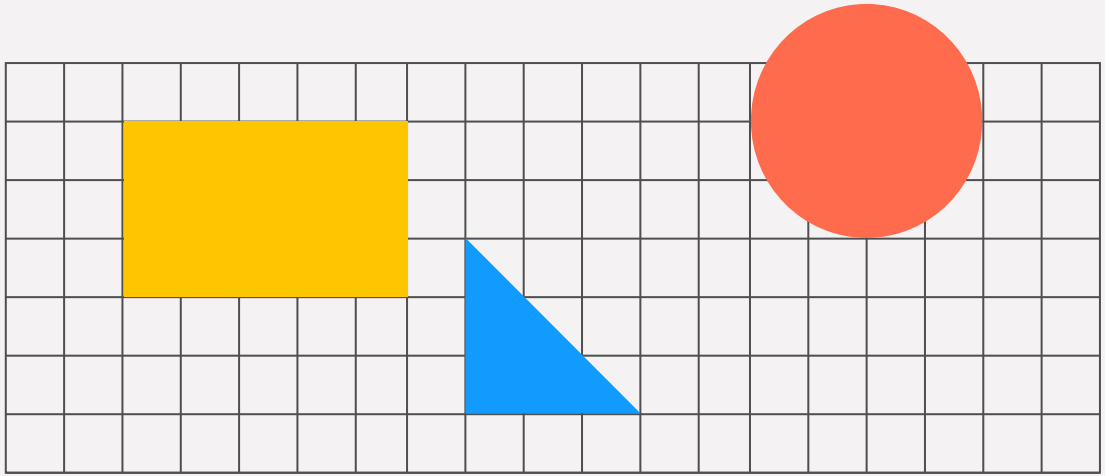
Open Science Guide for Parkinson's Research

Matthew J. Kmiecik, PhD
Hirotaka Iwaki, MD



What is Open Science?

Use the chat to write a brief sentence answering the question:

"What does Open Science mean to you?"

Agenda

https://mjff-researchcommunity.github.io/open-science-guide/

- Why pursue an open science guide?
- Definition of Open Science
- What to consider when doing Open Science (Hirotaka demo)
- Tools to help achieve Open Science
- Examples of Open Science in Parkinson's research
- How to contribute

Have a question or comment?

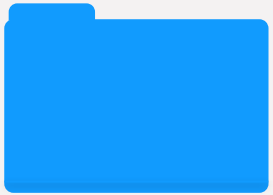
Feel free to come off mute or ask in the chat!

Tip!

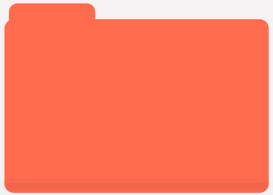
Why write a guide for Open Science?

- Where/how do I start?
- What should I consider when making data/code publicly available?
- Is there a right or wrong way to do Open Science?
- Are some tools better than others? Could I be using a better tool?
- Will someone be able to replicate my analyses given my documentation?

Principles



Considerations



Tools

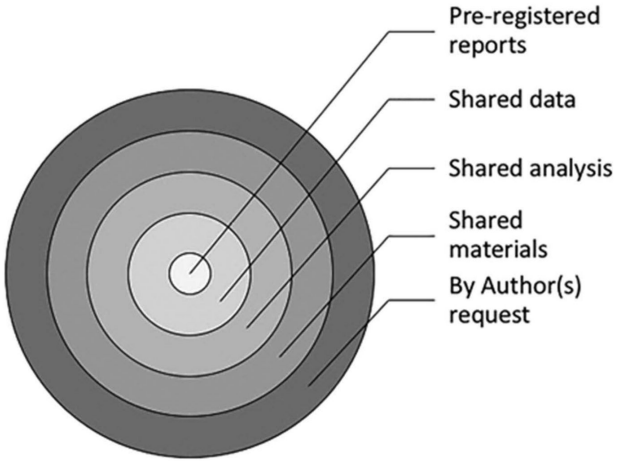
Examples

What is Open Science?

"Open Science is transparent and accessible knowledge that is shared and developed through collaborative networks."

(Vicente-Saez and Martinez-Fuentes 2018)

What is Open Science?

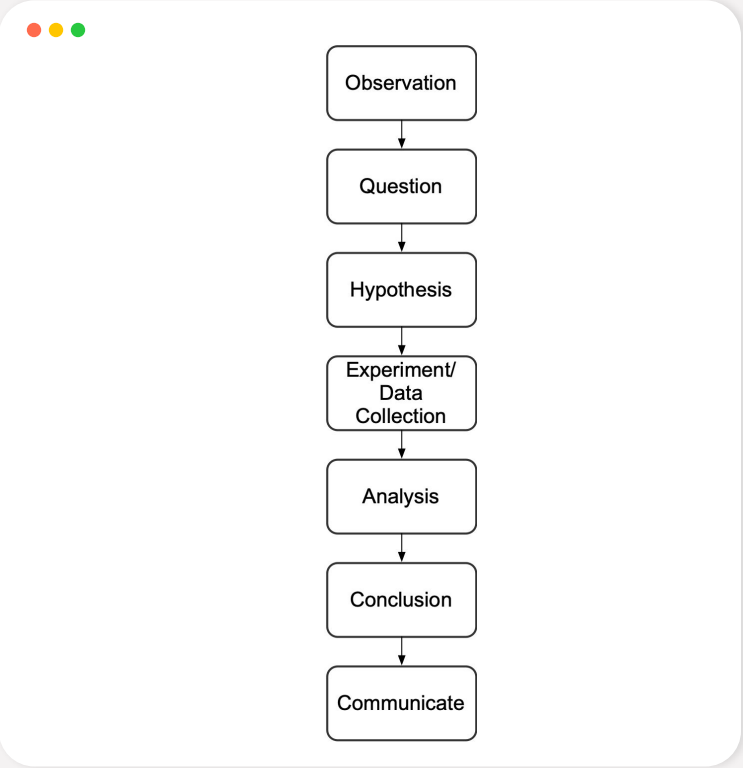


Investigators can approach Open Science in a layered, gradient approach with increasing levels of transparency and accessibility

(Bowman and Keene 2018)

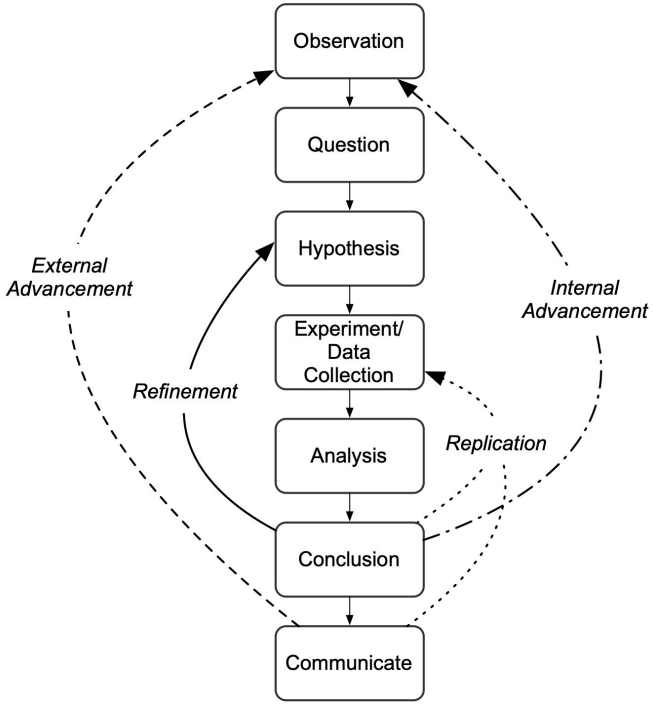
Why Open Science?

- Increased effectiveness of the scientific method
- Increased trust in scientific work
- Increased benefit to the public and common good



Why Open Science?

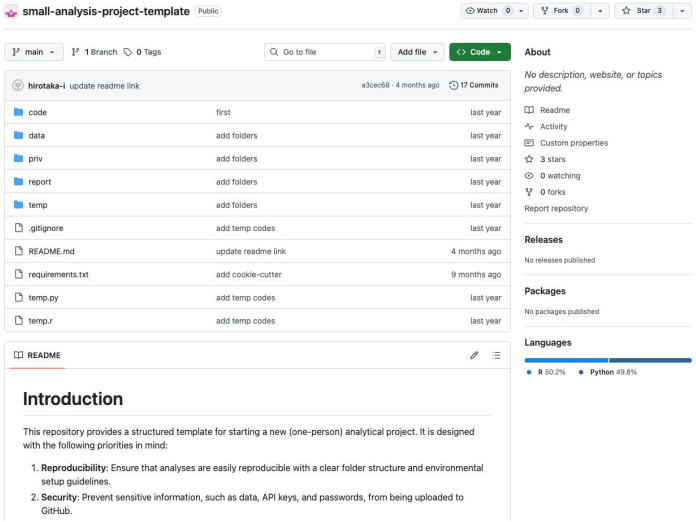
- Increased effectiveness of the scientific method
- Increased trust in scientific work
- Increased benefit to the public and common good



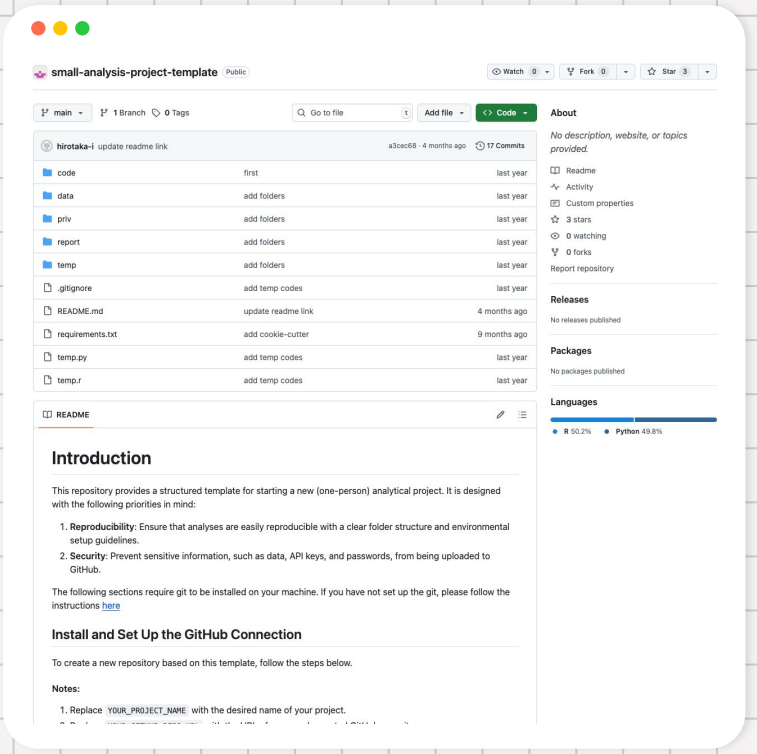
What to consider when doing Open Science?

- Data Sharing & Privacy
- Programming / Code
- Documentation
- Pre-registration
- Making Projects Citeable / Localized

Deep Dive on Organizing Project Directories by Hirotaka!

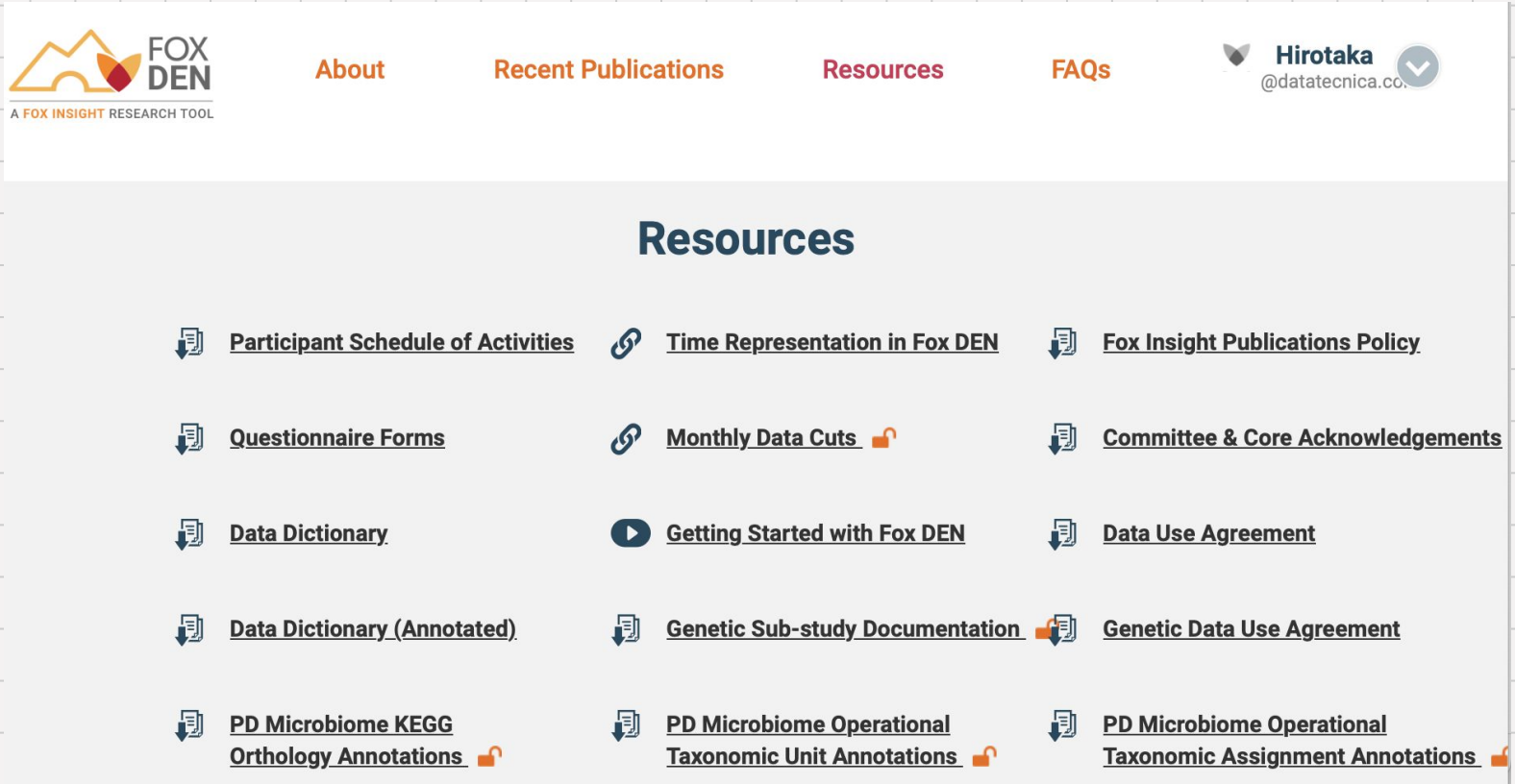


Hirotaka's Demo



https://github.com/MJFF-ResearchCommunity/small-analysis-project-template

Step1: FAIR data



Monthly Data Cuts

Data cuts are generated monthly and comprise all Fox DEN data available at that time. Data cuts can be useful for reproducibility of analysis results. A Fox DEN account is required to access these datasets.

To reference use of a Fox DEN Monthly Data Cut, please include the corresponding Archive Date and the following DOI: ***https://doi.org/10.25549/bxya-6133***

Archive Date	Download Link
August 2025	Download
July 2025	Download
June 2025	Download
May 2025	Download
April 2025	Download
March 2025	Download

FOX DEN
A FOX INSIGHT RESEARCH TOOL

Part

Que

Data

Data

PD M
Orth

FoxInsight_01Aug2025

Name	Date Modified	Size	Kind
About.csv	Aug 1, 2025 at 7:24 AM	39.4 MB	CSV Document
AcuteHealth.csv	Aug 1, 2025 at 7:24 AM	4 MB	CSV Document
Assessing_Discrimination.csv	Aug 1, 2025 at 7:24 AM	287 KB	CSV Document
Brief_Motor_Screen.csv	Aug 1, 2025 at 7:25 AM	8.2 MB	CSV Document
Cannabis_Use_in_PD.csv	Aug 1, 2025 at 7:25 AM	752 KB	CSV Document
Care_Partner_Experiences.csv	Aug 1, 2025 at 7:24 AM	1.4 MB	CSV Document
Clinical_Global_Impression_of_Change_Non_PD.csv	Aug 1, 2025 at 7:24 AM	1.4 MB	CSV Document
Clinical_Global_Impression_of_Change_PD.csv	Aug 1, 2025 at 7:24 AM	3.8 MB	CSV Document
Compensation_Strategies.csv	Aug 1, 2025 at 7:24 AM	1.3 MB	CSV Document
COVID_19_Experience_Part_2.csv	Aug 1, 2025 at 7:24 AM	3.1 MB	CSV Document
COVID_19_Experience.csv	Aug 1, 2025 at 7:25 AM	2.9 MB	CSV Document
CurrentHealth.csv	Aug 1, 2025 at 7:24 AM	20.1 MB	CSV Document
DailyActivity.csv	Aug 1, 2025 at 7:25 AM	9.1 MB	CSV Document
DailyLiving.csv	Aug 1, 2025 at 7:24 AM	9.6 MB	CSV Document
DBS_in_Early_Stage_PD.csv	Aug 1, 2025 at 7:25 AM	318 KB	CSV Document
EEQ_alcohol.csv	Aug 1, 2025 at 7:25 AM	166 KB	CSV Document
EEQ_Anti_Inflammatory_Medication_History.csv	Aug 1, 2025 at 7:25 AM	257 KB	CSV Document
EEQ_caffeine.csv	Aug 1, 2025 at 7:24 AM	688 KB	CSV Document
EEQ_calcium_channel_blocker_medication_history.csv	Aug 1, 2025 at 7:24 AM	196 KB	CSV Document
EEQ_female_reproductive_health.csv	Aug 1, 2025 at 7:25 AM	126 KB	CSV Document
EEQ_Head_Injury_or_Concussion.csv	Aug 1, 2025 at 7:25 AM	246 KB	CSV Document
EEQ_height_and_weight.csv	Aug 1, 2025 at 7:25 AM	180 KB	CSV Document
EEQ_occupation.csv	Aug 1, 2025 at 7:25 AM	188 KB	CSV Document
EEQ_pesticides_at_work.csv	Aug 1, 2025 at 7:25 AM	1.2 MB	CSV Document
EEQ_physical_activity.csv	Aug 1, 2025 at 7:25 AM	196 KB	CSV Document

Data repos

2.1.1 Where to share your data

- Zenodo
- Kaggle
- OSF (Open Science Framework)
- Dryad

If data is publicly available, provide a link to the source. If data cannot be shared, consider providing a **sample dataset** in the repository.

Step2: Coding for open science

Start with clear folder structure and instructions for everyone

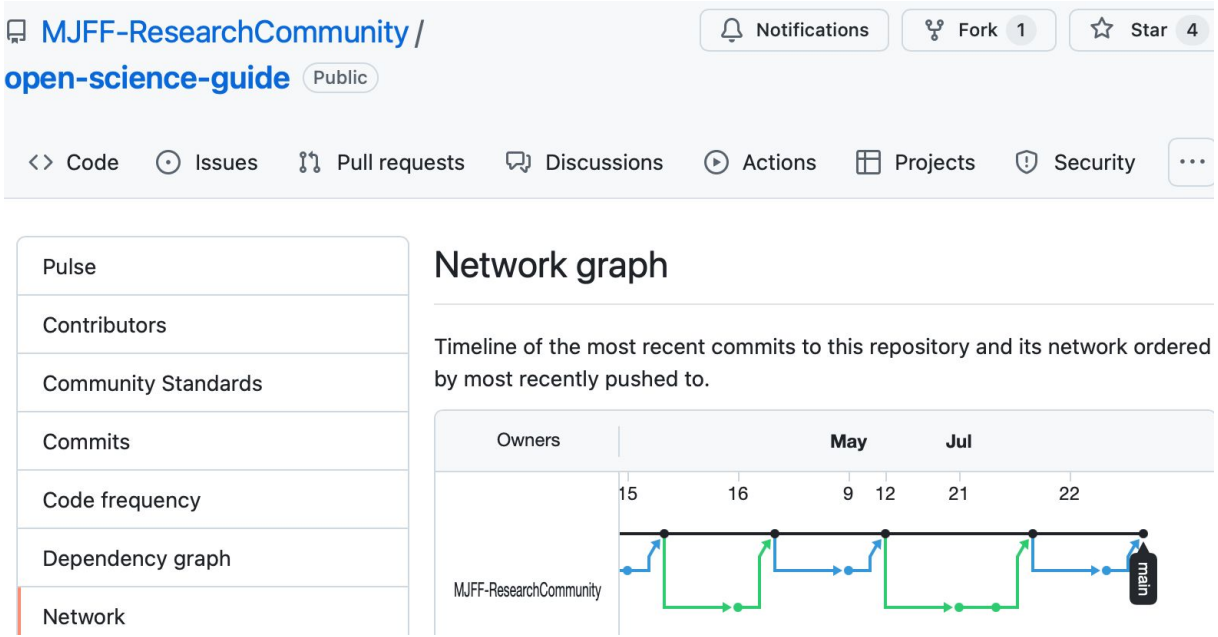
2.2.1 Project Folder Structure

A well-structured project is **transparent, reproducible, and reusable**. A clear and consistent folder structure makes collaboration easier and ensures reproducibility. Here's a basic template for a data science project:

```
├── data/           # Raw & processed datasets
├── scripts/        # Code and analysis scripts
├── results/        # Figures, tables, and outputs
├── docs/           # Documentation and notes
├── env/            # Dependency files (requirements.txt, environment.yml)
├── README.md       # Project overview
└── LICENSE         # License for open-source sharing
```

Step2: Coding for open science

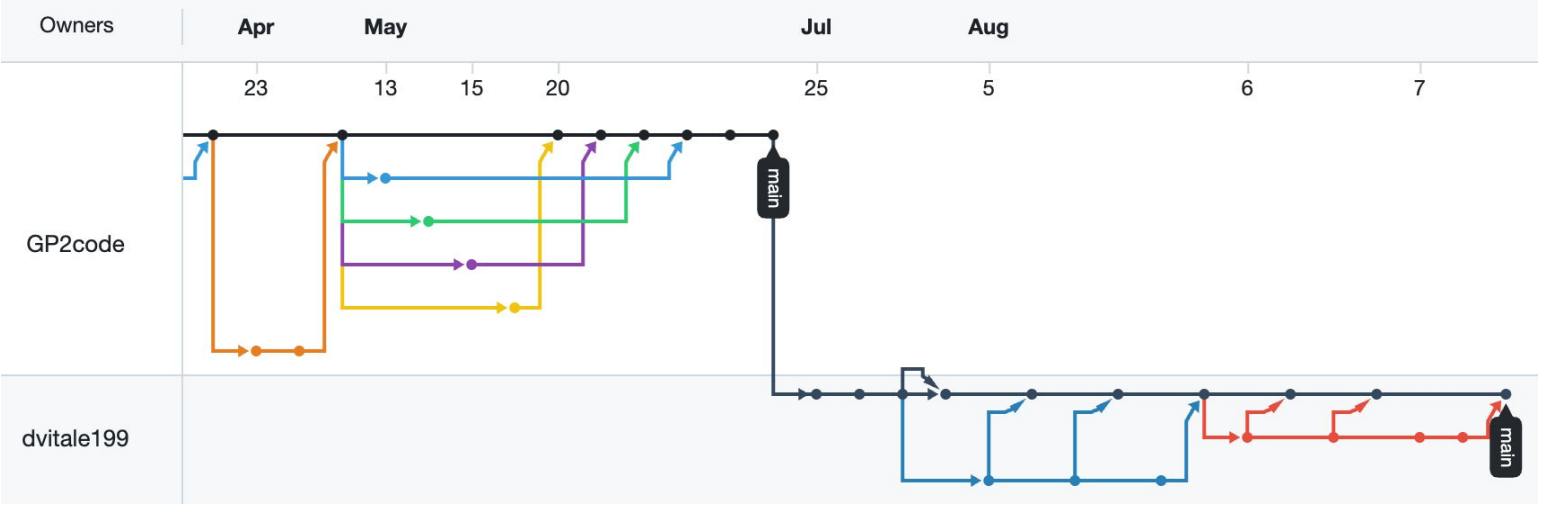
Git/GitHub: Version control tool useful for collaborative work



Step2: Coding for open science

Github: Version control tool useful for collaborative work

https://github.com/GP2code/GenoTool



Step3: Ensure traceability and reproducibility

You can just share your final github repo, with one but big caveat - DATA PRIVACY!!

2.1.2 Data privacy

While open science promotes transparency, some data must remain private:

- **Personally Identifiable Information (PII)**: Follow legal guidelines (e.g., GDPR, HIPAA).
- **Sensitive datasets**: Use controlled-access repositories when needed.
- **Anonymization**: If sharing is restricted, remove identifiable details or aggregate data.

Step3: Ensure traceability and reproducibility

You can just share your final github repo, with one but big caveat - DATA PRIVACY!!

Introduction to small-analysis-project template

It helps you:

✅ Maintain a clear, consistent folder structure that you and collaborators can navigate easily

✅ Prevent accidental sharing of sensitive raw data and secrets (e.g., API keys)

✅ Set up an analytical environment with required Python packages (venv)

Introduction to small-analysis-project template

It helps you:

✅ Maintain a clear, consistent folder structure that you and collaborators can navigate easily

✅ Prevent accidental sharing of sensitive raw data and secrets (e.g., API keys)

⬜ Set up an analytical environment with required Python packages (venv)

```
new_analysis
├── README.md
├── code                    # Code snippets and scripts
│   ├── test.py
│   ├── test.r
│   └── main.sh             # Shell script to run the main analysis
├── data                    # Input data (gitignored)
│   └── testdata.csv
├── priv                    # Private information (gitignored)
│   ├── exports.sh          # Script to export environment variables
│   └── private.txt         # Sensitive information, e.g., API keys
├── report                  # Analysis outputs (figures, summaries)
│   └── report.txt
├── temp                    # Temporary files (gitignored)
│   └── temp1.txt
├── requirements.txt
└── .gitignore
```

Introduction to small-analysis-project template

It helps you:

✅ Maintain a clear, consistent folder structure that you and collaborators can navigate easily

✅ Prevent accidental sharing of sensitive raw data and secrets (e.g., API keys)

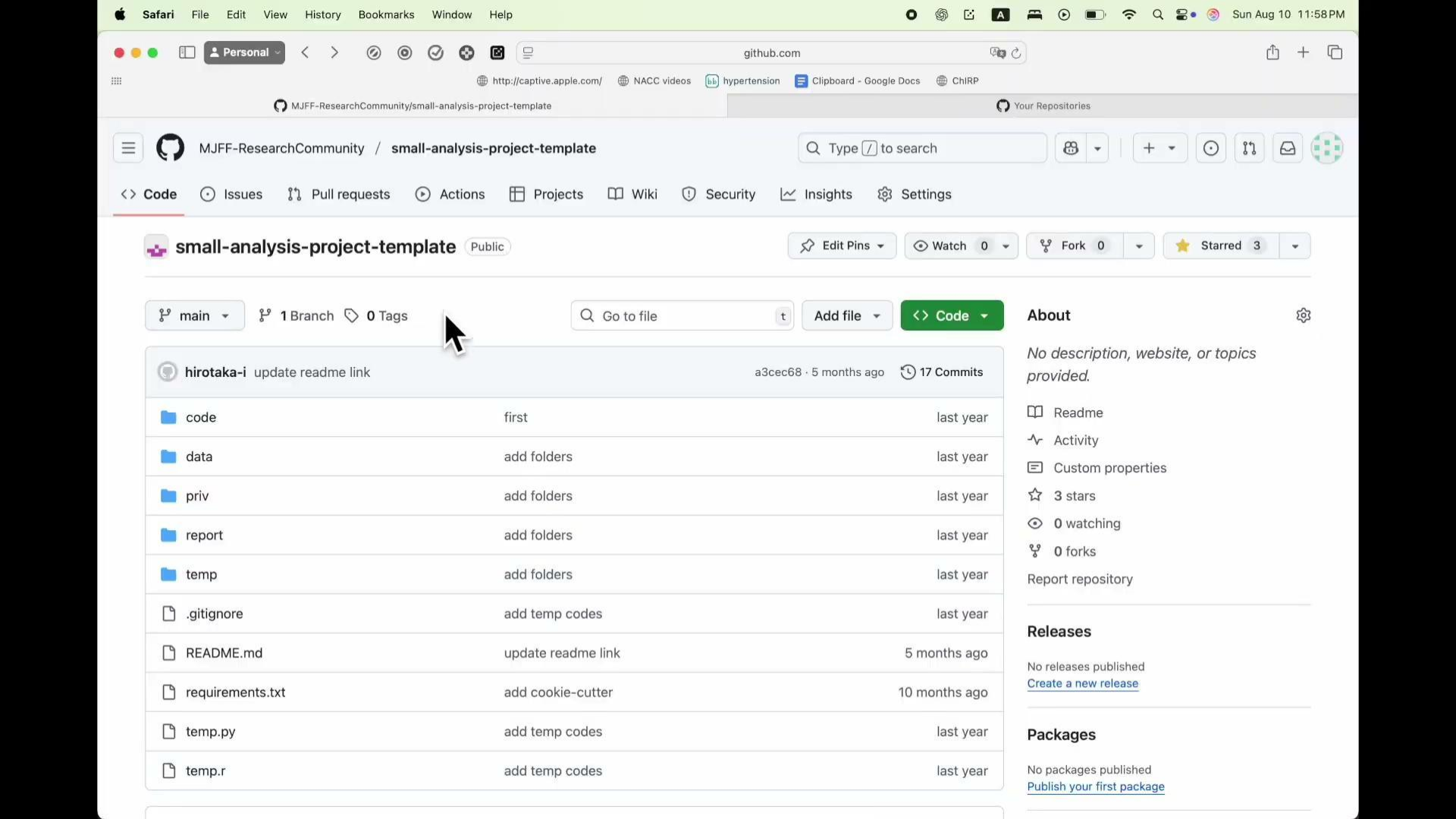
✅ Set up an analytical environment with required Python packages (venv)

```
# Create a virtual environment in the .venv director
python3 -m venv .venv # or "python -m venv .venv"

# Activate the virtual environment
source .venv/bin/activate
```

```
# Install dependencies listed in requirements.txt
python -m pip install -r requirements.txt
```

```
pip install <package_name>   # Install a new package
pip freeze > requirements.txt # Update dependencies in requirements.
```


Safari File Edit View History Bookmarks Window Help
Sun Aug 10 11:58 PM
Personal
github.com
http://captive.apple.com/
NACC videos
hypertension
Clipboard - Google Docs
ChIRP
MJFF-ResearchCommunity/small-analysis-project-template
Your Repositories
MJFF-ResearchCommunity / small-analysis-project-template
Type / to search
Code
Issues
Pull requests
Actions
Projects
Wiki
Security
Insights
Settings
small-analysis-project-template
Public
Edit Pins
Watch 0
Fork 0
Starred 3
main
1 Branch
0 Tags
Go to file
t
Add file
Code
hirotaka-i update readme link
a3cec68 · 5 months ago
17 Commits
code first last year
data add folders last year
priv add folders last year
report add folders last year
temp add folders last year
.gitignore add temp codes last year
README.md update readme link 5 months ago
requirements.txt add cookie-cutter 10 months ago
temp.py add temp codes last year
temp.r add temp codes last year
About
No description, website, or topics provided.
Readme
Activity
Custom properties
3 stars
0 watching
0 forks
Report repository
Releases
No releases published
Create a new release
Packages
No packages published
Publish your first package

Introduction to small-analysis-project template

Things to consider:

- Documentations: README should explain everything people need to know to reproduce the results.
- Jupyter notebook/lab: They are useful for data exploration but not so much for reproducing the results. Also the output could contain the PIV data. Only using it in the untracked folder would be safe.
- In collaborative development, a little bit more knowledge of Git is required.

Other considerations in the guide

2.4 Pre-registration & Study Design Transparency

Pre-registration strengthens research integrity by documenting hypotheses and methods **before** data collection. Pre-registration does **not** limit flexibility—it simply provides a record of initial research intentions.

2.5 Making Projects Citeable

We recommend establishing or creating a Digital Object Identifier (DOI) to enable researchers and the public to easily cite and access your work. A DOI is a permanent, unique identifier assigned to digital objects such as research papers, datasets, software, and code repositories. It provides a stable and citable link to the content, even if the location (URL) changes.

For example, a DOI link will look like this: https://doi.org/10.5281/zenodo.14984668 with **10.5281/zenodo.14984668** representing the DOI. It will always resolve to the same location.

Check the guide for more details

Tools for doing Open Science

Examples of Open Science in Parkinson's Research

Example 1: European Parkinson's GWAS Meta-analysis

By: GP2 and Leonard, H. L.

Open Science Principles:

- GitHub repo (transparent)
- Preprint (accessible)
- Summary statistics (shared)
- GP2 (collaborative)

Example 2: Fox Insight

By: Michael J. Fox Foundation for Parkinson's Research

Open Science Principles:

- Data use agreement (accessible)
- Fox DEN; data (shared)
- DCOP; 23andMe (collaborative)

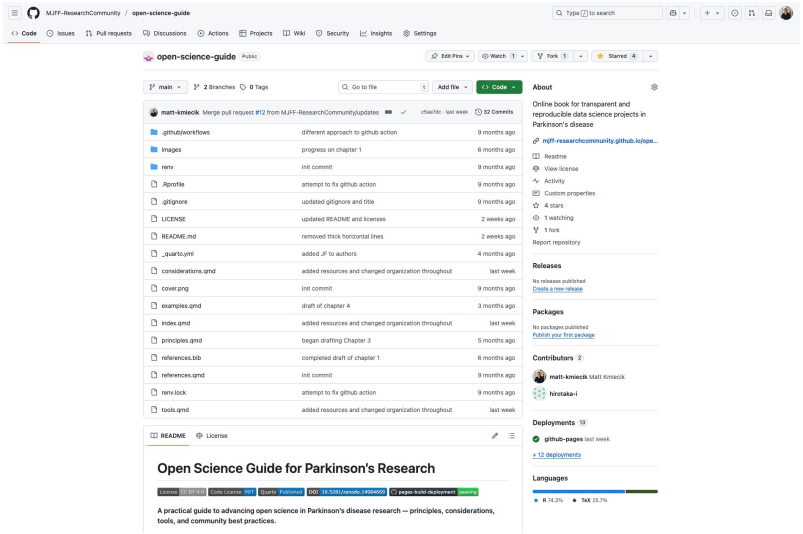
Contributing to the Open Science Guide

1. Following Open Science principles, this webbook was designed to be developed collaboratively

2. Three methods to contribute:

    a. File an issue on GitHub
    b. Make edits directly on GitHub
    c. Clone > branch > edit > commit > pull request (*advanced users*)

Deep Dive on Contributing to the OSG by Matt!



Summary

- "Open Science is transparent and accessible knowledge that is shared and developed through collaborative networks"
- There are many things to consider: data sharing and privacy, preregistration, reproducible workflows, documentation (just to name a few!)
- Many tools are available, with new ones being developed constantly!
- Parkinson's research has many great examples of Open Science projects
- Open Science Guide for Parkinson's Research is a working document; please feel free to contribute!

Discussion Topic

Into the DeLorean!

Think about your current project. If you were to start again, what would you do differently?