schedule day 1

colomb_julien

10/25/2018

If there is not enough time:

version control on day 1 is moved to day 2 Reproducible research and data analysis II (your data) in day 2 is skipped.

09:15-10:00: Open science: the what and why.

Introduction of each participant: 1min for name, institute, reason to do science (card), exeriment to present (card with name)

From ethical values collected -> science should be collaborative, transparent and inclusive -> open science

what is open science. What is open data, who ask for it

Why open data !! - presentation

workshop outline, code of conduct, technical questions, link to collaborative notes (hack.md)

10:00-10:45

Open collaboration. Mozilla open source project. Linux

Introduction to github: get you page ready in the website. Presentation of the website, front and back

step 0: login, be part of the gitlab group, (fork the website from github?) step 1: change the yml page, pull request it. step 2: add picture, modify yml step 3: add your page (create new file), write shortly what your experiment is about. sep 4: access to the repo (admin rights)

- blogdown created
- goal: get all output of workshop there, it will be yours afterwards!

10:45-11:00

BREAK

11:00-11:45

Introduction: video, overview RDM

Open Data I (raw data) Keynote presentation: 1_0_rawdata exercise with all participants (explain the data flow) exercice in group of 3 (explain the data flow)

output: 2 sentences summary: experiment, raw data and its format

11:45-12:30

Open Data II (spreadsheets)

Keynote presentation: 1_1_spreadsheets exercise with all participants (create one spreadsheet) exercice in group of 3 (create one spreadsheet)

output: one spreadsheet design ## tools goodtables.io , openrefine (know any other?)

12:30-13:30

LUNCH

13:30-14:15

Open Data III (FAIR data)

Keynote presentation: 1_2_fairdata exercice in group of 3 (Folder organisation) exercice in group of 3 (file naming)

Output: master file explaining naming and organisation of files

14:15-15:00

Open Data IV (open data)

Keynote presentation: 1_3_open_fairdata exercise with all participants individual exercice (DMP)

Output: a data management plan (dmptools.org)

15:00-15:15

BREAK

15:15-16:00

version control

Keynote presentation: gitbasics.html exercise: clone the repository, add the DMP, link in your page, git commit, git push, pull request and accept pull request.

16:00-16:45

Day review/discussion: elevator pitch each participant (with summary of data flow) Q and A

DAY 2

09:15-10:00: version control

Keynote presentation: gitbasics.html

look at issues (mozilla, zotero), comment

exercise: clone the repository, modify your page, link in your page, git commit, git push, pull request and accept pull request.

10:00-10:45 Open Publishing (open access and licensing)

open questions What do you know ? presentation (keynote)

free to read versus free to reuse: licensing colors of open access

ouptut: add a license to your part of the website, change readme file accordingly.

10:45-11:00

BREAK

11:00-11:45 Reproduciblity (Intro)

- Good scientific practice: the official side with recap first day, (GSP.html) **cards: 1 thing forbiden**
- What is reproducibility:

presentatation_d2.html

11:45-12:30 Reproducible reports I (material and methods)

PID for everything ! open materials: RRID, MGI numbers open methods and protocols.io (+ versioning) this info must be in metadata of each experiment !!!

Key reagent table (elife)

output: - krot for one experiment - is there a protocol on protocols.io you can use ? - get a PID for your protocol (?) - add it in the metadata file for your data.

12:30-13:30

LUNCH

13:30-14:15 Reproducible research and data analysis II (comment)

reproducibleanalysis_intro.html 15 min

duty: get some data and (re-)produce a graph, with all documentation data: https://figshare.com/articles/self-learning/4201452

learning data present.
use excel or R: you have 10 min (boxplots allowed)

check the documentation with other group members: 5 min.

Let's do it with R: 7 min

add your data to the website, in a folder with your name.

14:15-15:00 Reproducible reports I (Rmarkdown)

look at Rmarkdown example

add blogpost in Rmarkdown make an analysis, either with own data or the one above use markdown for commenting

15:00-15:15 BREAK

15:15-16:00 Reproducible reports II (Rmarkdown)

add material and method to the report. get a presentation version

16:00-16:45 Day review/discussion

Give a list of more information there: Open source and leadership (Public engagement / Open Communities) documents to add Collaboration with issues open leadership (mozilla)

citizen science open science example Mozilla community GOSH)

Look at each people summary, get feedback, make changes if needed. Give the website admin rights away, publish via netify ?

DAY 2 (old version)

09:15-10:00: Reproduciblity (Intro)

- Good scientific practice: the official side with recap first day, (GSP.html) **cards: 1 thing forbiden**

question 2: what to do if you see misconduct?

- What is reproducibility: **open question** presentatation_d2.html

10:00-10:45 Reproducible research and data analysis I (exp. design)

Experimental design and p-hacking (keynote: 2_1_expdesign)

data analysis registration: osf.io

10:45-11:00

BREAK

If not done on first day, version control part comes here and everything is pushed back

11:00-11:45 Reproducible research and data analysis II (comment)

reproducibleanalysis_intro.html 15 min

duty: get some data and (re-)produce a graph, with all documentation data: https://figshare.com/articles/self-learning/4201452

learning data present.
use excel or R: you have 10 min (boxplots allowed)

check the documentation with other group members: 5 min.

Let's do it with R: 7 min

add your data to the website, in a folder with your name.

11:45-12:30 Reproducible research and data analysis II (your data)

Let's look at your data and data analysis, and do some R (or python) and documentation

12:30-13:30

LUNCH

13:30-14:15 Open Publishing (open access and licensing)

open questions What do you know ? presentation (not ready yet)

free to read versus free to reuse: licensing colors of open access

ouptut: add a license to your part of the website, change readme file accordingly.

14:15-15:00 Reproducible reports I (material and methods)

PID for everything ! open materials: RRID, MGI numbers open methods and protocols.io (+ versioning) this info must be in metadata of each experiment !!!

Key reagent table (elife)

output: - krot for one experiment - is there a protocol on protocols.io you can use ? - get a PID for your protocol (?) - add it in the metadata file for your data.

15:00-15:15 BREAK

15:15-16:00 Reproducible reports II (Rmarkdown)

Create a Rmarkdown file: material and methods, protocols, data analysis, results, figures add it to the website. add a blogpost to the website to say it worked !

16:00-16:45 Day review/discussion

Give a list of more information there: Open source and leadership (Public engagement / Open Communities) documents to add Collaboration with issues open leadership (mozilla)

citizen science open science example Mozilla community GOSH)

Look at each people summary, get feedback, make changes if needed. Give the website admin rights away, publish via netify ?

reproducible research workflow: - Easier to write reports/papers - Easier reuse of material for future projects - Helps referees to review - Builds your reputation