, Εε, Ζζ, Ηη, Θθ, Ιι, Κκ, Λλ, Μμ, Νν, Ξξ, Οο, Ππ, Ρρ, Σσ
Ττ, Υυ, Φφ, Χχ, Ψψ, Ωω

Instead of having one working directory for each computer...



```
1 setwd("C://Users/Ver?nica/Dropbox/Veronica_Benvenutto_TFM/")
2 setwd("C://Users/Raul/Dropbox/Veronica_Benvenutto_TFM/")
3 setwd("C://Users/rocho/Dropbox/Veronica_Benvenutto_TFM/")
4 setwd("C://Users/USER/Dropbox/Veronica_Benvenutto_TFM/")
5 library(nlme)
6 zm(list=ls())
7
```

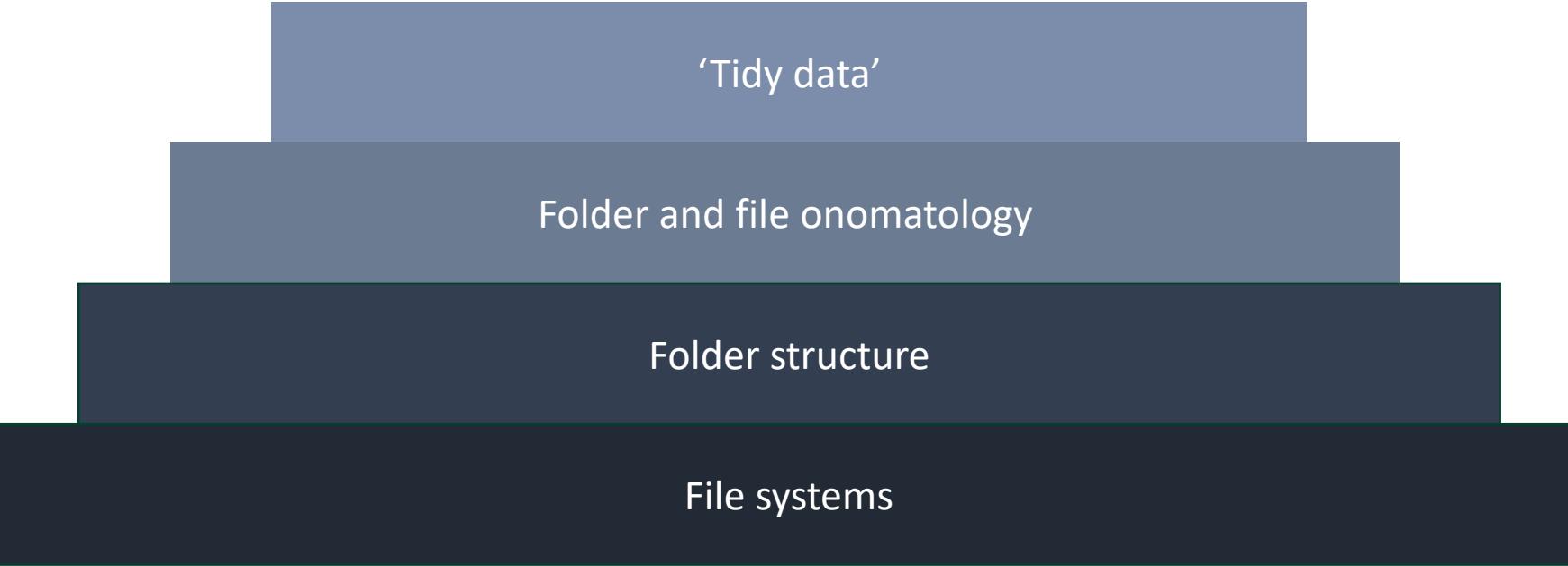
Veronicas computer
Raul's computer at home
Raul's computer at IVAGRO
Cloud directory

... we can use .Rproj and work with relative paths



```
1 # Load packages
2 library(nlme)
3 library(tidyverse)
4
5 # Load data
6 datos <- read_csv("veronica/Activ_Ens_Regajal.csv")
7 datos <- as.data.frame(datos)
8
```

Any user on the planet can acces it,
if has access to the Rproj



'Tidy data'

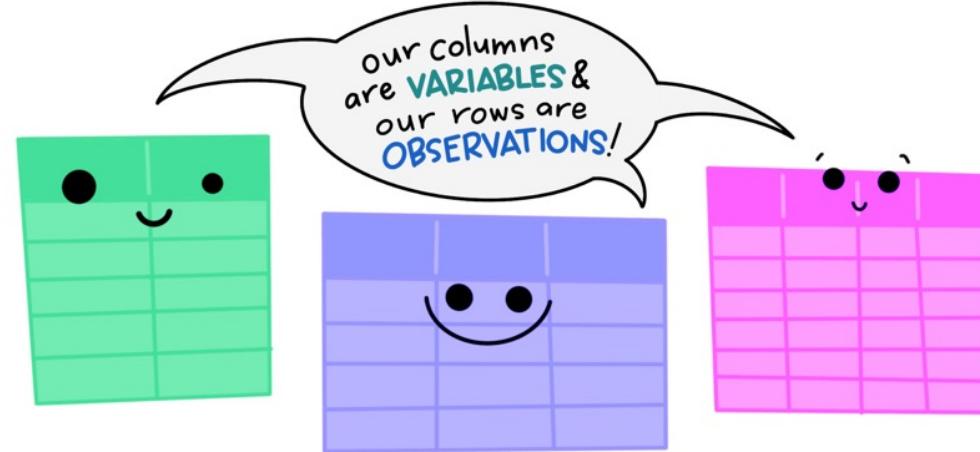
Folder and file onomatology

Folder structure

File systems

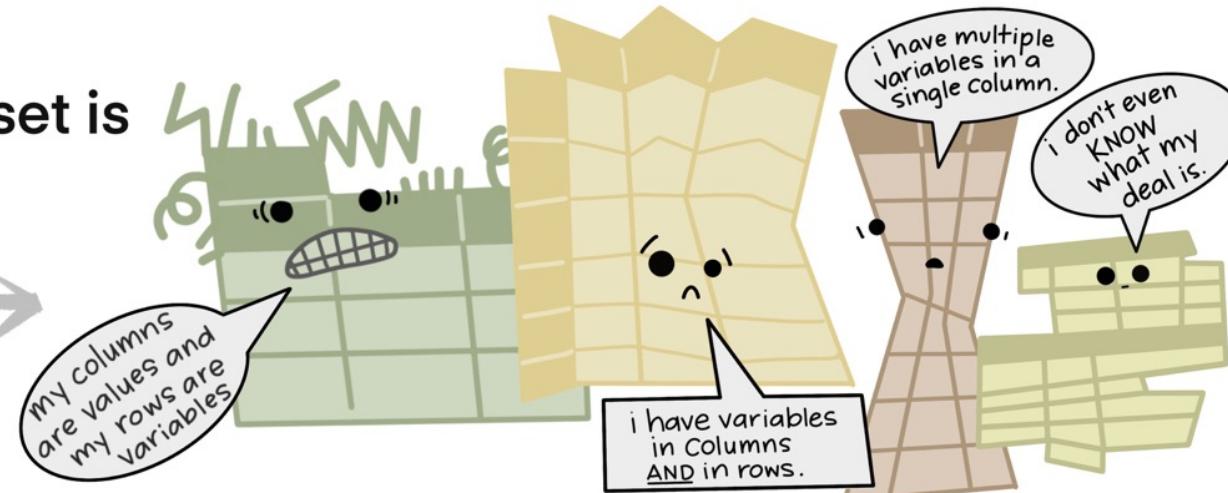
Tidy data

The standard structure of
tidy data means that
“tidy datasets are all alike...”



“...but every messy dataset is
messy in its own way.”

-HADLEY WICKHAM

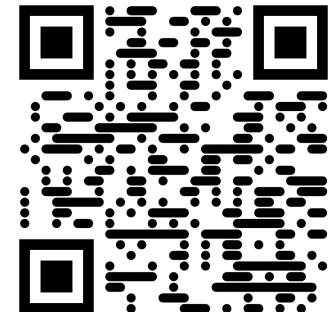




Journal of Statistical Software

MMMMMM YYYY, Volume VV, Issue II.

<http://www.jstatsoft.org/>



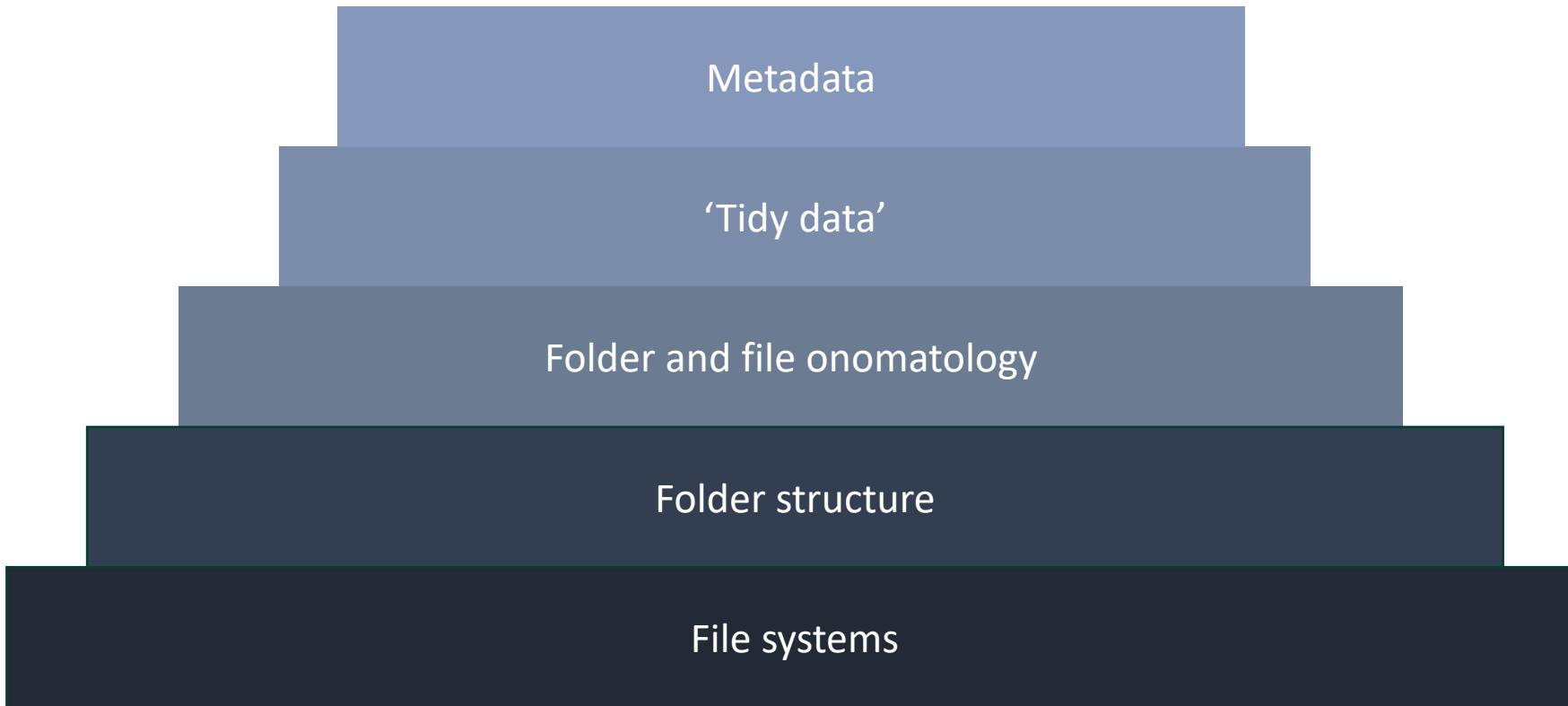
Tidy Data

Hadley Wickham
RStudio

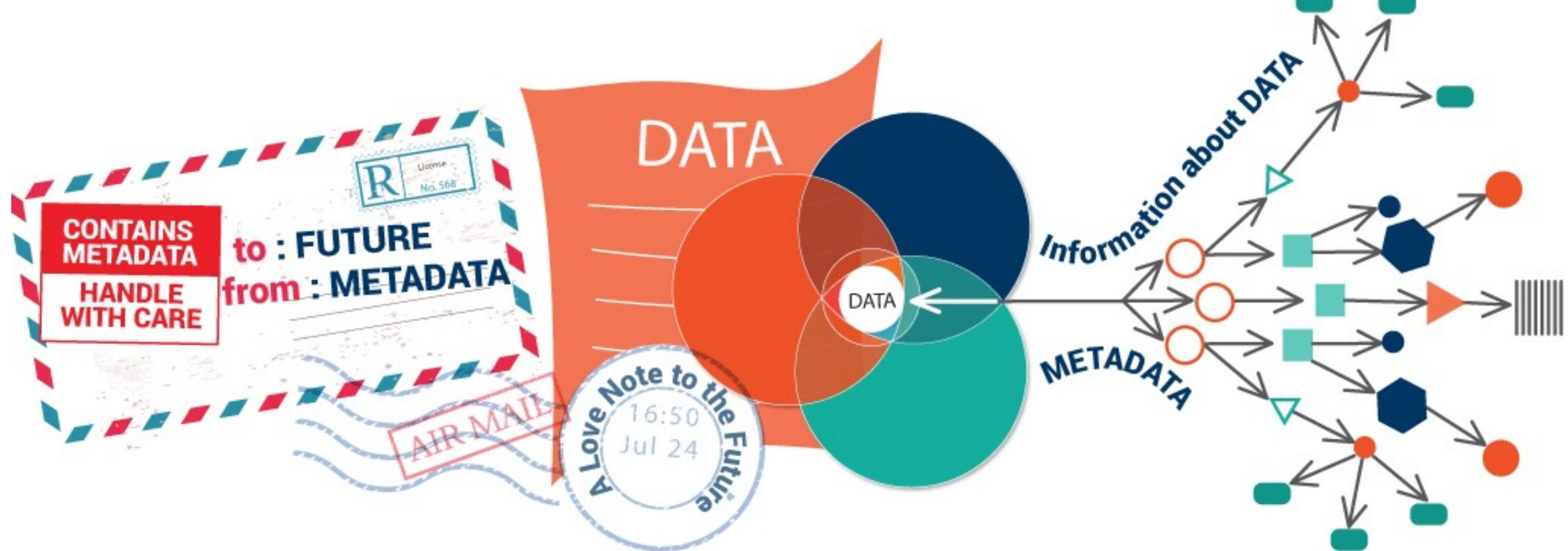
Abstract

A huge amount of effort is spent cleaning data to get it ready for analysis, but there has been little research on how to make data cleaning as easy and effective as possible. This paper tackles a small, but important, component of data cleaning: data tidying. Tidy datasets are easy to manipulate, model and visualize, and have a specific structure: each variable is a column, each observation is a row, and each type of observational unit is a table. This framework makes it easy to tidy many datasets because only a small set of tools are needed to deal with a wide range of untidy datasets. This structure also makes it easier to develop tidy tools for data analysis, tools that both input and output tidy datasets. The advantages of a consistent data structure and matching tools are demonstrated with a case study free from mundane data manipulation chores.

Keywords: data cleaning, data tidying, relational databases, R.



Metadata



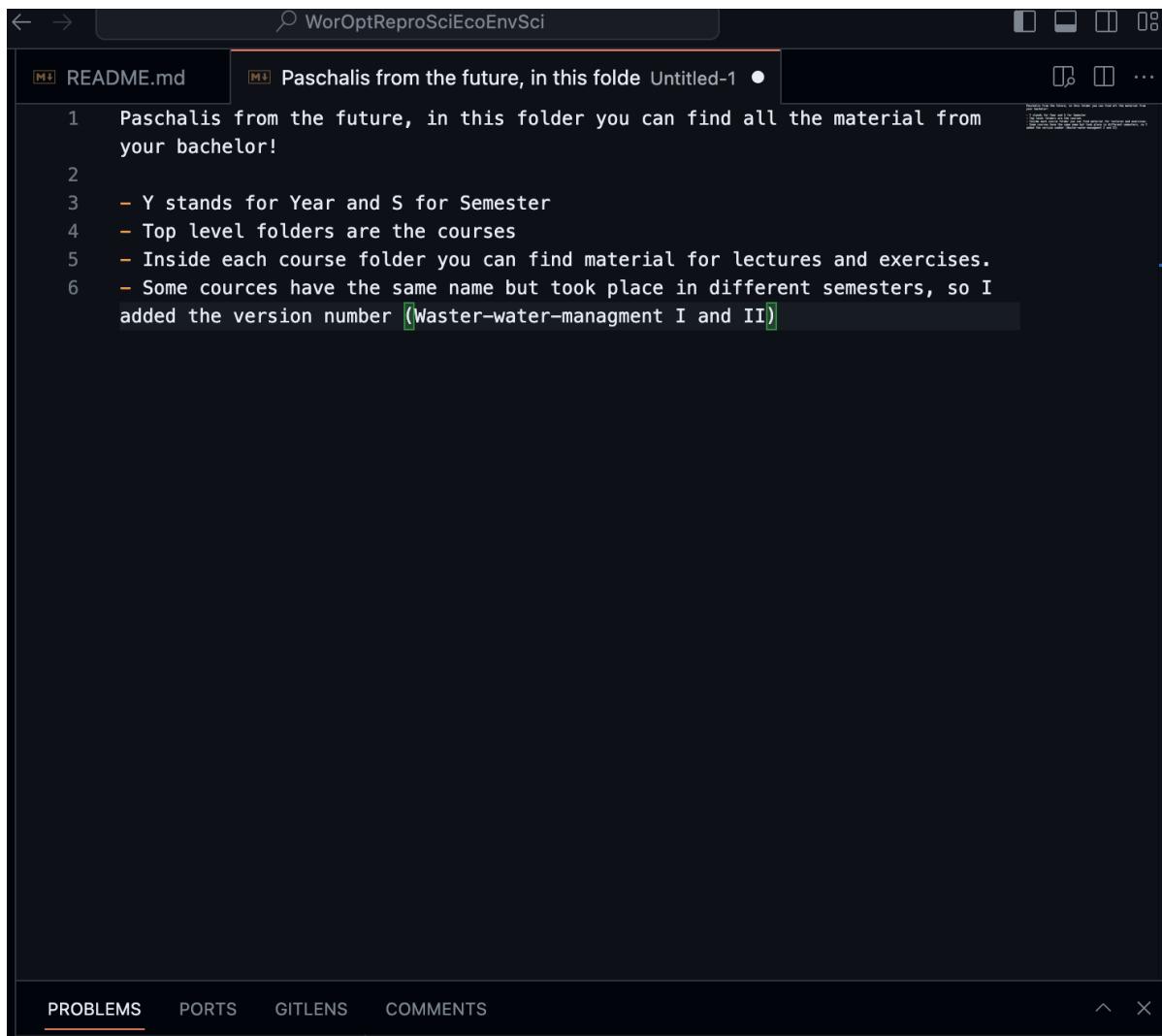
Data which describe data

- For a manuscript metadata are:
 - References
 - Table of content
- For a spreadsheet:
 - Columns and the explanation
 - Units of measure

A	B
Column	description
#	unique number of each site, different from N_folder
N_folder	I numbered each folder in order to simplify the identification
rev	Y or blanck. If revised or not
folder_old	Name of folder as in dropbox
country_code	unique code per country followin ISO
country	Country name
state	for countries like USA, Spain, Uk simplifys the identification of the sites
study_site	Study site as described in the raw data
latitude	
langitude	
biome	
PI_name	PI of the site. Sometimes were 2, I used only the name of the one who provided email
PI_last_name	PI last name
PI_contact	email adress
year_biochemistry	Year of data collection of biochemistry data. If NA, then biochemistry data are missing
year_community	Year of data collection of community data. If NA, then community data are missing
year_plant_elm_cont	Year of data collection of plant_elements data. If NA, then plant_element data are missing
year_functioning_metrics	Year of data collection of functiona_metris data. If NA, then functional_metric are missing
comment	personall comment on folder and data strstructure

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
1	#	N_folder	rev	rev_by	assistance	decision	comment	fodler_namer	country_code	country	state	study_site	latitude	longitude	biome
2	1	1	Y					01_Alberto_Canarini	AUT	Austria		ClimGrass	47.129167	11.305556	grassland
3	2	2	Y					02_Alexander_Tischer	ECU	Ecuador		Ferpast, Ecuador Loja	-3.95734	-79.038532	grassland
4	3	3	Y				Spanish coordinates	03_Alicia_Montesinos	ESP	Spain		Murcia	37.724125	1.70496111	shrubland
5	4	4	Y					04_Aline_Frossard	NAM	Namibia		Namibia Desert Fog zone	-23.0152	14.6716	Hyperarid Desert
6	5	5	Y					05_Amanda_Koltz	USA	USA	Alaska	Toolik Lake	68.624021	-149.621698	tundra (open grassland)
7	6	6	Y					06_Amit_Kumar	DEU	Germany		Friedland	51.493667	9.926917	Agricultural
8	7	7	Y					07_Andrea_Britton	GBR	UK	Scotland	Culardoch	57.07115694	-3.347469731	alpine
9	8	8	Y					08_Anna_Heintz-Buschart	ZAF	South Africa		Ukulinga	-29.67	30.4	grassland
10	9	9	Y					09_Anne_Ebeling	DEU	Germany		Jena Experiment	50.95130066	11.62131181	grassland

Metadata which describe the folder structure



WorOptReproSciEcoEnvSci

README.md

Paschalis from the future, in this folder you can find all the material from your bachelor!

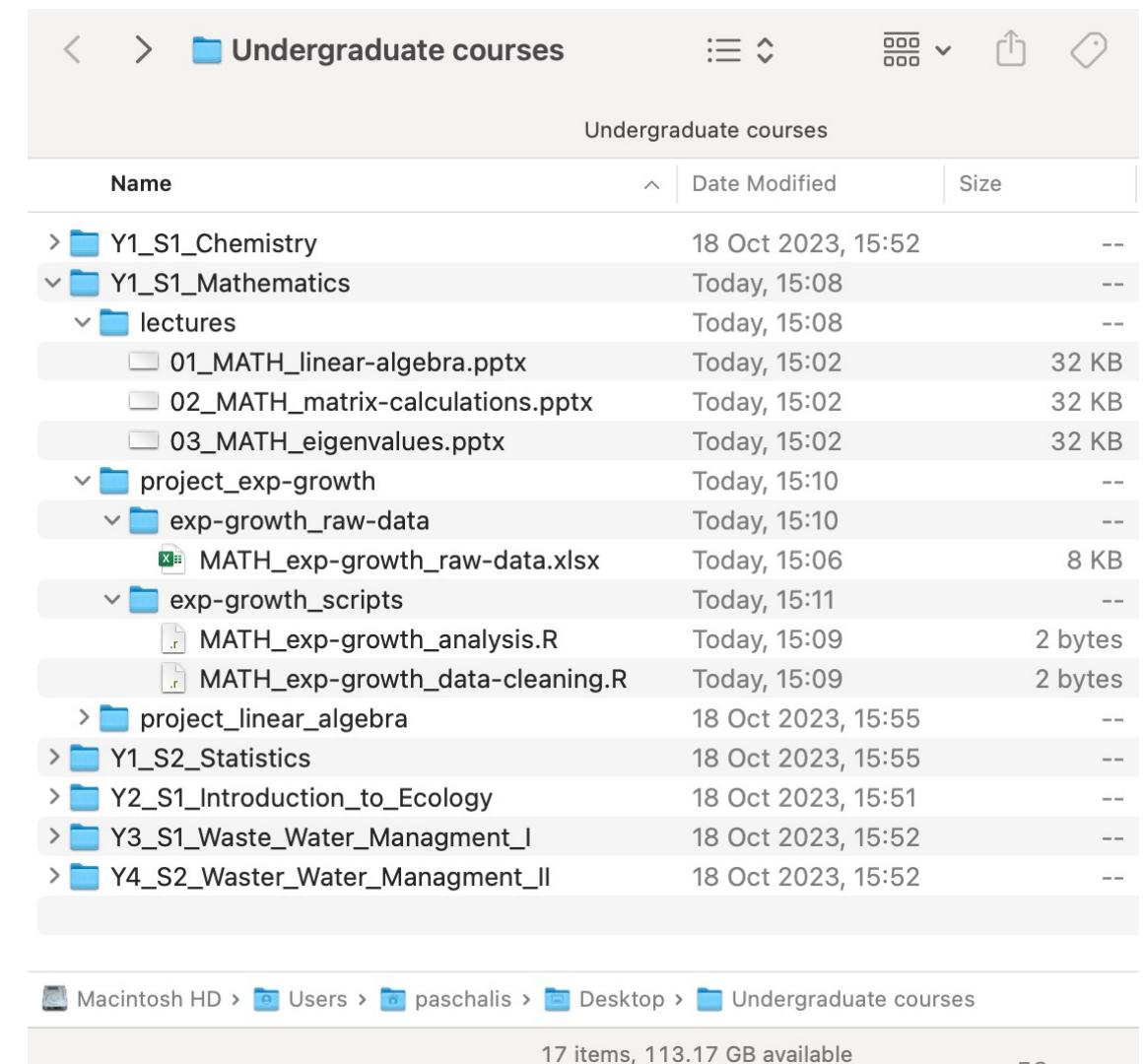
Y stands for Year and S for Semester

Top level folders are the courses

Inside each course folder you can find material for lectures and exercises.

Some courses have the same name but took place in different semesters, so I added the version number (Waster-water-managment I and II)

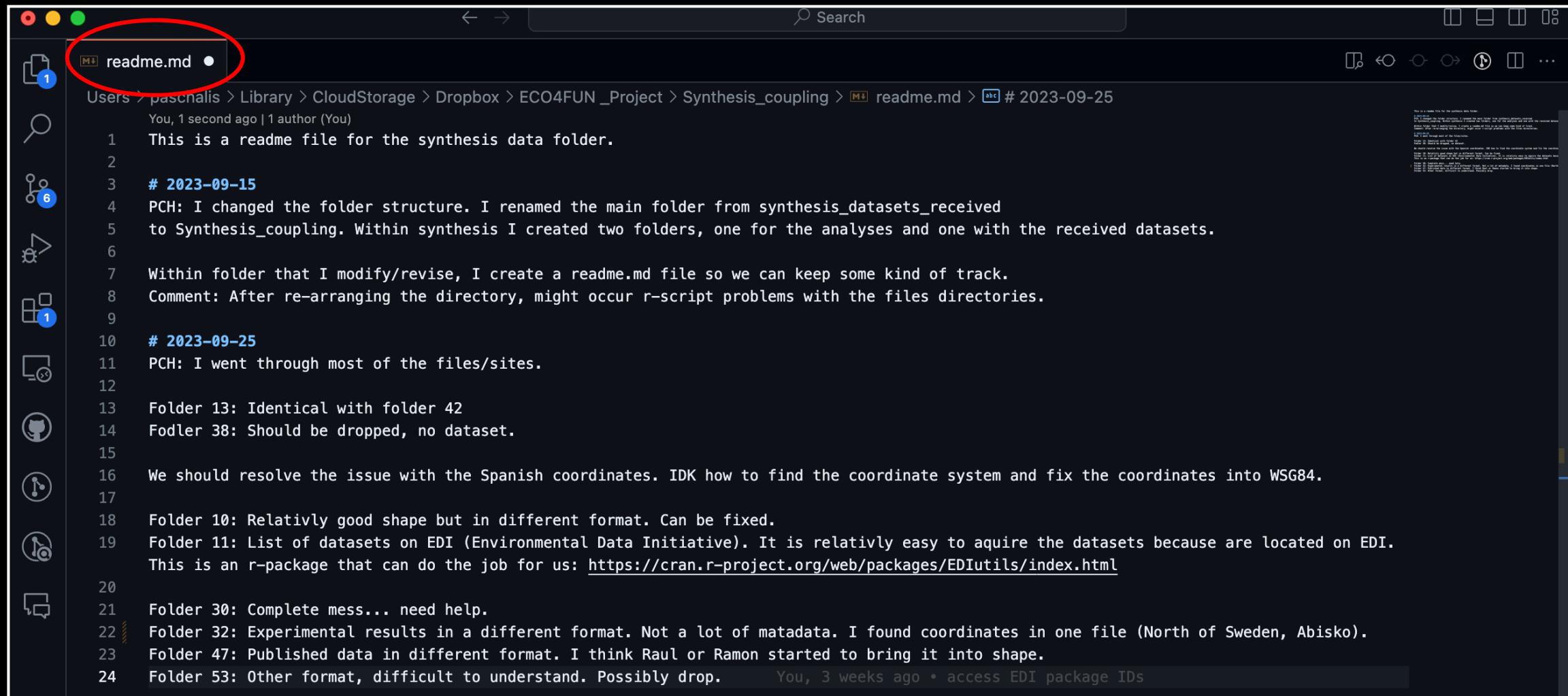
PROBLEMS PORTS GITLENS COMMENTS



Name	Date Modified	Size
> Y1_S1_Chemistry	18 Oct 2023, 15:52	--
Y1_S1_Mathematics	Today, 15:08	--
lectures	Today, 15:08	--
01_MATH_linear-algebra.pptx	Today, 15:02	32 KB
02_MATH_matrix-calculations.pptx	Today, 15:02	32 KB
03_MATH_eigenvalues.pptx	Today, 15:02	32 KB
project_exp-growth	Today, 15:10	--
exp-growth_raw-data	Today, 15:10	--
MATH_exp-growth_raw-data.xlsx	Today, 15:06	8 KB
exp-growth_scripts	Today, 15:11	--
MATH_exp-growth_analysis.R	Today, 15:09	2 bytes
MATH_exp-growth_data-cleaning.R	Today, 15:09	2 bytes
> project_linear_algebra	18 Oct 2023, 15:55	--
Y1_S2_Statistics	18 Oct 2023, 15:55	--
Y2_S1_Introduction_to_Ecology	18 Oct 2023, 15:51	--
Y3_S1_Waste_Water_Managment_I	18 Oct 2023, 15:52	--
Y4_S2_Waster_Water_Managment_II	18 Oct 2023, 15:52	--

Macintosh HD > Users > paschalis > Desktop > Undergraduate courses

17 items, 113.17 GB available



readme.md

Users > pascnalis > Library > CloudStorage > Dropbox > ECO4FUN _Project > Synthesis_coupling > readme.md > # 2023-09-25

You, 1 second ago | 1 author (You)

```
1 This is a readme file for the synthesis data folder.
2
3 # 2023-09-15
4 PCH: I changed the folder structure. I renamed the main folder from synthesis_datasets_received
5 to Synthesis_coupling. Within synthesis I created two folders, one for the analyses and one with the received datasets.
6
7 Within folder that I modify/revise, I create a readme.md file so we can keep some kind of track.
8 Comment: After re-arranging the directory, might occur r-script problems with the files directories.
9
10 # 2023-09-25
11 PCH: I went through most of the files/sites.
12
13 Folder 13: Identical with folder 42
14 Fodler 38: Should be dropped, no dataset.
15
16 We should resolve the issue with the Spanish coordinates. IDK how to find the coordinate system and fix the coordinates into WSG84.
17
18 Folder 10: Relativly good shape but in different format. Can be fixed.
19 Folder 11: List of datasets on EDI (Environmental Data Initiative). It is relativly easy to aquire the datasets because are located on EDI.
This is an r-package that can do the job for us: https://cran.r-project.org/web/packages/EDIutils/index.html
20
21 Folder 30: Complete mess... need help.
22 Folder 32: Experimental results in a different format. Not a lot of metadata. I found coordinates in one file (North of Sweden, Abisko).
23 Folder 47: Published data in different format. I think Raul or Ramon started to bring it into shape.
24 Folder 53: Other format, difficult to understand. Possibly drop.
```

Metadata in the form of notebook

Our worst enemy is an older version of ourself

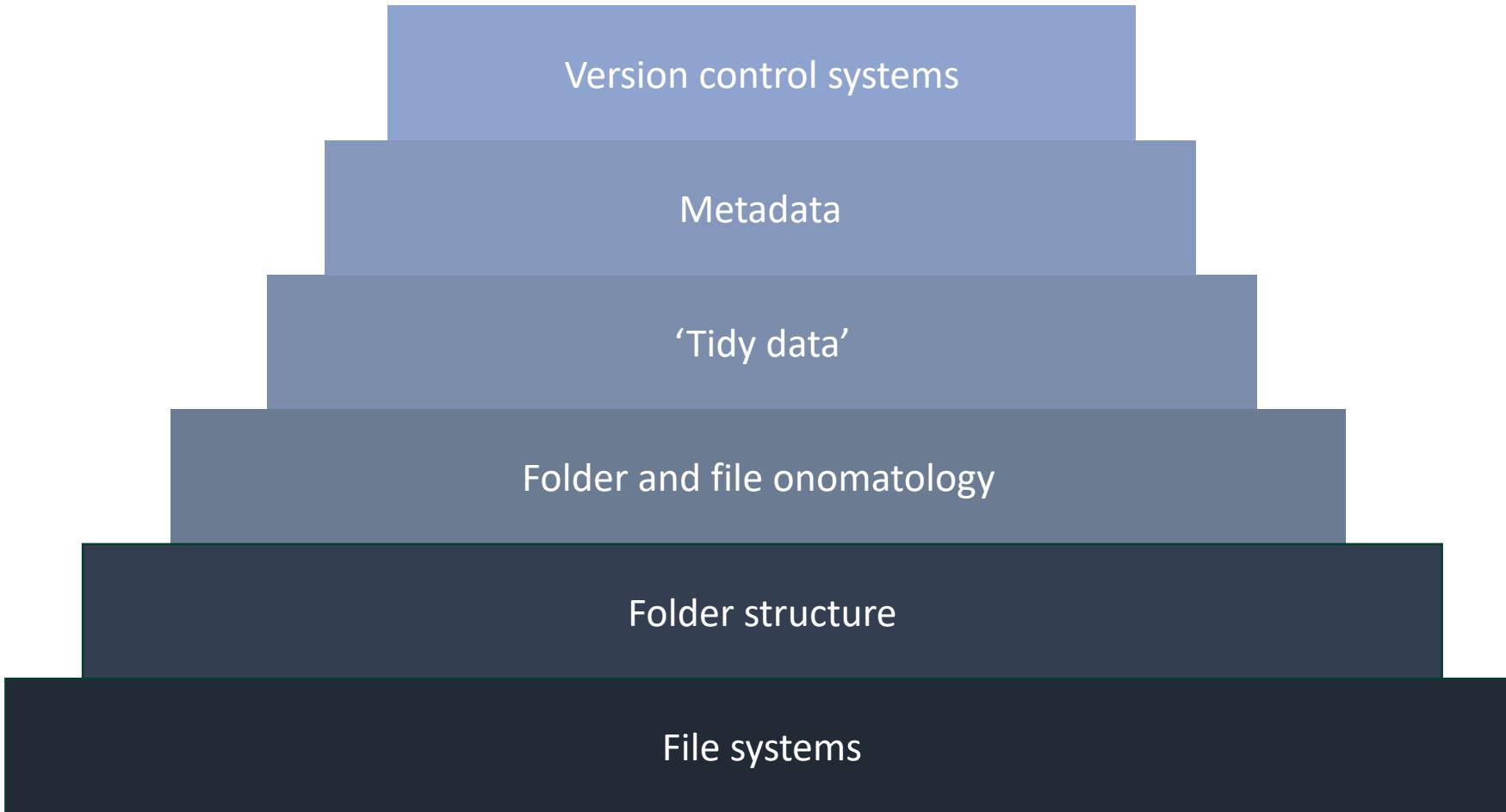
Think about it a bit more...

Our best friend is an older version of ourself
because taught us how to do or **not** do things

TRACK PROJECT HISTORY



Scriberia



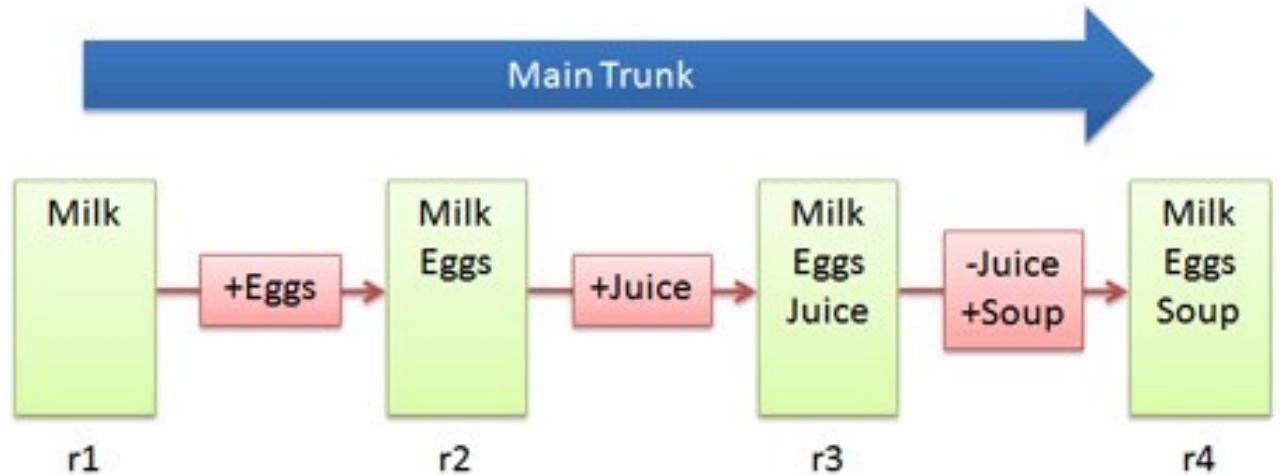
Final_final_final

 draft	>	 manuscript_v1.docx
 manuscript	>	 manuscript_v2.docx
 submitted_manuscript	>	 manuscript_v3.docx
		 manuscript_v5.docx
		 manuscript_v6.docx
		 manuscript_v7_rev_1.docx
		 manuscript_v8_final.docx
		 manuscript_v8_final2.docx
		 manuscript_v9_final_final.docx
		 manuscript_v10_final_final_reviewer_com.docx

- Confusing
- Resource hungry
- Make sense to you... and only you
- For manuscripts with many authors, you might have 20-30 versions

Version control systems

- Within a version control system, you essentially set the working directory to a **remote server** that maintains a record for every change that occurs for each file
- Every change in a file is stored as **a new version**
- Any moment you can **revert changes** and ‘go back’
- Handy for remote **collaboration**
- Must for those working with **code**
- Support many (all?) programming languages
(R, Pascal, HTML, Julia, C, C++, python, Fortran)



The image consists of three side-by-side screenshots demonstrating how to access version history in different document editing platforms.

Top Left: A screenshot of the Google Docs interface. A red circle highlights the "Version History" icon (a circular arrow) located in the top right corner of the toolbar.

Middle Left: A screenshot of Microsoft Word. A red circle highlights the "Browse Version History" button, which is part of the ribbon's "Review" tab.

Right: A screenshot of the Dropbox interface showing the "Version history" for a file named "participating_experiments.xlsx". The sidebar shows a list of versions, and the main area displays the details of a specific version from September 27, 2023, edited by Paschalis Chatzopoulos at 3:03 PM.



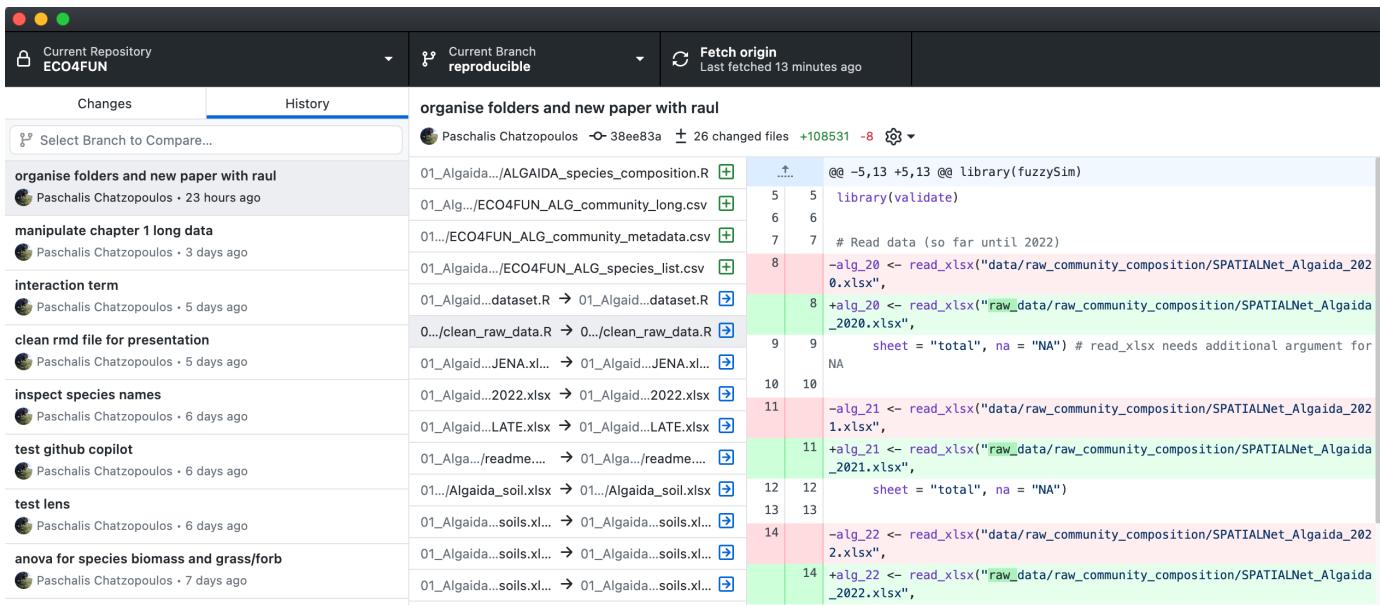
What is GitHub

GitHub is for git what Rstudio is for R

- Is a version of ‘version control systems’ (yes there are many!)
- User friendly graphic interface
- Its free
- Many additional services for academic stuff

```
# On branch master
# Changes to be committed:
#   (use "git reset HEAD <file>..." to unstage)
#
#       modified:   public/index.html
#       deleted:   public/sitemap.xml
#       new file:  public/stylesheets/mobile.css
#
# Changes not staged for commit:
#   (use "git add/rm <file>..." to update what will be committed)
#   (use "git checkout -- <file>..." to discard changes in working directory)
#
#       deleted:   app.rb
#       deleted:   test/add_test_crash_report.sh
#
# Untracked files:
#   (use "git add <file>..." to include in what will be committed)
#
#       public/javascripts/
```

git



GitHub desktop



github alternatives



Images

Free

Videos

Desktop

Open source

For android

Reddit

News

Codespaces

All filters ▾ | Tools

About 151,000,000 results (0.34 seconds)

From sources across the web



GitLab



Bitbucket



SourceForge



Gitea



Launchpad



GitBucket



Apache Allura



Phabricator



AWS Elastic Beanstalk



Amazon Web Services



GitKraken



Assembla



Trac



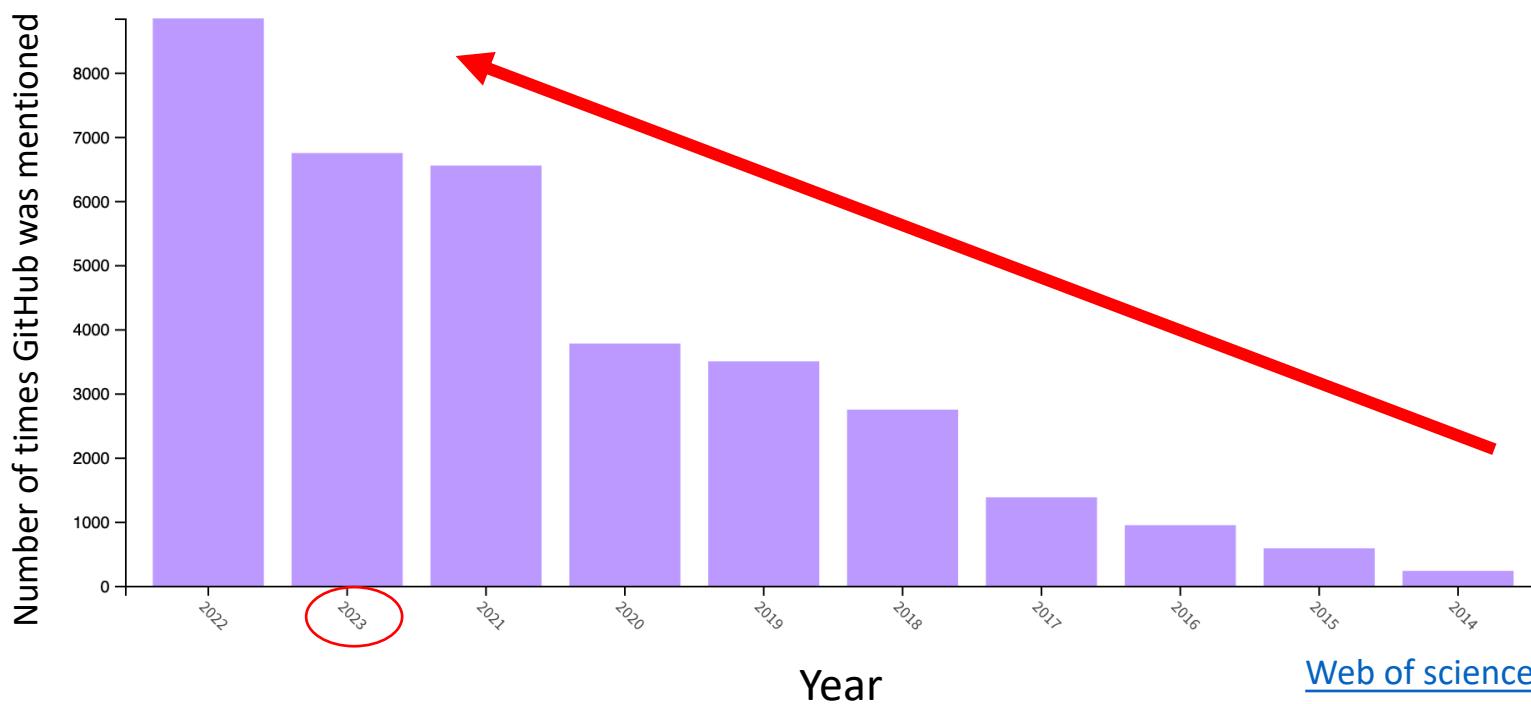
Gogs



Show less ^

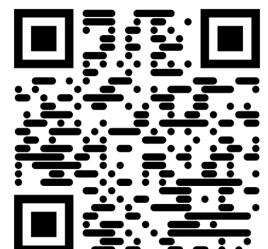
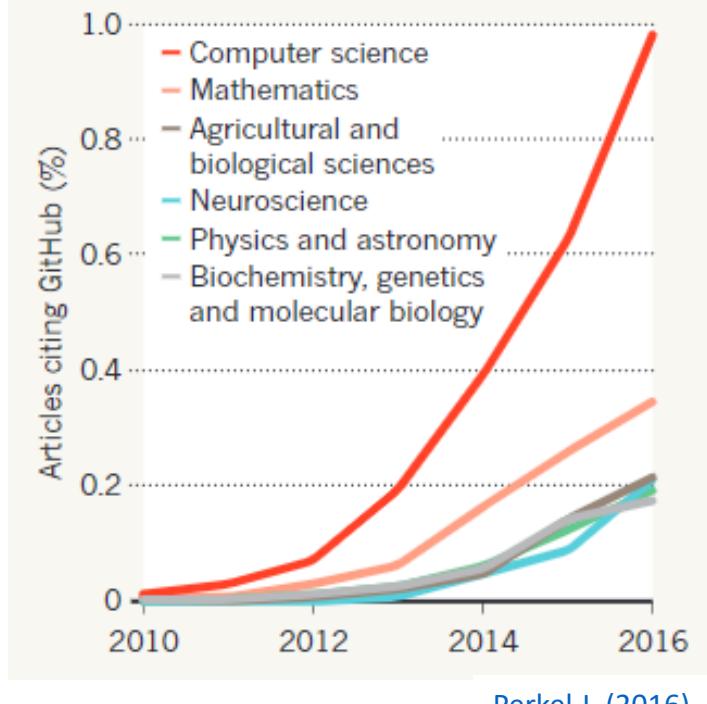
Feedback

- GitHub is around since 2013 and is an everyday tool for programmers
- Growing influence within many scientific fields



GROWING INFLUENCE OF GITHUB

An increasing proportion of research articles cite GitHub in their references.



REVIEW |  Open Access | 

Not just for programmers: How GitHub can accelerate collaborative and reproducible research in ecology and evolution

Pedro Henrique Pereira Braga , Katherine Hébert, Emma J. Hudgins, Eric R. Scott, Brandon P. M. Edwards, Luna L. Sánchez Reyes, Matthew J. Grainger ... [See all authors](#) 

First published: 21 April 2023 | <https://doi.org/10.1111/2041-210X.14108> | Citations: 2

Pedro Henrique Pereira Braga and Robert Crystal-Ornelas contributed equally to this work. Katherine Hébert, Emma J. Hudgins, Eric R. Scott, Brandon P. M. Edwards and Luna L. Sánchez Reyes contributed equally to this work.

Handling Editor: Aaron M. Ellison

Abstract

EN FR PT

1. Researchers in ecology and evolutionary biology are increasingly dependent on computational code to conduct research. Hence, the use of efficient methods to share, reproduce, and collaborate on code as well as document research is fundamental. GitHub is an online, cloud-based service that can help researchers track, organize, discuss, share, and collaborate on software and other materials related to research production, including data, code for analyses, and protocols. Despite these benefits, the use of GitHub in ecology and evolution is not widespread.
3. We outline features ranging from low to high technical difficulty, including storing code, managing projects, coding collaboratively, conducting peer review, writing a manuscript, and using automated and continuous integration to streamline analyses. Given that members of a research team may have different technical skills and responsibilities, we describe how the optimal use of GitHub features may vary among members of a research collaboration.

Create an account

Individuals

Students

Learn using real-world development tools

- ✓ **FREE** GitHub Pro while you are a student
- ✓ Valuable GitHub Student Developer Pack partner offers
- ✓ GitHub Campus Expert training for qualified applicants

[Get student benefits](#)

Teachers

Teach your students with the industry-standard tools

- ✓ **FREE** GitHub Team for courses, coding clubs, and nonprofit research
- ✓ GitHub Classroom for managing assignments

[Get teacher benefits](#)

[Join GitHub](#)

First, let's create your user account

Username *

Email address *

Password *

Make sure it's at least 15 characters OR at least 8 characters including a number and a lowercase letter. [Learn more](#).

Email preferences

Send me occasional product updates, announcements, and offers.

Verify your account



Overview Release Notes Help

GitHub Desktop

Focus on what matters instead of fighting with Git. Whether you're new to Git or a seasoned user, GitHub Desktop simplifies your development workflow.

[Download for macOS](#)



Feeling brave?

Try new features in the Beta Channel before they're released.

Apple silicon?

Download for an Apple silicon Mac.
See the [Apple docs](#) about Apple vs Intel chips.

Windows?

[Download for Windows](#)



GitHub



Paschalis Chatzopoulos

paschatz

Currently based in Cadiz, Spain. PhD candidate at
@Ecosystem-Ecology-Lab-UCA.

Edit profile

8 followers · 14 following

Cadiz university

Spain

@pascal_161

<https://orcid.org/0000-0002-3456-7203>

Highlights

PRO

Popular repositories

WorOptReproSciEcoEnvSci Public

Welcome to the "Information and data architecture in Ecology: Folder structure, metadata and project management through GitHub" repository

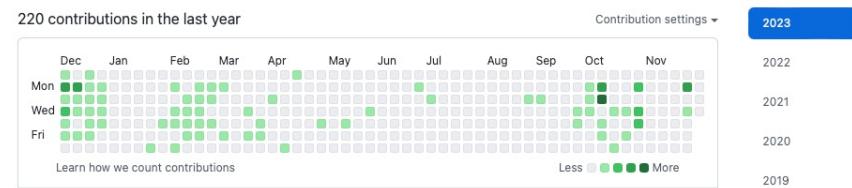
R 3

Group-Project-Acid-deposition-in-Sweden Public

This is a public version of our repository for our assignment for the course of Ecological methods.

R

220 contributions in the last year



2023

2022
2021
2020
2019

Contribution activity

November 2023

- Created 16 commits in 4 repositories
- paschatz/WorOptReproSciEcoEnvSci 12 commits
 - paschatz/Understanding-palm-forest-ca... 2 commits
 - Ecosystem-Ecology-Lab-UCA/Synthesis_c... 1 commit
 - paschatz/DRAGNET 1 commit

Show more activity

Organizations



GitHub desktop

Current Repository ECO4FUN

Current Branch reproducible

Last fetched just now

No local changes

There are no uncommitted changes in this repository. Here are some friendly suggestions for what to do next.

Preview the Pull Request from your current branch
The current branch (reproducible) is already published to GitHub. Preview the changes this pull request will have before proposing your changes.
Branch menu or ⌘ ⌘ P

Open the repository in your external editor
Select your editor in Preferences
Repository menu or ⌘ ⌘ A Open in RStudio

View the files of your repository in Finder
Repository menu or ⌘ ⌘ F Show in Finder

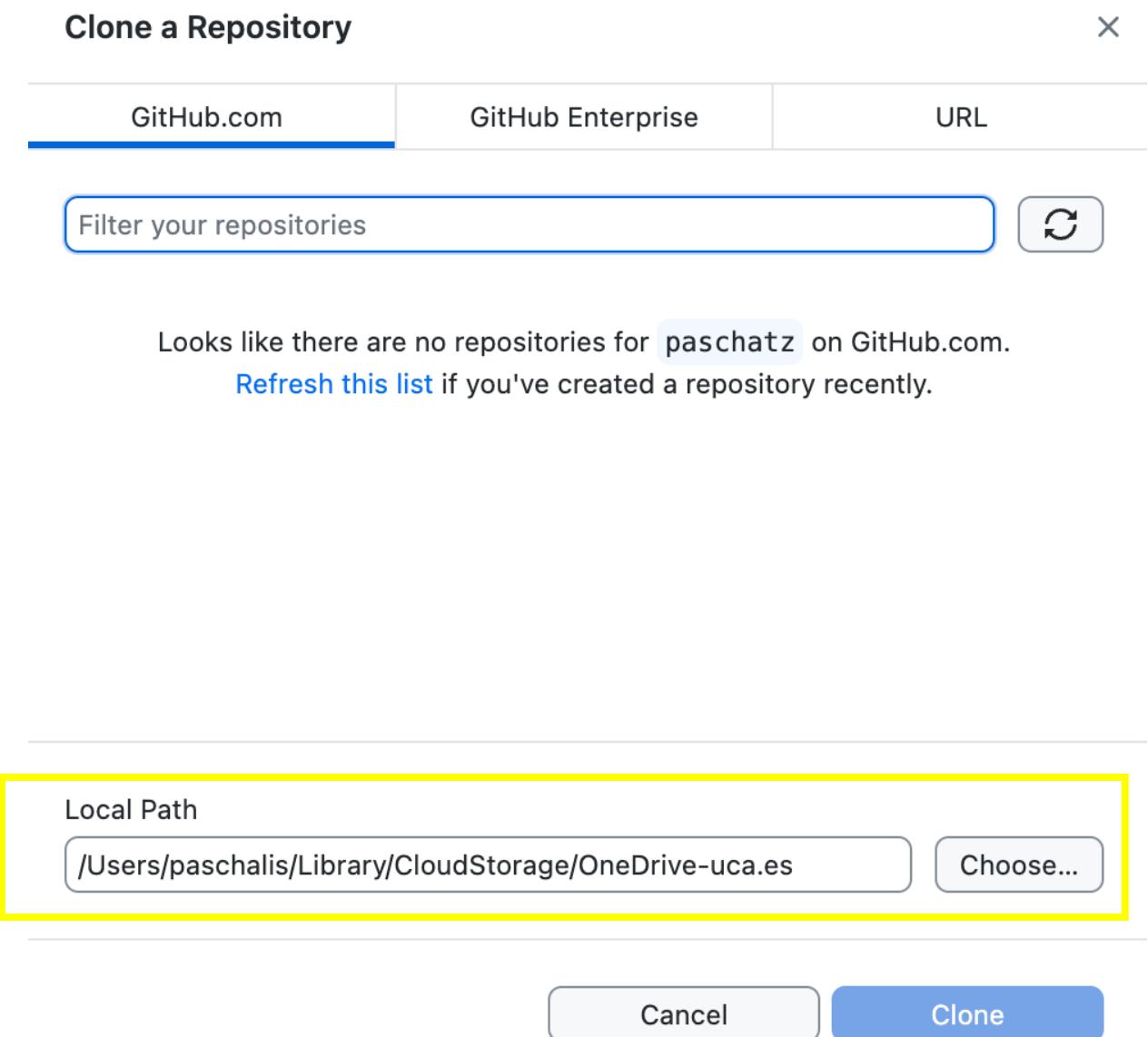
Open the repository page on GitHub in your browser
Repository menu or ⌘ ⌘ G View on GitHub

Summary (required)

Description

Commit to reproducible

- To start with GitHub, first you have to define the directory (local) in which the repository (remote) will be stored
- In git universe, a ‘repository’ is equivalent to a folder or directory
- Any working directory is locally stored in your HD
- Once synchronized with GitHub will have a remote location (on the cloud)





paschatz / WorOptReproSciEcoEnvSci



Type ⌘ to search



Code

Issues 3

Pull requests

Actions

Projects 1

Wiki

Security

Insights

Settings



WorOptReproSciEcoEnvSci

Public



Pin



Unwatch

2



0



Star

0



main ▾



3 branches



0 tags

Go to file

Add file ▾

< > Code ▾



Your main branch isn't protected

Protect this branch

About



Welcome to the "Information and data architecture in environmental sciences: Folder structure, metadata and project management through GitHub" repository, scheduled to be hosted by the University of the Aegean on October 26, 2023.

Readme

Activity

0 stars

2 watching

0 forks

Releases

No releases published

[Create a new release](#)

Packages

No packages published

[Publish your first package](#)

Languages



paschatz Merge branch 'main' of https://github.com/paschatz/WorOptReproSci...



...

0788cc7 2 hours ago



33 commits



lecture_notes

notes

2 weeks ago



poster

poster

2 days ago



raw_data

update script

2 hours ago



scripts

update script

2 hours ago



.DS_Store

update script

2 hours ago



.gitignore

ign

2 weeks ago



README.md

Update README.md

3 hours ago



WorOptReproSciEcoEnvSci.Rproj

add script and raw data

2 weeks ago

README.md



Information and data architecture in environmental sciences: Folder structure, metadata and project management through GitHub ↗



paschatz / WorOptReproSciEcoEnvSci

 Type ⌘ to search[Code](#) [Issues 3](#) [Pull requests](#) [Actions](#) [Projects 1](#) [Wiki](#) [Security](#) [Insights](#) [Settings](#)

WorOptReproSciEcoEnvSci

Public

[Pin](#) [Unwatch 2](#) [Fork 0](#) [Star 0](#)

	paschatz Merge branch 'main' of https://github.com/paschatz/WorOptReproSci...	...	0788cc7	2 hours ago	🕒 33 commits	out
	lecture_notes	notes			2 weeks ago	Welcome to the "Information and data architecture in environmental sciences: folder structure, metadata and project management through GitHub"
	poster	poster			2 days ago	repository, scheduled to be hosted by University of the Aegean on October 2023.
	raw_data	update script			2 hours ago	Readme
	scripts	update script			2 hours ago	Activity
	.DS_Store	update script			2 hours ago	0 stars
	.gitignore	ign			2 weeks ago	2 watching
	README.md	Update README.md			3 hours ago	0 forks
	WorOptReproSciEcoEnvSci.Rproj	add script and raw data			2 weeks ago	

The screenshot shows a file manager interface with the following details:

- Left Sidebar:** Favourites (AirDrop, Recents, Applications, Desktop, Documents, Downloads), README.md, and WorOptReproSciEcoEnvSci.
- Central Area:** A preview window showing the README.md file content: "Information and data architecture in environmental sciences: Folder structure, metadata and project management through GitHub".
- Right Area:** A file list titled "WorOptReproSciEcoEnvSci" with the following items:

Name	Date Modified	Size	Kind
lecture_notes	3 October 2023, 17:13	--	Folder
poster	Today, 13:33	--	Folder
raw_data	Yesterday, 14:12	--	Folder
README.md	Today, 13:33	10 KB	Markdown
scripts	18 October 2023, 14:29	--	Folder
WorOptReproSciEcoEnvSci.Rproj	Today, 14:31	205 bytes	R Project



Current Repository
WorOptReproSciEcoEnvSci

Current Branch
main

Fetch origin
Last fetched 8 minutes ago

Changes History

0 changed files

Every change appears here

For every change you have to provide a description

Commit to main

No local changes

There are no uncommitted changes in this repository. Here are some friendly suggestions for what to do next.

Open the repository in your external editor
Select your editor in [Preferences](#)
Repository menu or ⌘ ⌘ A

View in RStudio

View the files of your repository in Finder
Repository menu or ⌘ ⌘ F

Show in Finder

Open the repository page on GitHub in your browser
Repository menu or ⌘ ⌘ G

View on GitHub

Summary (required)

82



Current Repository
WorOptReproSciEcoEnvSci

We can have as many repositories as we want

main

Last fetched 17 minutes ago

Filter

Add ▾

Recent

Understanding-palm-forest-carbon-dynamics-through-...

ECO4FUN

Synthesis_coupling

Ecosystem-Ecology-Lab-UCA

Synthesis_coupling

paschatz

DRAGNET

ECO4FUN

julia_project

PhD

thesis_paschalidis

Understanding-palm-forest-carbon-dynamics-through-...

WorOptReproSciEcoEnvSci

One repository for each project

No local changes

There are no uncommitted changes in this repository. Here are some friendly suggestions for what to do next.

Open the repository in your external editor

Select your editor in [Preferences](#)

Repository menu or ⌘ ⌘ A

[Open in RStudio](#)

View the files of your repository in Finder

Repository menu or ⌘ ⌘ F

[Show in Finder](#)

Open the repository page on GitHub in your browser

Repository menu or ⌘ ⌘ G

[View on GitHub](#)



Current Repository: WorOptReproSciEcoEnvSci

Current Branch: main

Last fetched 3 minutes ago

Changes 4 History

4 changed files

.DS_Store poster/UOA_lecture_poster_to_print.png presentation/Presentation1.pptx README.md

Summary (required)

Description

Commit to main

Changes 4 History

4 changed files

.DS_Store poster/UOA_lecture_poster_to_print.png presentation/Presentation1.pptx README.md

README.md

@@ -1,11 +1,11 @@

1 1 # Information and data architecture in environmental sciences: Folder structure, metadata and project management through GitHub

2 2 - SHALL I MAKE A CHANGE IN LINE 2?

3 3 This serves as an open-access repository for the forthcoming lecture entitled 'Information and data architecture in environmental sciences: Folder structure, metadata and project management through GitHub,' scheduled to be hosted by the University of the Aegean on October 26, 2023.

4 4

5 5 Within this repository, you can find all the course materials, references, and lecture handouts. Additionally, this repository to serve as a practical demonstration of building and structuring a course from the ground up in a reproducible environment.

6 6

7 7 Please fill in the pre-lecture form which you can find [\[here\]](https://forms.gle/zFdoxFhezXb5ETm66)

8 8 -

9 9 ## Inspiration and background information:

10 10 Throughout my academic journey in the field of ecology, I had the privilege of participating in numerous research projects and meeting extraordinary collaborators, many of whom evolved into friendships. Frequently, I found myself getting involved in discussions that pertained to ecological research, with a particular emphasis on subjects related to the spatial dimension of our field. These interactions allowed me to explore and engage with the intricate aspects of science, broadening my perspective and deepening my understanding of this dynamic discipline. Once I had the opportunity to debate with my classmate Yves. He was noticeably vexed by an assignment for our master's program, and his reaction led to a rather comical moment of exasperation for him and very inspirational for me: "I don't understand this. I used to do ecology I thought I would be in the jungle or in a natural ecosystem collecting data or observing plants and animals, not programming and statistics". Since then I met amazing ecologists, very skilled in the field-work but also frustrated with the amount of statistical knowledge they had to cope with. Over a course of four years dedicated to this field, I began to discern that there is a broader difficulty and frustration among junior ecologists to deal with the computational aspect of our field. The way I approach this observation is relevant to advancements and changes that have occurred in the last decade. Worldwide data volume doubled nine times between 2006 and 2011 with this trend to continue.

Four files have been modified:

- Yellow for changes
- Green for additions
- Red for deletion

Summary (required)

Description

Commit to main



Current Repository
WorOptReproSciEcoEnvSci

Changes 4 History poster/UOA_lecture_poster_to_print.png

4 changed files

- .DS_Store
- poster/UOA_lecture_poster_to_print.png
- presentation/Presentation1.pptx
- README.md

This file has been deleted

Summary (required)

Description

Commit to main

Current Branch
main

Fetch origin
Last fetched 2 minutes ago

Deleted

ΤΟ ΕΡΓΑΣΤΗΡΙΟ ΔΙΑΧΕΙΡΙΣΗΣ ΒΙΟΠΟΙΚΙΛΟΤΗΣ
ΣΑΣ ΠΡΟΣΚΑΛΕΙ ΣΤΗ ΔΙΑΛΕΙΞΗ

**ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΠΛΗΡΟΦΟΡΙΩΝ ΚΑΙ ΔΕΔΟΜΕΝΩΝ
ΣΤΙΣ ΠΕΡΙΒΑΛΛΟΝΤΙΚΕΣ ΕΠΙΣΤΗΜΕΣ**

ΔΟΜΗ ΦΑΚΕΛΩΝ, ΚΑΝΟΝΕΣ ΣΥΝΤΑΞΗΣ ΜΕΤΑΔΕΔΟΜΕΝΩΝ
ΚΑΙ ΔΙΑΧΕΙΡΙΣΗΣ PROJECT ΜΕΣΩ GITHUB
ΓΙΑ ΜΙΑ (ΠΙΟ) ΑΝΟΙΚΤΗ ΚΑΙ ΑΝΑΠΑΡΑΣΙΜΗ ΕΠΙΣΤΗΜΗ

ΠΕΜΠΤΗ, 26 ΟΚΤΩΒΡΙΟΥ 12:00

ΜΙΚΡΟΥΠΟΛΟΓΙΣΤΙΚΟ ΕΡΓΑΣΤΗΡΙΟ, ΙΣΟΓΕΙΟ, ΤΜΗΜΑ ΠΕΡΙΒΑΛΛΟΝΤΟΣ

ΠΕΡΙΣΣΟΤΕΡΕΣ ΠΛΗΡΟΦΟΡΙΕΣ

ONLINE ΕΓΓΡΑΦΗ

ECO SYSTEM ECOLOGY



Current Repository **WorOptReproSciEcoEnvSci**

Current Branch **main**

Fetch origin
Last fetched 5 minutes ago

Changes 4 History presentation/Presentation1.pptx

4 changed files

.DS_Store

poster/UOA_lecture_poster_to_print.png

presentation/Presentation1.pptx +

README.md

What about presentations?

This binary file has changed.

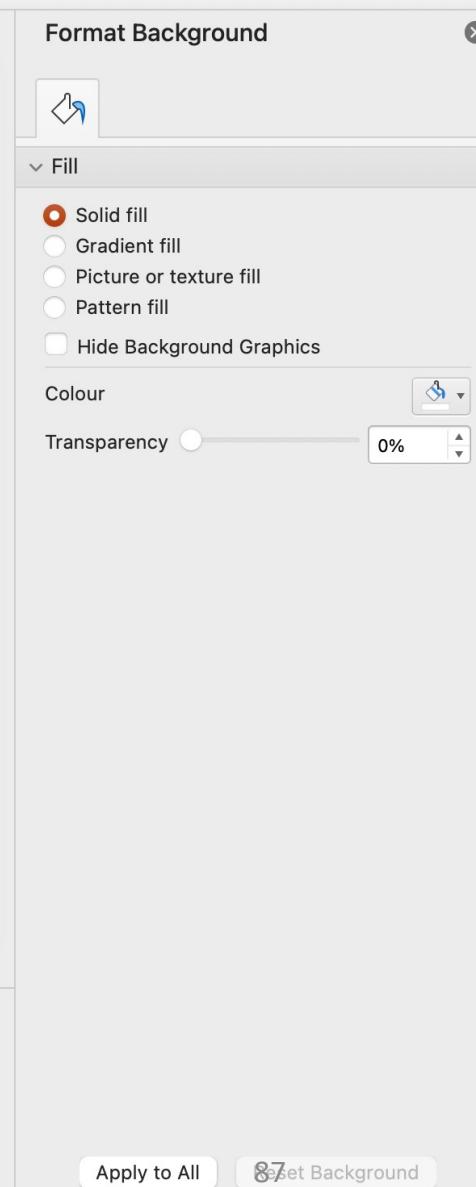
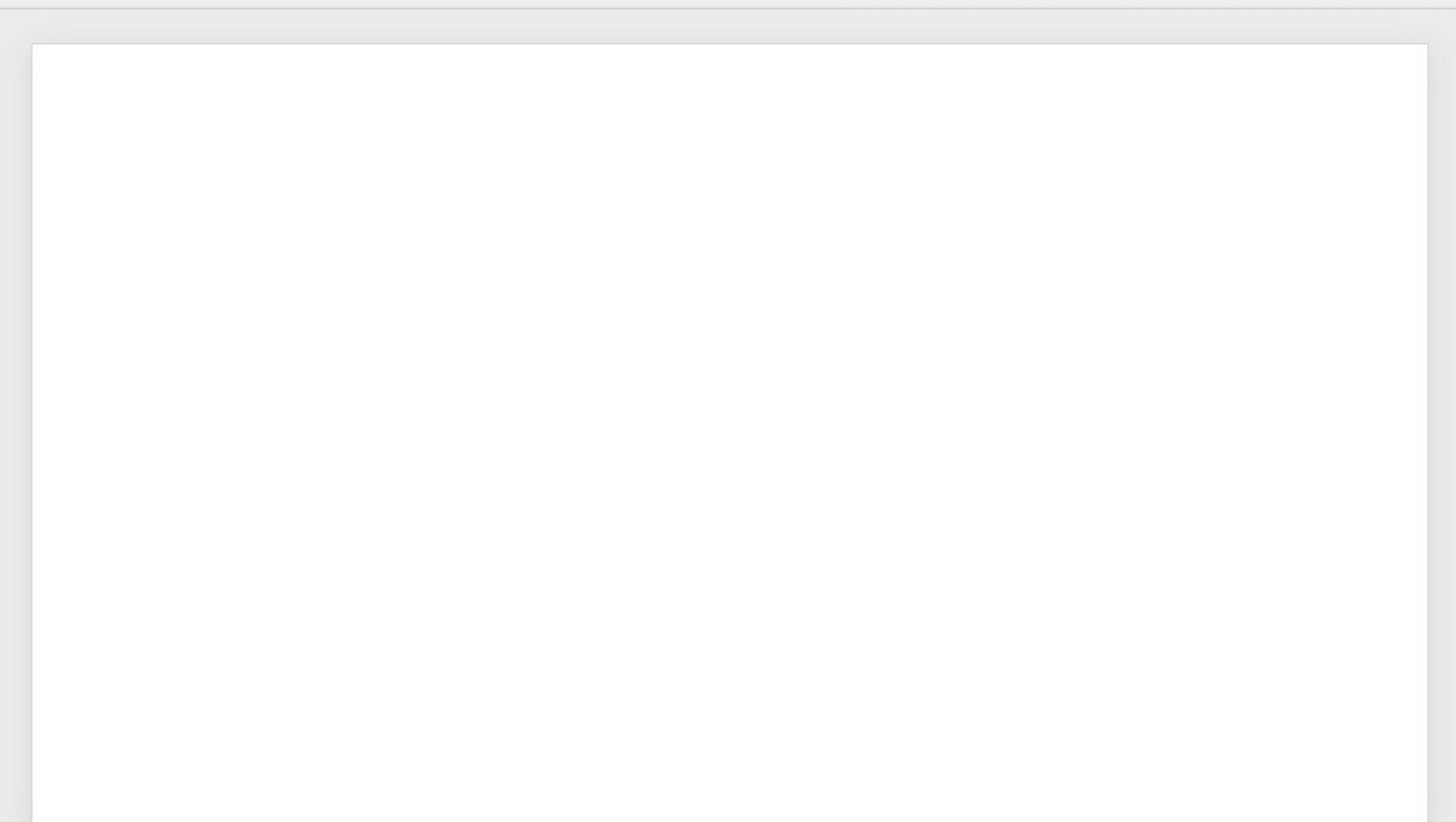
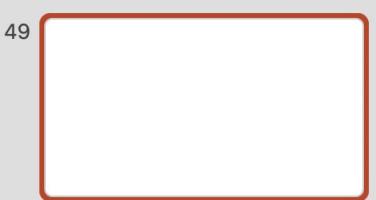
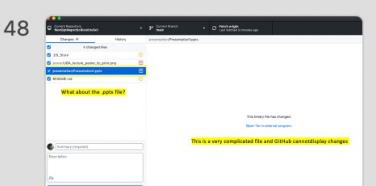
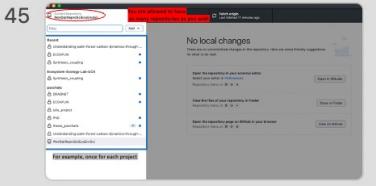
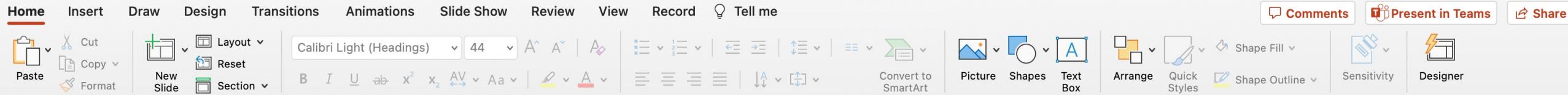
[Open file in external program.](#)

Very complicated file to track line by line changes

Summary (required)

Description

Commit to main





Current Repository **WorOptReproSciEcoEnvSci**

Current Branch **main**

Fetch origin
Last fetched 27 minutes ago

Changes 2 History Select Branch to Compare...

update readme
Paschalis Chatzopoulos • Oct 6, 2023

local path
Paschalis Chatzopoulos • Oct 5, 2023

ign
Paschalis Chatzopoulos • Oct 3, 2023

Merge pull request #6 from paschatz/pasc...
Paschalis Chatzopoulos • Oct 3, 2023

script comments
Paschalis Chatzopoulos • Oct 3, 2023

Update README.md
Paschalis Chatzopoulos • Oct 3, 2023

Merge pull request #5 from paschatz/pasc...
Paschalis Chatzopoulos • Oct 3, 2023

add script and raw data
Paschalis Chatzopoulos • Oct 3, 2023

Merge pull request #4 from paschatz/kosth...
Paschalis Chatzopoulos • Oct 3, 2023

Update README.md
I.insulaire • Oct 3, 2023

Update README.md

Update README.md

Update README.md

I.insulaire ddf63f9 ± 1 changed file +4 -4 ⚙

File	Line	Change Type	Text
README.md	18	+	Once you have set up R and the IDE of your choice, it is time to setup GitHub. First you should create your personal account here . As a student you have free access to GitHub premium which can be obtained with our academic email address. I highly recommend you to take advantage of the premium services as soon as you can follow courses and have free access to GitHub copilot. Remember that GitHub can also serve as a portfolio or CV. Very often, when you apply for jobs, you can showcase your coding skills by linking your GitHub account to your CV. For many job openings, having experience with Git and GitHub is considered a significant advantage. That definitely goes beyond its intended purpose, but in a highly competitive environment, it's always good to be prepared. In case you are having trouble, you can take a look at this video on how to sign up.
	19	## GitHub desktop	
	20	-	Assuming you have signed up for GitHub and you claimed your account, now you can download the desktop application . GitHub Desktop offers an easy-to-use and navigable working environment that doesn't require advanced coding skills. You can find a nice introduction on how to install GitHub desktop and set up your first README files here .
	20	+	Assuming you have signed up for GitHub and claimed your account, you can now download the desktop application . GitHub Desktop offers an easy-to-use and navigable working environment that doesn't require advanced coding skills. You can find a nice introduction on how to install GitHub desktop and set up your first README files here .
	21	## Further reading:	
	22		

Any moment changes can be reverted and return to a previous version

Track changes made by other users

WorOptReproSciEcoEnvSci - main - RStudio

participatin_sites.R

```
1 # Load packages
2 library(ggplot2) # to produce the map
3 library(maps)
4 library(tidyverse) # manipulate data
5 library(readxl) # read raw data from .xlsx file
6 library(leaflet) # alternative map preview
7
8 # Load file with coordinates
9 # example with paths:
10 absolute_path_KTH <- "/user/theodorou/documents/WorOptReproSciEcoEnvSci/raw_data"
11
12 absolute_path_PCH <- "~/Library/CloudStorage/OneDrive-uca.es/WorOptReproSciEcoEnvSci/raw
13
14 # Relative path
15 file_path <- "raw_data/participating_experiments.xlsx"
16 file <- read_xlsx(file_path, sheet = 2, na = "NA")
17
18 # Clean data
19 file$longitude <- as.numeric(file$longitude)
20 file$latitude <- as.numeric(file$latitude)
21 str(file)
22 coord <- file %>%
23   select(latitude, longitude)
24
25 # Plot
26 world_map <- ggplot() +
27   borders("world", fill = "white", colour = "red") +
28   geom_point(data = file, aes(longitude, latitude), size = 2) +
29   coord_fixed(1.3) +
30   theme_void()
```

12:1 (Top Level) R Script

Environment History Connections Git Tutorial

Import Dataset 130 R Global Environment

Console Background Jobs

R 4.3.0 · ~/Library/CloudStorage/OneDrive-uca.es/WorOptReproSciEcoEnvSci/

K is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |

Files Plots Packages Help Viewer Presentation

New Folder New Blank File Delete Rename More

Home Library CloudStorage OneDrive-uca.es WorOptReproSciEcoEnvSci raw_data

Name	Size	Modified
..		
participating_experiments.xlsx	112.3 KB	Oct 19, 2023, 2:12 PM
README.md	285 B	Oct 3, 2023, 5:46 PM

RStudio: Review Changes

Changes History main Stage Revert Ignore

Pull Push

participatin_sites_map.R x

Source on Save .DS_Store

Load packages
library(ggplot2) # to produce the map
library(maps)
library(tidyverse) # manipulate data
library(readxl) # read raw data
library(leaflet) # alternative rendering

Load file with coordinates
example with paths:
absolute_path_KTH <- "/user/theo/.../
absolute_path_PCH <- "~/Library/.../

Relative path
file_path <- "raw_data/participatin_sites_map.xlsx"
file <- read_xlsx(file_path, sheet = 1)

Clean data
file\$longitude <- as.numeric(file\$longitude)
file\$latitude <- as.numeric(file\$latitude)
str(file)
coord <- file %>%
 select(latitude, longitude)

Plot
world_map <- ggplot() +
 borders("world", fill = "white")
 geom_point(data = file, aes(longitude, latitude))
 coord_fixed(1.3) +
 theme_void()

(Top Level)

Environment History Connections Git Tutorial

Diff Commit Pull Push History

Staged Status Path .DS_Store

Commit message

Show Staged Unstaged Context 5 lines Ignore Whitespace Stage All Discard All

Amend previous commit Commit

Similar working environment with GitHub desktop

Size Modified

112.3 KB Oct 19, 2023, 2:12 PM
285 B Oct 3, 2023, 5:46 PM

Go to file/function

RStudio: Review Changes

Changes History main (all commits) | C

Search Pull

participatin_sites.R

Source on Save | 🔍 | ✎ | 📁

```

1 # Load packages
2 library(ggplot2) # to produce t
3 library(maps)
4 library(tidyverse) # manipulate
5 library(readxl) # read raw data
6 library(leaflet) # alternative
7
8 # Load file with coordinates
9 # example with paths:
10 absolute_path_KTH <- "/user/the
11 absolute_path_PCH <- "~/Library
12
13
14 # Relative path
15 file_path <- "raw_data/particip
16 file <- read_xlsx(file_path, she
17
18 # Clean data
19 file$longitude <- as.numeric(fil
20 file$latitude <- as.numeric(file
21 str(file)
22 coord <- file %>%
23   select(latitude, longitude)
24
25 # Plot
26 world_map <- ggplot() +
27   borders("world", fill = "white")
28   geom_point(data = file, aes(lo
29   coord_fixed(1.3) +
30   theme_void()

```

(Top Level) ▾

Environment History Connections Git Tutorial

Diff | Commit | Pull | Push | History

Staged Status Path .DS_Store

Commit History:

- HEAD -> refs/heads/main (origin/main, origin/HEAD) Merge branch 'main' of https://github.com/paschatz Paschalis Chatzopoulos <58082380+paschatz@2023-10-20 0788cc74
- update script
- Update README.md
- poster
- update readme
- local path
- paschatz/issue2 ign
- Merge pull request #6 from paschatz/paschalis_r_script
- origin/paschalis_r_script paschalis_r_script script comments
- Update README.md
- Merge pull request #5 from paschatz/paschalis_r_script
- add script and raw data

Commits 1–33 of 33

Commit Details:

SHA	0788cc743440b98ebc6f4788a0a77337ec409755
Author	Paschalis Chatzopoulos <58082380+paschatz@users.noreply.github.com>
Date (UTC)	2023-10-20 09:34
Subject	Merge branch 'main' of https://github.com/paschatz/WorOptReproSciEcoEnvSci
Parent	3362d32147c61638ecd020bb5a1c4e6264c95c1a fbfa10e5321329d8273f6e12929c0e90e2cbdd79

Diff Viewer:

README.md

```

@@ -1,6 +1,6 @@
1 # Workflow optimisation for reproducible science in ecology and environmental science
2 1 # Information and data architecture in environmental sciences: Folder structure, metadata and project
3 management through GitHub
4 2
5 3 This serves as an open-access repository for the forthcoming lecture entitled 'Workflow Optimization for
6 Reproducible Science in Ecology and Environmental Science,' scheduled to be hosted by the University of
7 the Aegean on October 26, 2023.
8 3 This serves as an open-access repository for the forthcoming lecture entitled 'Information and data
9 architecture in environmental sciences: Folder structure, metadata and project management through
10 GitHub,' scheduled to be hosted by the University of the Aegean on October 26, 2023.
11 4
12 5 Within this repository, you can find all the course materials, references, and lecture handouts.
13 Additionally, I envision this repository to serve as a practical demonstration of building and
14 structuring a course from the ground up in an open and reproducible environment.
15 6
16 @@ -39,4 +39,4 @@ Here, you can access a list of references and a compilation of courses I have pr
17 39 - [How to use R Markdown ](https://www.rforecology.com/post/how-to-use-rmarkdown-part-one/?fbclid=IwAR0Axz3yb5Ql0Z0Koz6baJnN3P0a6ut4FNo1jbJP0B4NWXeyNSVMXwMZDem)
18 40 - [Tidy data for efficiency, reproducibility, and collaboration](https://openscapes.org/blog/2020-10-12-tidy-data/)
19 41 - [Some good practices for research with R](https://github.com/etiennebacher/good-practices)

```

Size Modified

112.3 KB Oct 19, 2023, 2:12 PM
285 B Oct 3, 2023, 5:46 PM

WorOptReproSciEcoEnvSci

91

Update Nov-2023: Rstudio V2023.09.1+494.pro2

The screenshot shows the RStudio interface with the GitHub Copilot feature enabled. The code editor on the left contains R script code, and the options panel on the right is titled "Copilot". A red circle highlights the status bar at the bottom of the code editor, which displays the message "Copilot: No completions available." Below the status bar, there are three buttons: "OK", "Cancel", and "Apply".

GitHub Copilot (Preview)

Enable GitHub Copilot
You are currently signed in as: paschatz [Sign Out](#)

Index project files with GitHub Copilot

Copilot Completions

Show code suggestions after keyboard idle (ms):

This feature is in preview. If you'd like to provide feedback or report an issue, please [file an issue](#) on the RStudio GitHub repository.

By using GitHub Copilot, you agree to abide by their terms of service.

GitHub Copilot: Terms of Service

General

Code

Console

Appearance

Pane Layout

Packages

R Markdown

Python

Sweave

Spelling

Git/SVN

Publishing

Terminal

Accessibility

Copilot

Copilot: No completions available. R Script

The screenshot shows a code editor interface with two main panes displaying R scripts. A red box highlights the left pane, and a green box highlights the right pane. A red circle highlights the line `select(latitude, longitude)` in the left pane.

Left Pane (Red Box):

```
scripts > participatin_sites_map.R
1 # Load packages
2 library(ggplot2) # to produce the map
3 library(maps)
4 library(tidyverse) # manipulate data
5 library(readxl) # read raw data from .xlsx file
6 library(leaflet) # alternative map preview
7
8 # Load file with coordinates
9 # example with paths:
10 absolute_path_KTH <- "/user/theodorou/documents/WorOptReproSciEcoEnvSci/raw_data"
11
12 absolute_path_PCH <- "~/Library/CloudStorage/OneDrive-uca.es/WorOptReproSciEcoEnvSci/raw
13
14 # Relative path
15 file_path <- "raw_data/participating_experiments.xlsx"
16 file <- read_xlsx(file_path, sheet = 2, na = "NA")
17
18 # Clean data
19 file$longitude <- as.numeric(file$longitude)
20 file$latitude <- as.numeric(file$latitude)
21 str(file)
22 coord <- file %>%
23   select(latitude, longitude) # Line highlighted by red circle
24
25 # Plot
26 world_map <- ggplot() +
27   borders("world", fill = "white", colour = "red") +
28   geom_point(data = file, aes(longitude, latitude), size = 2) +
29   coord_fixed(1.3) +
30   theme_void() +
```

Right Pane (Green Box):

```
# Load packages
library(ggplot2) # to produce the map
library(maps)
library(tidyverse) # manipulate data
library(readxl) # read raw data from .xlsx file
library(leaflet) # alternative map preview

# Load file with coordinates
# example with paths:
absolute_path_KTH <- "/user/theodorou/documents/WorOptReproSciEcoEnvSci/raw_data"

absolute_path_PCH <- "~/Library/CloudStorage/OneDrive-uca.es/WorOptReproSciEcoEnvSci/raw

# Relative path
file_path <- "raw_data/participating_experiments.xlsx"
file <- read_xlsx(file_path, sheet = 2, na = "NA")

# Clean data
file$longitude <- as.numeric(file$longitude)
file$latitude <- as.numeric(file$latitude)
str(file)
coord <- file %>%
  select(latitude, longitude) # Line highlighted by red circle You, 6 seconds ago • Uncommitted changes
# Plot
world_map <- ggplot() +
  borders("world", fill = "white", colour = "red") +
  geom_point(data = file, aes(longitude, latitude), size = 2) +
  coord_fixed(1.3) +
  theme_void() +
```

Bottom Panel:

PROBLEMS 17 **PORTS** **GITLENS** **COMMENTS**

COMMIT GRAPH: WOROPTREPROSCIECOENVSCI

WorOptReproSciEcoEnvSci > main > Fetch (58 seconds ago)

Search commits (↑↓ for history), e.g. "Updates dependencies" author:eamodio

Try the Focus Preview

BRANCH / TAG	GRAPH	COMMIT MESSAGE	AUTHOR	CHANGES	COMMIT DATE / TIME	SHA
✓ main		Merge branch 'main' of https://github.com/pasc... update script Update README.md update title poster update readme local path	You	3 1 3 2 1	yesterday 3 days ago 2 weeks ago 2 weeks ago	0788cc7 3362d32 fbfa10e 6042696 ebd1f29 a82f36a

COMMIT GRAPH DETAILS

Commit 0788cc7 Working Changes

You committed yesterday

Merge branch 'main' of https://github.com/pascatz/WorC

AUTOLINKS 0 found

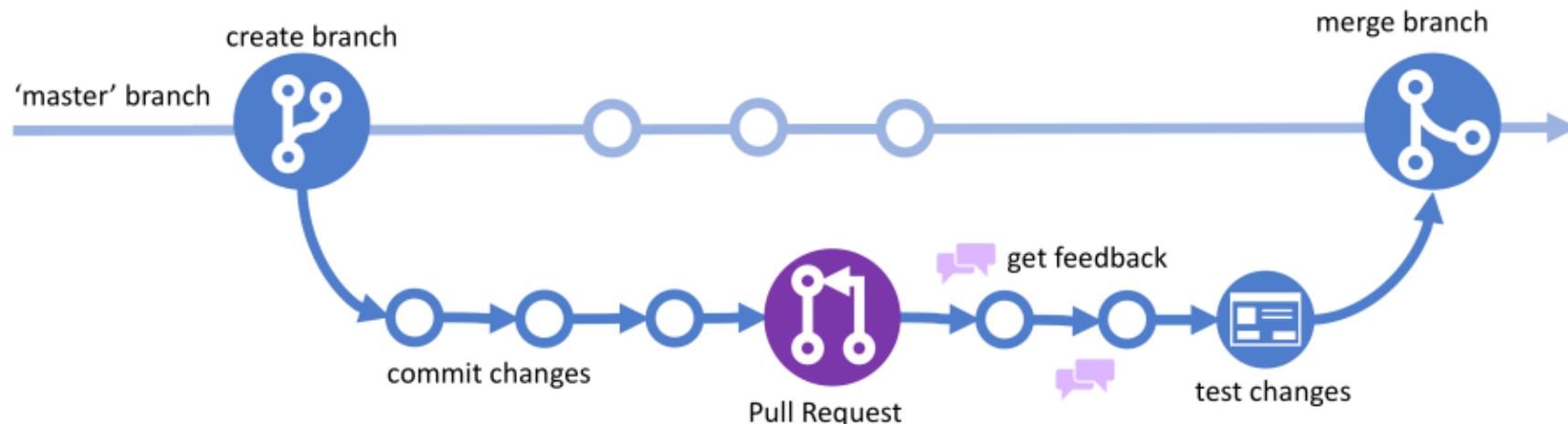
Use autolinks to linkify external references, like Jira issues or Zendesk tickets, in commit messages.

FILES CHANGED +0 ~1 -0

Branches

Let's assume you have been assigned to run an alternative *post-hoc* test (Dunnett instead of Tukey's)

- Everything in your code remains the same apart from the R function
 - *DunnettTest()* instead of *HSD.test()*
- Without affecting the existing script, you create a branch
- Make changes
- Evaluated by your supervisor/ collaborator
- Return to the main branch





main 2 branches 0 tags

Go to file Add file Code

Switch branches/tags [Find or create a branch...](#)

Branches Tags

main (default) review_2

View all branches

	scripts	update graphs_fix code spaces
	tables	update graphs_fix code spaces
	.DS_Store	Update .DS_Store
	.gitignore	organise files
	PREMON_allometry.Rproj	organise files
	README.md	Update README.md

README.md

Understanding palm forest carbon dynamics through height-diameter allometry, Puerto Rico

This is a publicly accessible implementation of our study, providing the R-script containing all statistical analyses, figures and tables. Upon the acceptance of our publication, we will also upload the raw data to online databases for further accessibility.

Abstract:

Tropical forests are major components of Earth's carbon stocks, but their diversity and structural complexity challenge our ability to accurately estimate carbon stocks and dynamics. Palms, in particular, are prominent components of many tropical forests that have unique anatomical, physiological, and allometric differences from dicot trees, which impede accurate estimates of their above-ground biomass (AGB). In this study, we focused on improving height estimates and, ultimately, AGB estimates for a highly abundant palm in Puerto Rico, *Prestoea acuminata*. Based on field measurements of 1,215 individuals, we found a strong relationship between stem height and diameter at breast height. We also found evidence that height-diameter allometry of *P. acuminata* is mediated

About

No description, website, or topics provided.

Readme

Activity

0 stars

2 watching

0 forks

Releases

No releases published
[Create a new release](#)

Packages

No packages published
[Publish your first package](#)

Contributors 2

paschatz Paschalis Chatzopoulos

bobmuscarella Bob Muscarella

Languages

R 100.0%

University < ⚡ 🔍

github.com

paschatz

Overview Repositories 14 Projects Packages Stars

Popular repositories

Group-Project-Acid-deposition-in-Sweden Public

This is a public version of our repository for our assignment for the course of Ecological methods.

R

WorOptReproSciEcoEnvSci Public

Welcome to the "Information and data architecture in environmental sciences: Folder structure, metadata and project management through GitHub" repository, scheduled to be hosted by the University o...

R

Customize your pins

209 contributions in the last year

Contribution settings ▾

Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct

Mon Wed Fri

Learn how we count contributions

Less More

Contribution activity

2023

October 2023

Created 35 commits in 3 repositories

paschatz/WorOptReproSciEcoEnvSci 32 commits

paschatz/Understanding-palm-forest-ca... 2 commits

paschatz/julia_project 1 commit

2022

2021

2020

2019

Highlights PRO

Organizations

paschatz/WorOptReproSciEcoEnvSci Oct 2 96

Paschalis Chatzopoulos
paschatz

Currently based in Cadiz, Spain. PhD candidate at
@Ecosystem-Ecology-Lab-UCA.

Edit profile

8 followers · 13 following

Cadiz university

Spain

@pascal_161

Highlights PRO

Organizations



hadley

Overview Repositories 331 Projects Packages Stars 114 Sponsoring 2



Hadley Wickham
hadley · he/him

Unfollow

Chief Scientist at **@posit-pbc**

24.9k followers · 0 following

 **@posit-pbc**
 Houston, TX
 05:38 - 8h behind
 hadley@posit.co
 <http://hadley.nz>
 @hadleywickham@fosstodon.org

Sponsoring

Pinned

 **tidyverse/ggplot2** Public
An implementation of the Grammar of Graphics in R
R 6.1k 2k

 **tidyverse/dplyr** Public
dplyr: A grammar of data manipulation
R 4.6k 2.1k

 **tidyverse/tidyverse** Public
Easily install and load packages from the tidyverse
R 1.5k 274

 **r4ds** Public
R for data science: a book
R 4.1k 4k

 **r-lib/devtools** Public
Tools to make an R developer's life easier
R 2.3k 754

 **adv-r** Public
Advanced R: a book
TeX 2.3k 1.7k

4,286 contributions in the last year

Less  More 

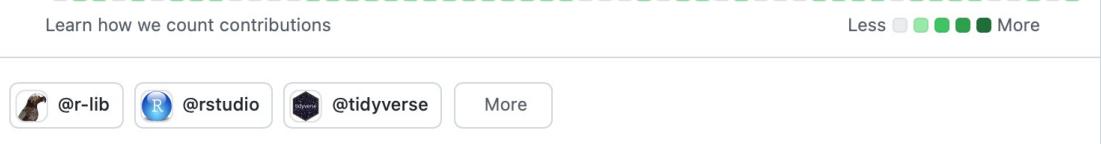
Learn how we count contributions

2023

Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct

Mon 

Wed 

Fri 

2022

2021

2020

2019

2018

2017

2016

97

Activity overview

Contributed to **rstudio/renv**, **rstudio/rsconnect**, **r-lib/testthat** and 284 other repositories

15% Code review

A screenshot of a GitHub repository page for `tidyverse / ggplot2`. The page shows the repository's activity, releases, contributors, and other metadata.

Repository Summary:

- Code:** 219
- Issues:** 219
- Pull requests:** 50
- Actions:**
- Wiki:**
- Security:**
- Insights:**

Code Tab (Active):

- Branch:** main
- Branches:** 71 branches
- Tags:** 40 tags
- Commits:** 5,155 commits (last week)
- Latest Commit:** bf0c64f last week
- Recent Activity:**
 - Do not require Hmisc on the CI with R < 4.1 (#5237) - 7 months ago
 - Merge in v3.4.4 to main (#5480) - last week
 - Update faithful.R (#5015) - last year
 - Release 3.4.0 (#5025) - last year
 - seq function calls should use full name of 'length.out' (#3817) - 4 years ago
 - update CITATION to new style - 8 months ago
 - Merge in v3.4.4 to main (#5480) - last week
 - Use retina logo and generate favicons - 5 years ago
 - Merge in v3.4.4 to main (#5480) - last week
 - Merge in v3.4.4 to main (#5480) - last week
 - Merge in v3.4.4 to main (#5480) - last month
 - Introduction vignette (#5159) - last year
 - Documentation improvements (#4851) - last year
 - Use union merge strategy for NEWS - last year
 - Add as.data.frame method to mapped_discrete that strips the class s... - 3 years ago
 - link to new location (#4836) - last year
 - Merge in v3.4.4 to main (#5480) - last week
 - link to new location (#4836) - last year

About Section:

- An implementation of the Grammar of Graphics in R
- ggplot2.tidyverse.org
- Tags: r, visualisation, data-visualisation
- Links:
 - Readme
 - Unknown, MIT licenses found
 - Code of conduct
 - Cite this repository
 - Activity
 - 6.1k stars
 - 309 watching
 - 2k forks
- Report repository

Releases Section:

- ggplot2 3.4.4 (Latest)
last week
- + 22 releases

Contributors Section:

- 296 contributors
- 98



Files

main

Go to file

.github

R

- aaa-.R
- aes-colour-fill-alpha.R
- aes-evaluation.R
- aes-group-order.R
- aes-linetype-size-shape.R
- aes-position.R
- aes.R**
- annotation-custom.R
- annotation-logticks.R
- annotation-map.R
- annotation-raster.R
- annotation.R
- autolayer.R
- autoplot.R
- axis-secondary.R
- backports.R
- bench.R
- bin.R
- compat-plyr.R
- coord-.R
- coord-cartesian-.R
- coord-fixed.R

ggplot2 / R / aes.R

Code Blame 464 lines (418 loc) · 13.5 KB

Raw ⌂ ⌄ ⌅ ⌆ ⌇ ⌈ ⌉ ⌊ ⌋

```
86  aes <- function(x, y, ...) {
87    xs <- arg_enquo("x")
88    ys <- arg_enquo("y")
89    dots <- enquo(...)
90
91    args <- c(xs, ys, dots)
92    args <- Filter(Negate(quo_is_missing), args)
93
94    # Pass arguments to helper dummy to throw an error when duplicate
95    # `x` and `y` arguments are passed through dots
96    local({
97      aes <- function(x, y, ...) NULL
98      inject(aes(!!!!args))
99    })
100
101   aes <- new_aes(args, env = parent.frame())
102   rename_aes(aes)
103 }
104
105 # Wrap symbolic objects in quooses but pull out constants out of
106 # quooses for backward-compatibility
107 new_aesthetic <- function(x, env = globalenv()) {
108   if (is_quosure(x)) {
109     if (!quo_is_symbolic(x)) {
110       x <- quo_get_expr(x)
111     }
112     return(x)
113   }
114
115   if (is_symbolic(x)) {
116     x <- new_quosure(x, env = env)
117     return(x)
118   }
119
120   x
121 }
122 new_aes <- function(x, env = globalenv()) {
123   check_object(x, is.list, "a .cls list")
124   x <- lapply(x, new_aesthetic, env = env)
125   structure(x, class = "uneval")
126 }
```

GitHub issues

tidyverse / ggplot2

Type ⌘ to search

Issues 206

Dismiss

Filters is:issue is:open

Labels 22

Milestones 1

New issue

206 Open ✓ 3,533 Closed

Author ▾ Label ▾ Projects ▾ Milestones ▾ Assignee ▾ Sort ▾

Issue	Author	Labels	Comments
Should legend.key inherit from panel.background?	#5549	feature	1
Feature request: colourbar as gradients	#5547	feature guides	1
Should theme() support argument splicing?	#5542	feature themes	1
Feature request: legend.byrow argument added to theme	#5541	feature guides themes	3
Feature request: legend.title.position argument added to theme	#5540	feature guides themes	6
Feature request: legend.key.spacing argument(s) added to theme	#5539	feature guides themes	2
ggplot_add.list() uses the %+% function instead of the + generic.	#5536	feature	6
First argument to scales is not longer the name of the scale	#5535	bug scales	3

Closed

read_data_entity_names of multiple package_ids #47

paschatz opened this issue last month · 7 comments



cInsmth commented last month

Contributor

...

This issue is occurring due to a change in the repository API 'Read Data Package Archive' method. The corresponding `read_data_package_archive` function has been updated and is available for immediate use by installing EDIutils with:

```
devtools::install_github(  
  repo = "rOpenSci/EDIutils",  
  ref = "refactor-read-data-package-archive"  
)
```



The function no longer uses a `transaction` identifier, so the new call becomes:

```
try_download <- read_data_package_archive(  
  packageId = "knb-lter-cdr.444.8",  
  path = "data_cleaning/11_Cedar_Creek/"  
)
```



Next steps are to update the docs, tests, and release into the development and main branches.

@paschatz, does this fix the issue? Is there anything else? Thanks again for reporting it!



paschatz commented last month

Author

...

Hey @cInsmth,

Now works smoothly!! 💯

I appreciate your support.



Reproducible science workshop

Handy for big projects

View 1

+ New View

Filter by keyword or by field

	Title	...	Assignees	...	Status	...	+
1	● add script for example #3		paschatz		Done		
2	● fix course date #1				Done		
3	● organise folders #2		paschatz		In Progress		
4	● presentation/collaboration #7		kostheo71 and paschatz				
5	● organise 'workshop' #8		paschatz		In Progress		

+ You can use Control + Space to add an item

I start writing methods, but I cannot recall the protocol for species identification

No worries, I am on top of it!

GitHub Docs

Help for wherever you are on your GitHub journey.



⚡ Get started

- [Get started](#)
- [Migrations](#)
- [Account and profile](#)
- [Authentication](#)
- [Billing and payments](#)
- [Site policy](#)

💬 Collaborative coding

- [GitHub Codespaces](#)
- [Repositories](#)
- [Pull requests](#)
- [GitHub Discussions](#)
- [GitHub Copilot](#)

⚙️ CI/CD and DevOps

- [GitHub Actions](#)
- [GitHub Packages](#)
- [GitHub Pages](#)

🛡️ Security

- [Code security](#)
- [Supply chain security](#)
- [Security advisories](#)
- [Dependabot](#)
- [Code scanning](#)
- [Secret scanning](#)

💻 Client apps

- [GitHub CLI](#)
- [GitHub Desktop](#)

周恩 Project management

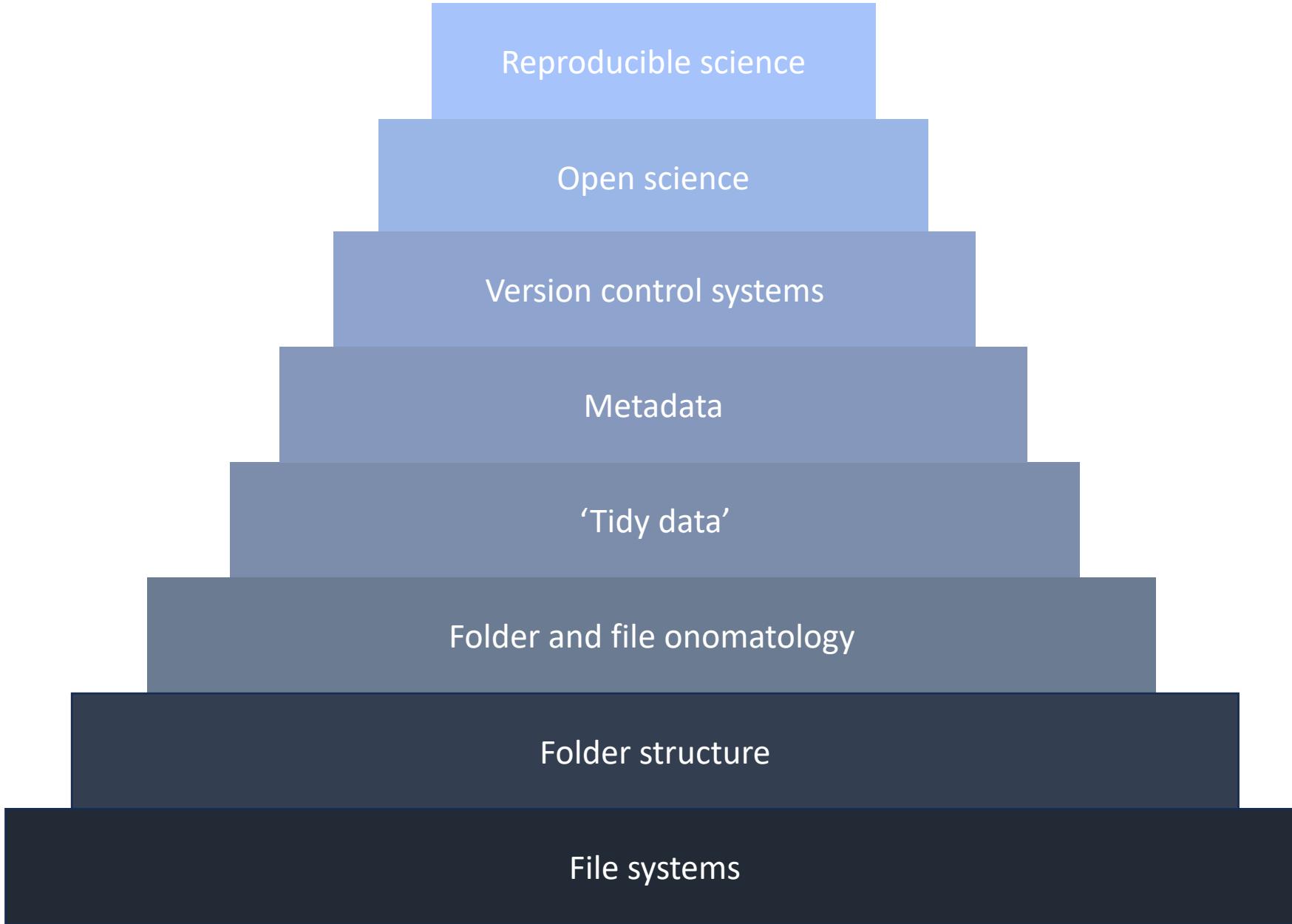
- [GitHub Issues](#)
- [Projects](#)
- [Tasklists](#)
- [Search on GitHub](#)

🔗 Developers

- [Apps](#)
- [REST API](#)
- [GraphQL API](#)
- [Webhooks](#)

🏢 Enterprise and Teams

- [Organizations](#)
- [Enterprise administrators](#)





Caitlin Rivers



Epidemiological
data in .pdf



Machine readable
files (.csv)



Publish on GitHub



[◀ News Releases](#)

FEATURE STORY

‘We Had to Get This Right’: How Johns Hopkins Built the Coronavirus Tracking Global Dashboard: An Oral History

Screenshot of the GitHub repository for CSSEGISandData / COVID-19. The repository has been archived by the owner on Mar 10, 2023, and is now read-only. It has 1,603 branches and 0 tags. The README.md file contains the following text:

```
Novel Coronavirus (COVID-19) Cases, provided by JHU CSSE
systems.jhu.edu/research/public-health...
engineering johns-hopkins-university
jhu csse 2019-ncov coronavirus
covid-19 systems-science
```

The repository has 7,691 commits, with the latest being an update on Mar 10, 2023. The repository is public and has 29.2k stars, 869 watching, and 18.7k forks. It is a report repository.

COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University

On March 10, 2023, the Johns Hopkins Coronavirus Resource Center ceased its collecting and reporting of global COVID-19 data. For updated cases, deaths, and vaccine data please visit the following sources:

- Global: [World Health Organization \(WHO\)](#)
- U.S.: [U.S. Centers for Disease Control and Prevention \(CDC\)](#)

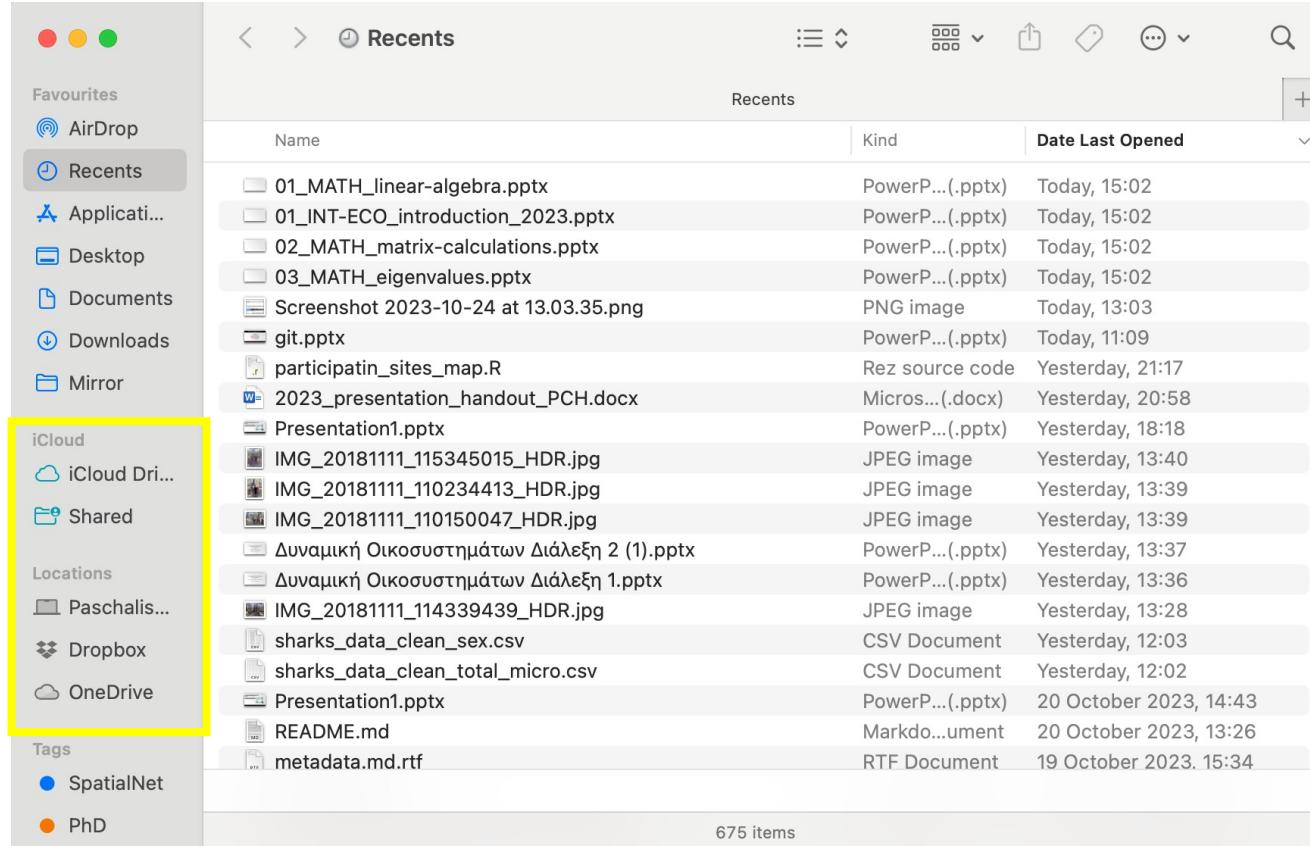
For more information, visit the [Johns Hopkins Coronavirus Resource Center](#).

This is the data repository for the 2019 Novel Coronavirus Visual Dashboard operated by the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE). Also, Supported by ESRI Living Atlas Team and the Johns Hopkins University Applied Physics Lab (JHU APL).

```
1 #install.packages("jsonlite")
2 library(jsonlite)
3 #install.packages("httpuv")
4 library(httpuv)
5 #install.packages("httr")
6 library(httr)
7
8 # Can be github, linkedin etc depending on application
9 oauth_endpoints("github")
10
11 # Change based on what you
12 myapp <- oauth_app(appname = "Coursera_John_Hopkins",
13                     key = "8758a6bf9a146e1da0c1",
14                     secret = "b9504edde46b794414495bd9c33ea28cbfd87824")
15
16 # Get OAuth credentials
17 github_token <- oauth2.0_token(oauth_endpoints("github"), myapp)
18
19 # Use API
20 gtoken <- config(token = github_token)
21 req <- GET("https://api.github.com/users/jtleek/repos", gtoken)
22
23 # Take action on http error
24 stop_for_status(req)
```

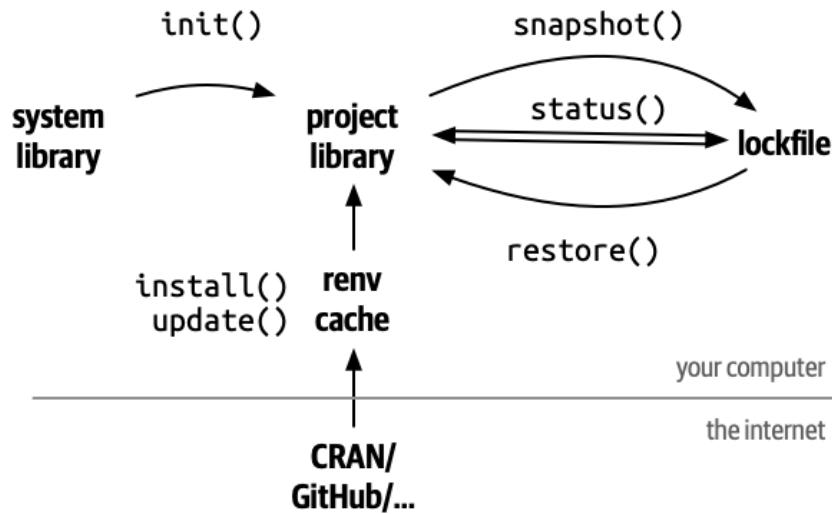
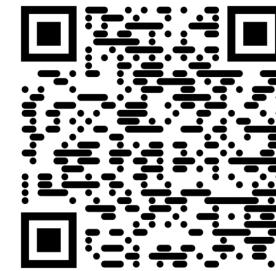


setwd() in a remote folder



- Set as local repository for GitHub a remote collaborative folder (google drive, Dropbox, iCloud)
- Allows users to benefit from version control systems, even if collaborators are not interested/participating/have the time to learn a new tool

Reproducible environment



☰ README.md

renv

lifecycle stable CRAN 1.0.3 R-CMD-check failing

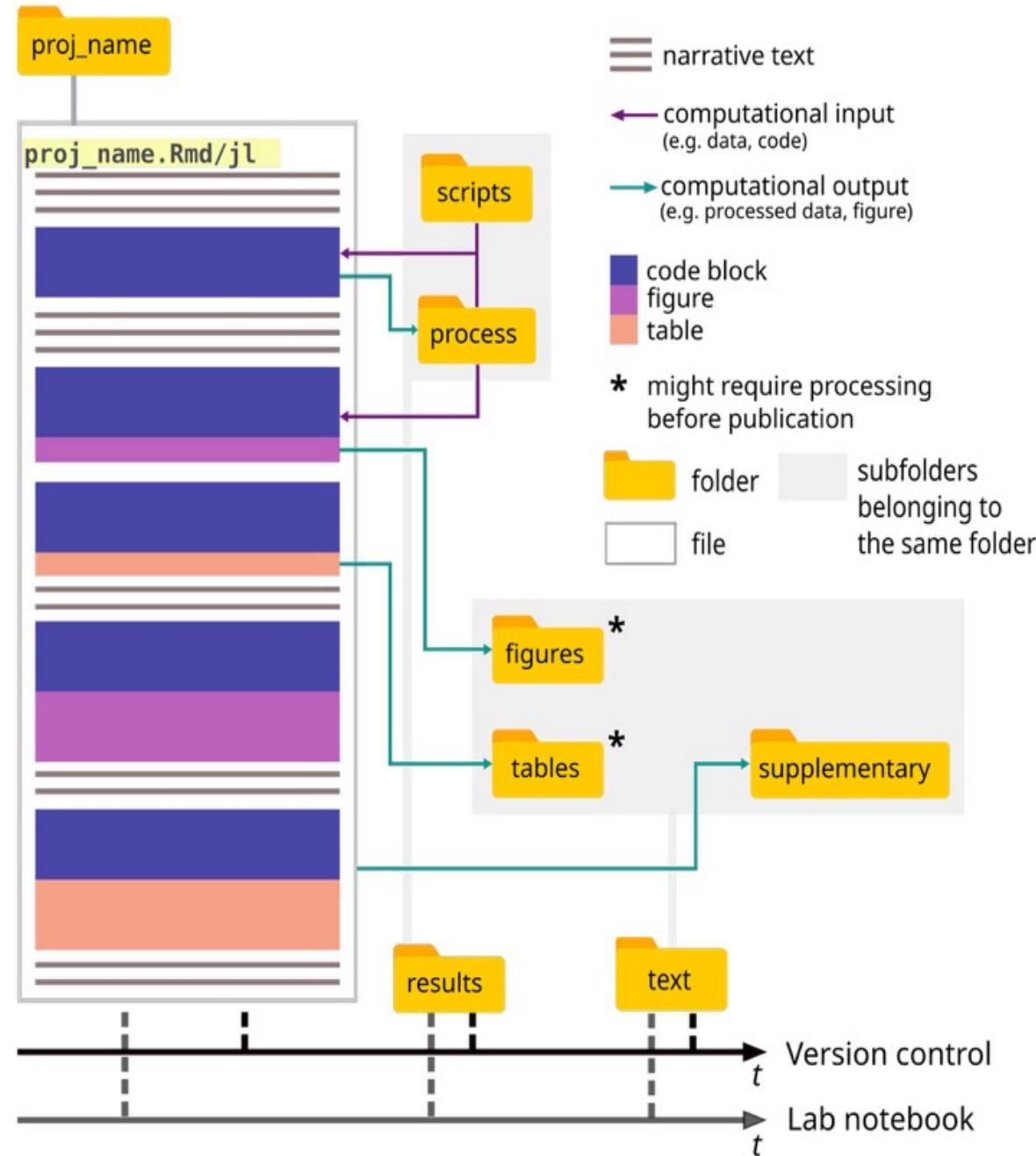


Overview

The `renv` package [\[1\]](#) helps you create reproducible environments for your R projects. Use `renv` to make your R projects more isolated, portable and reproducible.

- **Isolated**: Installing a new or updated package for one project won't break your other projects, and vice versa. That's because `renv` gives each project its own private library.
- **Portable**: Easily transport your projects from one computer to another, even across different platforms. `renv` makes it easy to install the packages your project depends on.
- **Reproducible**: `renv` records the exact package versions you depend on, and ensures those exact versions are the ones that get installed wherever you go.

```
install.packages("renv")
```



Available for:

- R
- Python
- Julia

[More info](#)

master

2 branches

5 tags

Go to file

Add file

Code



ludmillafigueiredo Docs: Update README ...

55b2517 on Sep 3 133 commits

	examples	Fix: Clean up repo	last year
	kits	fix setwd bug on set_kit.R	last year
	text/figures	Docs: File structure showing alternative branch	last month
	.gitignore	Update .gitignore	last month
	LICENSE	Fix: LICENSE had been deleted by mistake	2 years ago
	README.md	Docs: Update README	last month
	fw_sponsorship.png	Refactor: Figures (rendered and edit) in text folder; paths in README...	last year

README.md

A simple kit to use computational notebooks for more openness, reproducibility, and productivity in research

DOI [10.5281/zenodo.6977667](https://doi.org/10.5281/zenodo.6977667)

A minimal setup to get you started

This repository contains a starter-kit to use so called "computational notebooks" to structure your work. Computational notebooks are files that contain descriptive text, as well as code and its outputs, in a single, dynamic and visually appealing file (usually `.html`, but `.pdf` is also possible). Using this file to organize your work should facilitate:

- traceability: the notebook file contains all the reasoning behind the study. Moreover, the file formats

[More info](#)

About

Tools to facilitate the use of computational notebooks.

Readme

MIT license

Activity

32 stars

4 watching

8 forks

[Report repository](#)

Releases

 2.0.0 Latest
on Aug 10, 2022[+ 2 releases](#)

Packages

No packages published

Contributors

ludmillafigueiredo Ludmilla Figueiredo**z3tt** Cédric Scherer**alether** Alejandro Thérèse Navarro