

Roots for Resilience in the Year of **Open Science**: Lessons Learned

Tomasz Włodarczyk

11/25/2023



THE UNIVERSITY OF ARIZONA
RESEARCH, INNOVATION & IMPACT

**Arizona Institutes
for Resilience**



RESEARCH, INNOVATION & IMPACT

Data Science Institute



Roots for Resilience & Foundational Open Science Skills



THE UNIVERSITY OF ARIZONA
RESEARCH, INNOVATION & IMPACT

Arizona Institutes
for Resilience



RESEARCH, INNOVATION & IMPACT

Data Science Institute

Foundational Open Science Skills

CyVerse's 8-week virtual workshop teaches you the principles, practices, and how-tos for doing collaborative open science using cutting-edge, open source cyberinfrastructure, in a collaborative, hands-on setting. To see how our FOSS workshop can support your work, check out the [curriculum](#).

https://foss.cyverse.org/00_basics/

*The Roots for Resilience Program provides **training and support** to select graduate students on **open, reproducible science and computational infrastructure** to enhance research focused on **resiliency in the environment**.*

Open-Science

"Open Science is the movement to make scientific research (including publications, data, physical samples, and software) and its dissemination accessible to all levels of society, amateur or professional..." (Wikipedia)

Awesome List:

Original

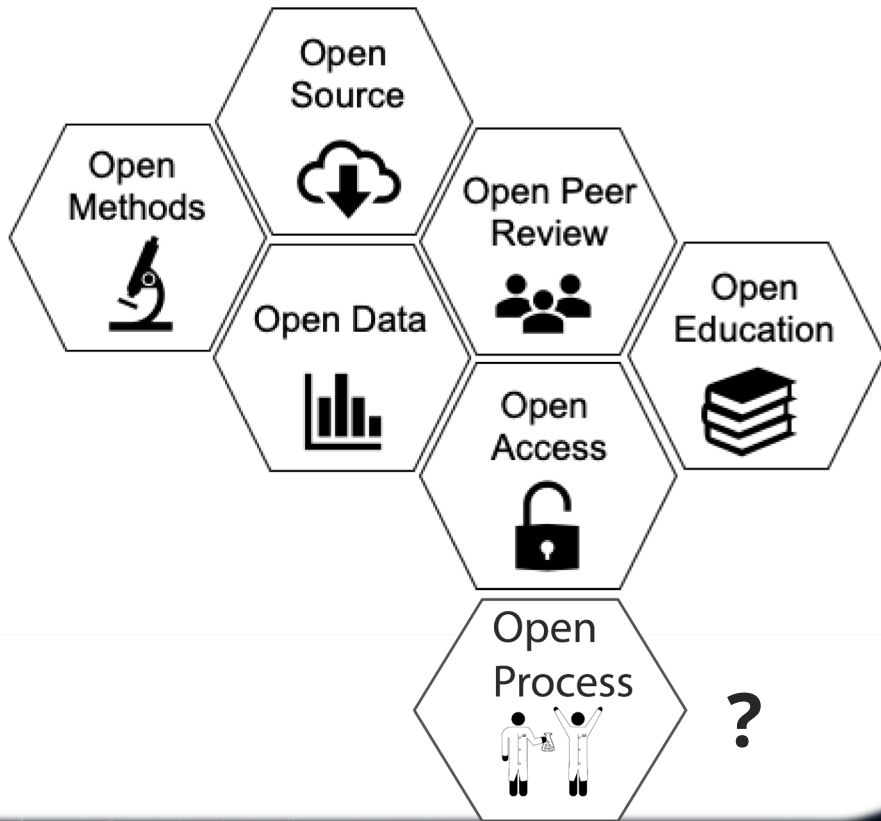
<https://github.com/sindresorhus/awesome>

Consolidated

<https://tyson-swetnam.github.io/awesome-open-science/>



Open, FAIR and CARE Science



DOI
Raw Data Archived
Open Protocols
Reproducible Codes
Clear Licensing
Open Collaboration

FAIR



Findable



Accessible



Interoperable



Reusable

CARE

Collective Benefit Authority to Control Responsibility Ethics



Open Process – In response to Reproducibility Crisis

Preregistration

- Prevents **HARKing** (Hypothesizing **A**fter the **R**esults are **K**nown)
- Separates hypothesis-generating from hypothesis-testing (because the same data can't be used to generate and test hypothesis)
- Mitigates **p-hacking** (manipulation of p-value)

<https://www.cos.io/initiatives/prereg>



Check Data Science Institute's events

RESEARCH, INNOVATION & IMPACT
Data Science Institute

Home About Research Directory Education Get Involved **Events** Analytics Power

Home > Calendar

Events

Calendar

Submit an Event

Calendar

DEC 06	CCT Sharing your research as Quarto reports 11 a.m. to 1 p.m., Dec. 6, 2023
DEC 07	Women in ML Symposium (Hosted by Google) 10:30 a.m. to 2 p.m., Dec. 7, 2023
DEC 27	End-to-end data science workflows with Posit Team 8 to 9 a.m., Dec. 27, 2023
JAN 19	CCT Bayesian Learning Group noon to 2 p.m., Jan. 19, 2024
JAN 24	CCT Demystifying APIs for Researchers 11 a.m. to 1 p.m., Jan. 24, 2024
JAN 31	End-to-end data science workflows with Posit Team

<https://datascience.arizona.edu/calendar>

CCT Data Science YouTube Channel

CCT Data Science
@cctdatascience2556 · 152 subscribers · 35 videos
This channel hosts videos produced by the University Of Arizona Data Science Team for th... >
datascience.cals.arizona.edu and 1 more link

Subscribed

Home Videos Playlists Community

Latest Popular Oldest

- Workshop Wednesdays August 2023: Making professional websites with Quarto ...
339 views · 3 months ago
- Summer Workshop Series: Exploring the Wide World of ggplot2 Extensions
Eric R. Scott · Kelsina Premier
92 views · 4 months ago
- Summer Workshop Series: Crafting Publication Quality Data Visualizations Wit...
247 views · 5 months ago
- Summer workshop series: Understanding the foundations of 'ggplot2'
310 views · 5 months ago

Learn R with Jeffrey Oliver (Every Tuesday 9-11am)

<https://jcoliver.github.io/learn-r/schedule>

Sign up for the UA Data Science Slack workspace

<https://jcoliver.github.io/uadatascience-slack/>

Contact your Data Science Ambassador

<https://datascience.arizona.edu/dsa>

Kelsey Gonzalez
@KelseyGonzalez · 994 subscribers · 12 videos
<https://kelseygonzalez.github.io/> >

Subscribed

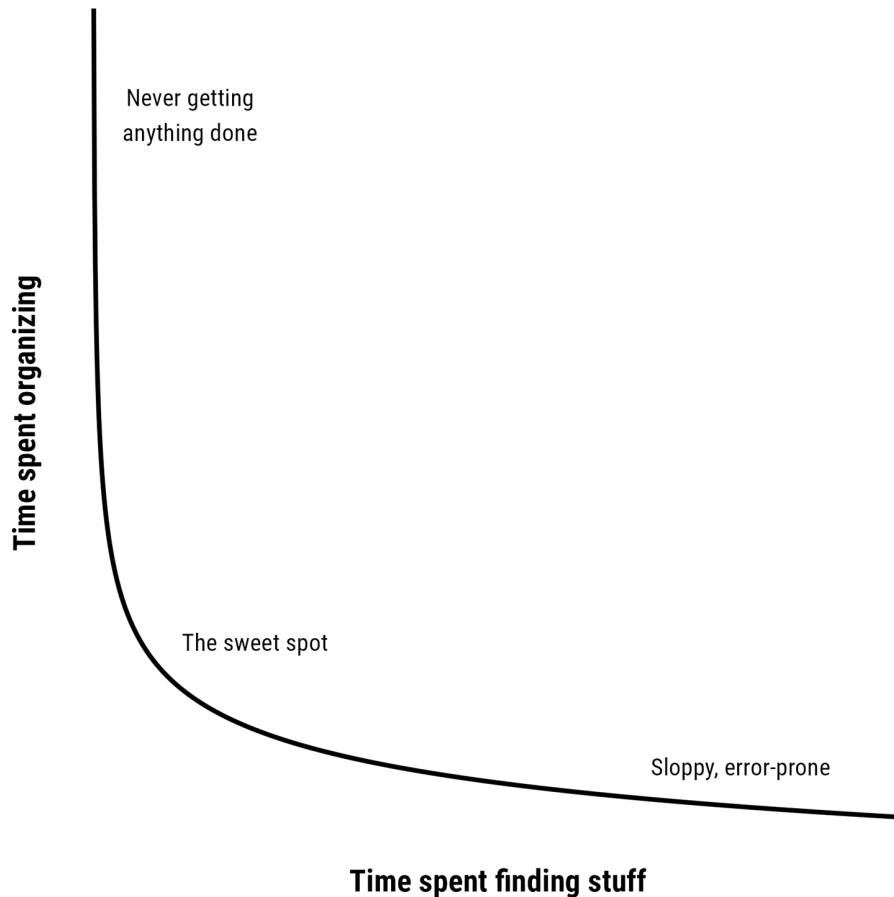
Home Videos Playlists Community

Latest Popular Oldest



Project Management

“The best time to organize is at the start, the 2nd best is right now”



1. Add README file to your project
2. Create consistent folder structure for each project
3. Keep raw data raw
4. Avoid manual (point-and-click) steps as much as possible
5. Create collaborative workspace
6. Backup your data (multiple locations)
7. Use AGILE



Project Structure

Option 1: Create a bash script to automate the process.

Structure.txt - Notepad

File Edit Format View Help

```
mkdir "$1"

cd "$1" || exit

echo "# $1" >> README.md

mkdir data

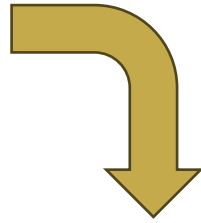
mkdir data/raw

mkdir data/cleaned

mkdir scripts

mkdir images

mkdir reports
```



Folder Structure

File Home Share View

Desktop > All documents > 1 PhD > CNRS + Synch > CyVerse > Folder Structure

Name	Status	Date modified
data	🔄	11/25/2023 2:57 PM
images	🔄	11/25/2023 2:57 PM
reports	🔄	11/25/2023 2:57 PM
scripts	🔄	11/25/2023 2:57 PM
README.md	🔄	11/25/2023 2:57 PM
Structure.sh	🔄	11/25/2023 2:55 PM

Option 2: Use CookieCutter



<https://drivendata.github.io/cookiecutter-data-science/>

```
LICENSE
Makefile      <- Makefile with commands like `make data` or `make train`
README.md     <- The top-level README for developers using this project.
data
├── external  <- Data from third party sources.
├── interim   <- Intermediate data that has been transformed.
├── processed <- The final, canonical data sets for modeling.
└── raw       <- The original, immutable data dump.

docs          <- A default Sphinx project; see sphinx-doc.org for details

models        <- Trained and serialized models, model predictions, or model summaries

notebooks     <- Jupyter notebooks. Naming convention is a number (for ordering),
                 the creator's initials, and a short '-' delimited description, e.g.
                 `1.0-jqp-initial-data-exploration`.

references    <- Data dictionaries, manuals, and all other explanatory materials.

reports
├── figures   <- Generated analysis as HTML, PDF, LaTeX, etc.
└──          <- Generated graphics and figures to be used in reporting

requirements.txt <- The requirements file for reproducing the analysis environment, e.g.
                 generated with `pip freeze > requirements.txt`

setup.py      <- Make this project pip installable with `pip install -e`
src           <- Source code for use in this project.
├── __init__.py <- Makes src a Python module
├── data        <- Scripts to download or generate data
│   └── make_dataset.py
├── features    <- Scripts to turn raw data into features for modeling
│   └── build_features.py
├── models      <- Scripts to train models and then use trained models to make
│   │           predictions
│   ├── predict_model.py
│   └── train_model.py
└── visualization <- Scripts to create exploratory and results oriented visualizations
    └── visualize.py

tox.ini       <- tox file with settings for running tox; see tox.readthedocs.io
```


Project Governance

Project Governance is the set of rules, procedures and policies that determine how projects are managed and overseen

GenoPhenoEnvo Governance and Operations Manual

Munoz-Torres, Monica C¹; Bartelme, Ryan P²; Behrisch, Michael³; Cain, Emily J²; Chang, Remco⁴; Debnath, Ishita⁵; Heidorn, Bryan²; Jaiswal, Pankaj¹; LeBauer, David S²; Mosca, Ab⁴; Ross, Arun⁵; Swetnam, Tyson L²; Thessen, Anne¹

[Show affiliations](#)

This Governance and Operations Manual (GOM) defines standard operating procedures and various policies created to clarify, support, and further the goals of the Genomics, Phenotypes, and Environment (GenoPhenoEnvo) Research Team.

Files

GenoPhenoEnvo Governance - SUBMITTED - 20200521.pdf

Page: 1 of 10 Automatic Zoom:

GenoPhenoEnvo Governance and Operations Manual

This is a living document. Changes are expected throughout the life of the project.

Introduction 1
Primary Office 2
Mission 2
Vision 2

GenoPhenoEnvo Governance and Operations Manual

This is a living document. Changes are expected throughout the life of the project.

Introduction	1
Primary Office	2
Mission	2
Vision	2
Goals	2
Participation and Roles	2
Our Team	3
Organizational Structure	5
Operations	5
Communications	5
Work Procedures	6
Scientific and Technical Decision making	7
Community Practices	7
Our Commitment to Open Science	7
Diversity Statement	7
Code of Conduct	8
Conflict resolution	9
Reporting Issues	9
Attribution, Authorship, and Ownership	9
Acknowledgements	10

Introduction

This Governance and Operations Manual (GOM) defines standard operating procedures and various policies created to clarify, support, and further the goals of the Genomics Phenotypes and Environment (GenoPhenoEnvo) Research Team.

<https://zenodo.org/records/3839120>



Work management - Kanban

<https://www.office.com/>

<https://trello.com/>



THE UNIVERSITY OF ARIZONA **Planner**

Kanban Plant-Soil-Environment Lab

Grid Board Charts Schedule ...

R MR KP MA +7 Members

To do

- Research Business Important**
- Schedule R4R workshop
- Due** WT
- Management**
- Lab Website improvements
- Due** MK
- Management**
- SharePoint workshop
- Due** MK
- Business**
- Reimbursement \$20 SME

In Progress

- Research**
- Project 1
- Test the instrument
- Generate Data
- Literature
- 0 / 3
- 12/15
- Research Important**
- Protocols
- Protocol - Plants XRF
- Protocol - Soil XRF
- Protocol - Plant survey
- 0 / 3
- 12/01

To Review

- Research**
- Paper - Methods
- Soil Statistics
- Plant Statistics
- Data Visualization
- Methodology
- 0 / 4
- Grinding/Analysis Protocol
- Test the protocol
- Finish the analysis section
- 0 / 2
- 01/31

Done (Saved)

- Completed tasks** 8

10

Microsoft Office 365

<https://www.office.com/>

Explore by category

Productivity

Utilities

Education

Communication

Content management

Project management

Developer tools

Employee Experience



Forms

Customize surveys and quizzes, get real-time results.



Sway

Create interactive reports and presentations.



Power BI

Create actionable, dynamic, and engaging data dashboards you can share with others.



Engage

Connect with coworkers and classmates, share information, and organize around projects.



Stream

Share videos of classes, meetings, presentations, and training sessions.



Visio

Simplify and communicate complex information visually.



Project

Develop project plans, assign tasks, track progress, and manage budgets.



Whiteboard

Ideate and collaborate on a freeform canvas designed for pen, touch and keyboard.



Lists

Allows users to create, share, and track data inside lists.



SharePoint

Share and manage content, knowledge, and applications to empower teamwork.



Insights

Improve your productivity and wellbeing with Microsoft Viva Insights.



Power Automate

Sync files and more to simplify your work.



Planner

Create plans, organize and assign tasks, share files, and get progress updates.



Loop

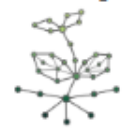
Enabling teams to think, plan, and create together.



Power Pages

Unleash the power of Copilot to craft secure, low-code business websites with ease.





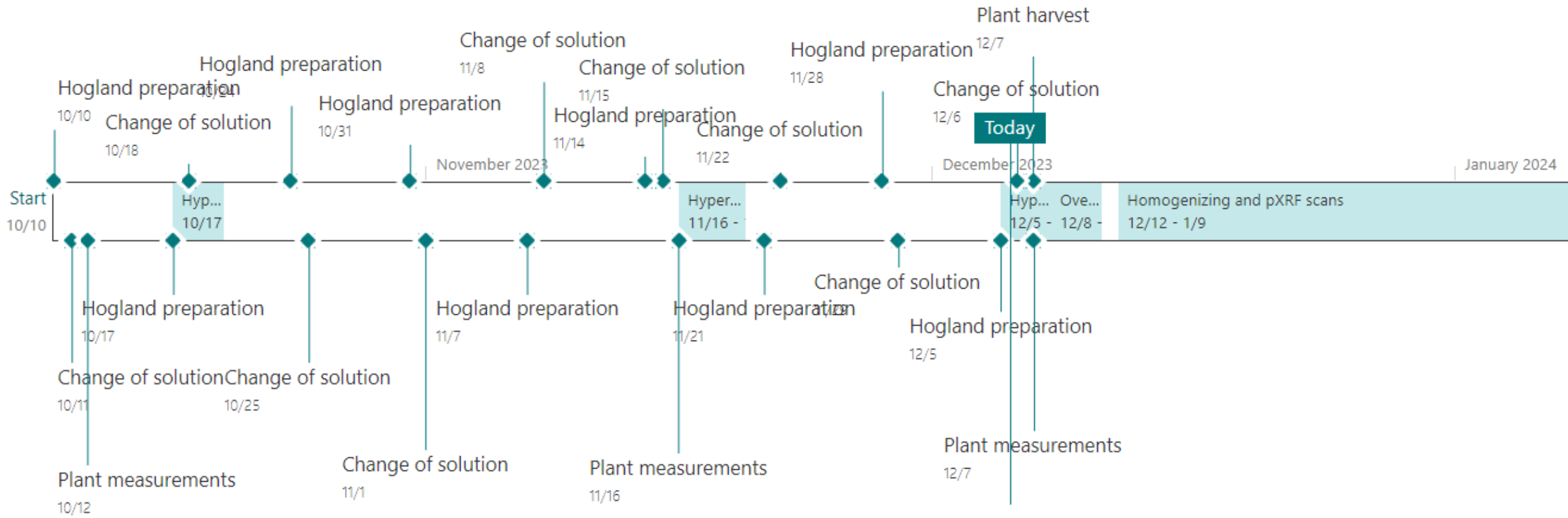
Tomek Research

- Home
- Notebook
- Documents
- Pages
- Obligations
- Workload Calendar
- Time Management
- Fieldwork**
- Halleri_Hydroponics
- Site contents
- Recycle bin
- Edit

[+ New](#)
[Upload](#)
[Edit in grid view](#)
[Sync](#)
[Add shortcut to OneDrive](#)
[Pin to Quick access](#)

Fieldwork

Name	Modified	Modified By	+ Add column
BHP	September 27, 2022	Wlodarczyk, Tomasz Tadeu	
Data	April 15, 2022	Wlodarczyk, Tomasz Tadeu	
Documents	May 31, 2022	Wlodarczyk, Tomasz Tadeu	
Photos	April 15, 2022	Wlodarczyk, Tomasz Tadeu	
Plants	June 1, 2022	Wlodarczyk, Tomasz Tadeu	
Presentations	May 31, 2022	Wlodarczyk, Tomasz Tadeu	
Measurement Table.xlsx	April 20	Rios Dominguez, Shaila Ja	
Schedule_TW.JPG	August 20, 2022	Wlodarczyk, Tomasz Tadeu	
Solitude_Measurement Table_and_Control.x...	July 28	Wlodarczyk, Tomasz Tadeu	
Week Time Table .xlsx	August 26, 2022	Wlodarczyk, Tomasz Tadeu	



new task or edit this list

All Tasks Calendar Completed

✓		Task Name	Due Date	Assigned To
✓		Hogland preparation	... October 3	<input type="checkbox"/> Galloway, Jacob Ryan - (jrgalloway5) <input type="checkbox"/> Włodarczyk, Tomasz Tadeusz - (twłodarczyk) <input type="checkbox"/> Babst-Kostecka, Alicja - (ababstkostecka) <input type="checkbox"/> Stokes, Owyn Robert - (ors)

Version history

Delete All Versions

No. ↓	Modified	Modified By	Size	Comments
224.0	4/20/2023 2:53 PM	<input type="checkbox"/> Rios Dominguez, Shaila Janeth - (sriosdominguez)	37.4 KB	
223.0	3/23/2023 3:04 PM	<input type="checkbox"/> Wlodarczyk, Tomasz Tadeusz - (twlodarczyk)	37 KB	
222.0	3/23/2023 2:51 PM	<input type="checkbox"/> Wlodarczyk, Tomasz Tadeusz - (twlodarczyk)	37 KB	
221.0	3/23/2023 2:46 PM	<input type="checkbox"/> Wlodarczyk, Tomasz Tadeusz - (twlodarczyk)	37.1 KB	
220.0	3/23/2023 2:33 PM	<input type="checkbox"/> Wlodarczyk, Tomasz Tadeusz - (twlodarczyk)	36.6 KB	
219.0	3/23/2023 2:22 PM	<input type="checkbox"/> Wlodarczyk, Tomasz Tadeusz - (twlodarczyk)	36.4 KB	
218.0	3/23/2023 2:09 PM	<input type="checkbox"/> Wlodarczyk, Tomasz Tadeusz - (twlodarczyk)	36.3 KB	
217.0	3/23/2023 1:48 PM	<input type="checkbox"/> Wlodarczyk, Tomasz Tadeusz - (twlodarczyk)	36 KB	
216.0	3/19/2023 10:53 AM	<input type="checkbox"/> Wlodarczyk, Tomasz Tadeusz - (twlodarczyk)	35.3 KB	
215.0	3/17/2023 1:29 PM	<input type="checkbox"/> Wlodarczyk, Tomasz Tadeusz - (twlodarczyk)	35.3 KB	
214.0	3/17/2023 1:26 PM	<input type="checkbox"/> Wlodarczyk, Tomasz Tadeusz - (twlodarczyk)	35.3 KB	
213.0	3/17/2023 1:15 PM	<input type="checkbox"/> Wlodarczyk, Tomasz Tadeusz - (twlodarczyk)	32.2 KB	
212.0	3/17/2023 12:45 PM	<input type="checkbox"/> Wlodarczyk, Tomasz Tadeusz - (twlodarczyk)	32.2 KB	



Version Control – Git and GitHub



1. It is a software	1. It is a service
2. It is installed locally on the system	2. It is hosted on Web
3. It is a command line tool	3. It provides a graphical interface
4. It is a tool to manage different versions of edits, made to files in a git repository	4. It is a space to upload a copy of the Git repository
5. It provides functionalities like Version Control System Source Code Management	5. It provides functionalities of Git like VCS, Source Code Management as well as adding few of its own features

Version Control – Git

Commit, Push, Pull, Clone, Fork

```
1117 Cu <- ggplot(dt_selected, aes(x = reorder(Scientific_Name, Predicted_Cu_ICP, FUN = median),
1118                                y = Predicted_Cu_ICP, Sceintific_Name = Scientific_Name)) +
1119   geom_boxplot(linewidth=0.3) +
1120   geom_point(aes(shape = Plot), size = 2.5) + # Adjust the size parameter here
1121   scale_shape_manual(values = c(21, 21, 21, 4)) +
1122   geom_hline(yintercept = 70, linetype = "dashed", color = "#9a9a9a", size = 0.4) +
1123   geom_hline(yintercept = 300, linetype = "dotdash", color = "#454545", size = 0.4) +
1124   scale_y_continuous(limits = c(0, 900), breaks = seq(0, 900, by = 150)) +
1125   coord_flip() +
1126   theme_classic()+
1127   theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
1128         axis.text.x = element_text(size=12),
1129         axis.title.x = element_text(size = 15),
1130         axis.text.y = element_text(size=12, face="italic"),
1131         axis.title.y = element_blank(),
1132         legend.key.size = unit(1, "lines"),
1133         legend.text = element_text(size = 13),
1134         legend.title = element_text(size=14, face = "bold"))+
1135   guides(color = guide_legend(override.aes = list(size = 3.5)),
1136         shape = guide_legend(override.aes = list(size = 3.5))) +
1137   ylab("Cu (mg kg-1)")
1138 Cu
```





Git

Create Git repository: *git init*
Add files to track versions: *git add .*
Checking the status of the file: *git status*
Committing after modifying: *git -m "..."*
Pushing changes to GitHub: *git push*

https://www.youtube.com/watch?v=tRZGeaHPoaw&ab_channel=KevinStratvert

```
MINGW64:/c/Users/twlodarczyk
twlodarczyk@CAL5-ENVS21-03 MINGW64 ~
$ git
usage: git [-v | --version] [-h | --help] [-C <path>] [-c <name>=<value>]
          [--exec-path[=<path>]] [--html-path] [--man-path] [--info-path]
          [-p | --paginate | -P | --no-pager] [--no-replace-objects] [--bare]
          [--git-dir=<path>] [--work-tree=<path>] [--namespace=<name>]
          [--super-prefix=<path>] [--config-env=<name>=<envvar>]
          <command> [<args>]

These are common Git commands used in various situations:

start a working area (see also: git help tutorial)
  clone      Clone a repository into a new directory
  init       Create an empty Git repository or reinitialize an existing one

work on the current change (see also: git help everyday)
  add        Add file contents to the index
  mv         Move or rename a file, a directory, or a symlink
  restore    Restore working tree files
  rm         Remove files from the working tree and from the index

examine the history and state (see also: git help revisions)
  bisect    Use binary search to find the commit that introduced a bug
  diff      Show changes between commits, commit and working tree, etc
  grep      Print lines matching a pattern
  log       Show commit logs
  show      Show various types of objects
  status    Show the working tree status

grow, mark and tweak your common history
  branch    List, create, or delete branches
  commit    Record changes to the repository
  merge     Join two or more development histories together
  rebase    Reapply commits on top of another base tip
  reset     Reset current HEAD to the specified state
  switch    Switch branches
  tag       Create, list, delete or verify a tag object signed with GPG

collaborate (see also: git help workflows)
  fetch     Download objects and refs from another repository
  pull      Fetch from and integrate with another repository or a local branch
  push      Update remote refs along with associated objects

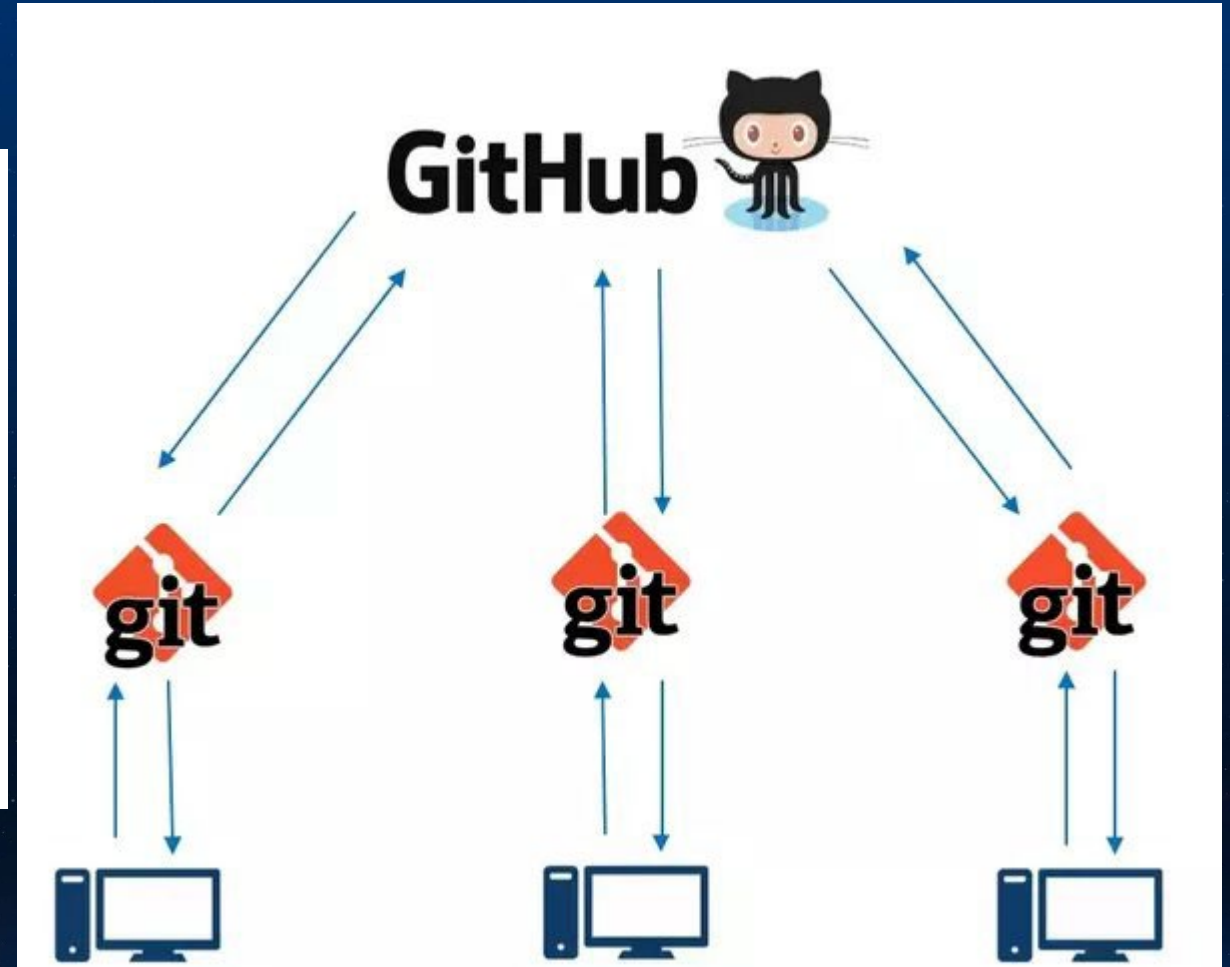
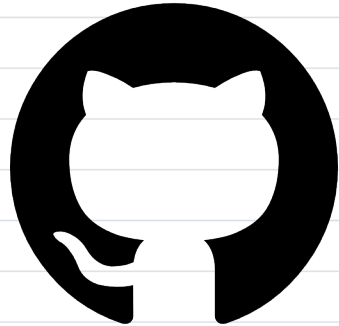
'git help -a' and 'git help -g' list available subcommands and some
concept guides. See 'git help <command>' or 'git help <concept>'
to read about a specific subcommand or concept.
See 'git help git' for an overview of the system.

twlodarczyk@CAL5-ENVS21-03 MINGW64 ~
$ |
```

Version Control – Git and Github

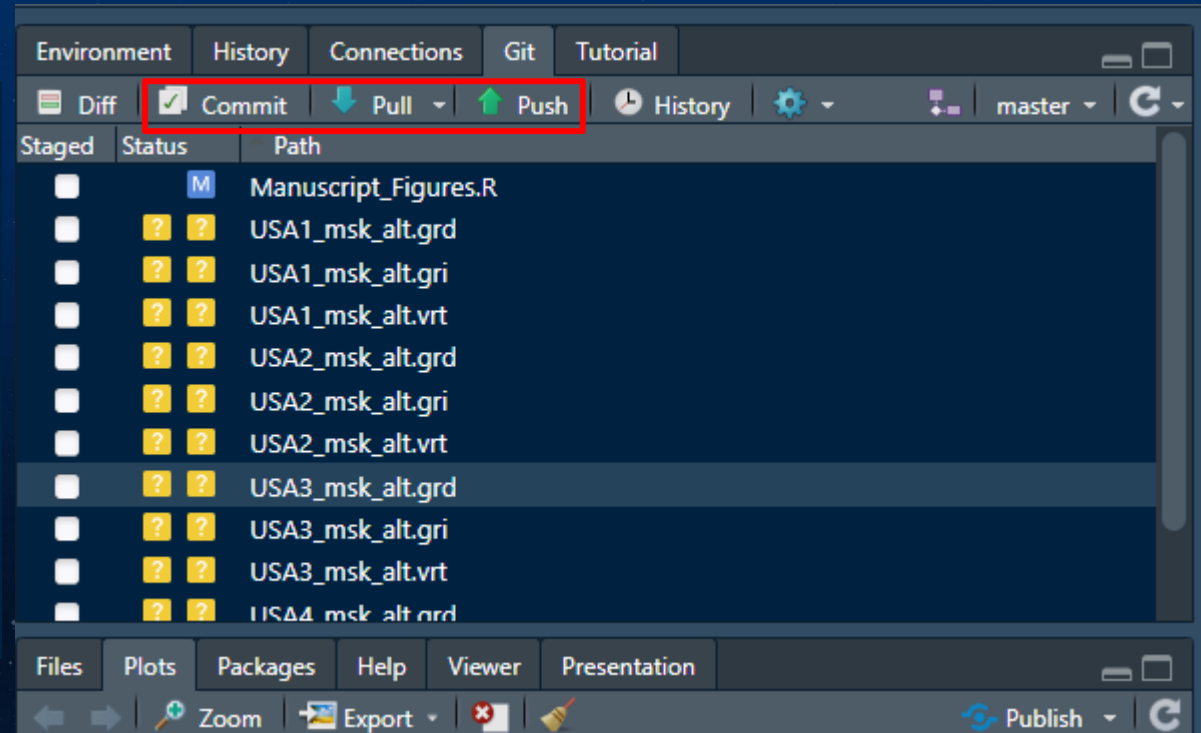
Commit, Push, Pull, Clone, Fork

Tomasz Add new figures		6db05c6 5 days ago	🕒 97 commits
images	new		8 months ago
.gitignore	Dataframe transformation and visualization of all species for individ...		9 months ago
CESMlogo.jpg	new		8 months ago
CESMlogo.png	Add files via upload		7 months ago
Dataset_LODs.R	Add new files		6 months ago
Figures_CESM.R	qwe		3 months ago
Hyperspectral_Test.R	Add new		3 months ago
LICENSE	Initial commit		9 months ago
Manuscript_Analyses.R	New analysis		2 weeks ago
Manuscript_Figures.R	Add kruskal		3 weeks ago
Manuscript_Mastefile.R	Add new loops		last week



Version Control – R Studio

```
1117 Cu <- ggplot(dt_selected, aes(x = reorder(Scientific_Name, Predicted_Cu_ICP, FUN = median),
1118                               y = Predicted_Cu_ICP, Sceintific_Name = Scientific_Name)) +
1119   geom_boxplot(linewidth=0.3) +
1120   geom_point(aes(shape = Plot), size = 2.5) + # Adjust the size parameter here
1121   scale_shape_manual(values = c(21, 21, 21, 4)) +
1122   geom_hline(yintercept = 70, linetype = "dashed", color = "#9a9a9a", size = 0.4) +
1123   geom_hline(yintercept = 300, linetype = "dotted", color = "#454545", size = 0.4) +
1124   scale_y_continuous(limits = c(0, 900), breaks = seq(0, 900, by = 150)) +
1125   coord_flip() +
1126   theme_classic()+
1127   theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
1128         axis.text.x = element_text(size=12),
1129         axis.title.x = element_text(size = 15),
1130         axis.text.y = element_text(size=12, face="italic"),
1131         axis.title.y = element_blank(),
1132         legend.key.size = unit(1, "lines"),
1133         legend.text = element_text(size = 13),
1134         legend.title = element_text(size=14, face = "bold"))+
1135   guides(color = guide_legend(override.aes = list(size = 3.5)),
1136         shape = guide_legend(override.aes = list(size = 3.5))) +
1137   ylab("Cu (mg kg-1)")
1138 Cu
```



The screenshot shows the R Studio Git interface. The top menu bar includes Environment, History, Connections, Git, and Tutorial. The Git menu is open, showing options for Diff, Commit (highlighted with a red box), Pull, and Push. Below the menu, a table lists staged files with their status and paths.

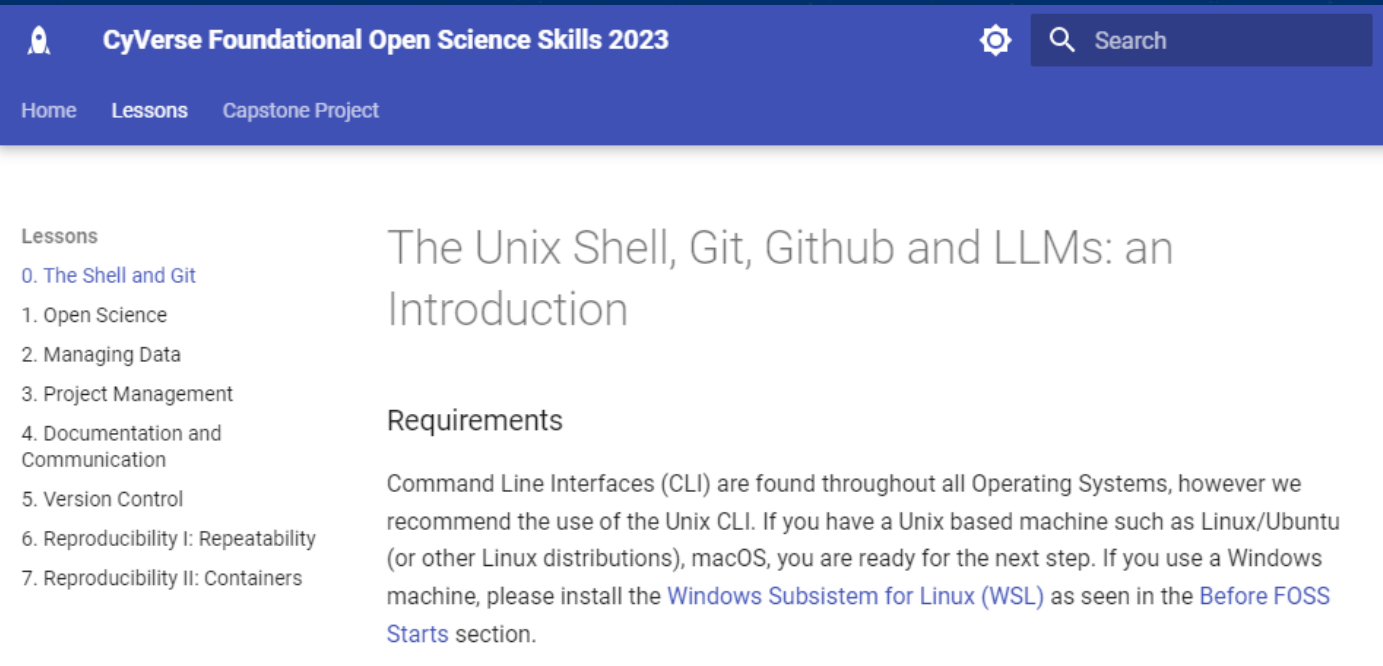
Staged	Status	Path
<input type="checkbox"/>	M	Manuscript_Figures.R
<input type="checkbox"/>	??	USA1_msk_alt.grd
<input type="checkbox"/>	??	USA1_msk_alt.gri
<input type="checkbox"/>	??	USA1_msk_alt.vrt
<input type="checkbox"/>	??	USA2_msk_alt.grd
<input type="checkbox"/>	??	USA2_msk_alt.gri
<input type="checkbox"/>	??	USA2_msk_alt.vrt
<input type="checkbox"/>	??	USA3_msk_alt.grd
<input type="checkbox"/>	??	USA3_msk_alt.gri
<input type="checkbox"/>	??	USA3_msk_alt.vrt
<input type="checkbox"/>	??	USA4_msk_alt.grd

The bottom of the interface shows a toolbar with icons for Files, Plots, Packages, Help, Viewer, and Presentation. The status bar at the bottom includes Zoom, Export, and Publish buttons.

<https://jcoliver.github.io/learn-r//010-github.html>



GitHub Pages



CyVerse Foundational Open Science Skills 2023

Home Lessons Capstone Project

Lessons

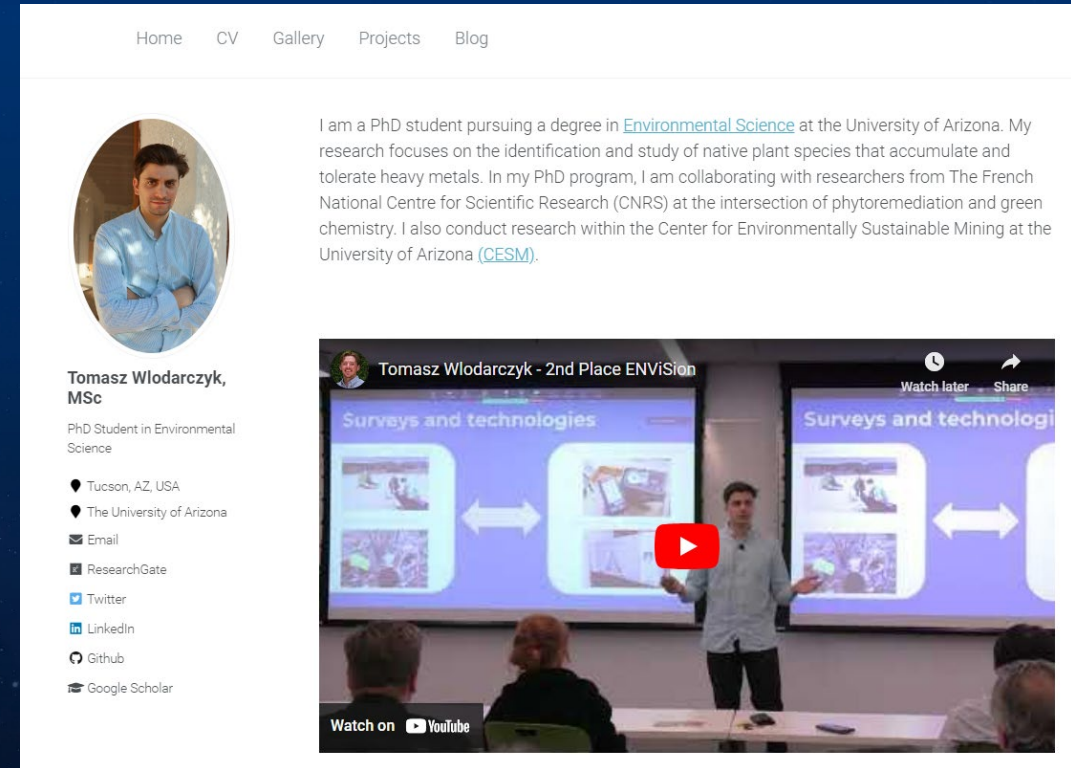
0. The Shell and Git
1. Open Science
2. Managing Data
3. Project Management
4. Documentation and Communication
5. Version Control
6. Reproducibility I: Repeatability
7. Reproducibility II: Containers

The Unix Shell, Git, Github and LLMs: an Introduction

Requirements

Command Line Interfaces (CLI) are found throughout all Operating Systems, however we recommend the use of the Unix CLI. If you have a Unix based machine such as Linux/Ubuntu (or other Linux distributions), macOS, you are ready for the next step. If you use a Windows machine, please install the [Windows Subsystem for Linux \(WSL\)](#) as seen in the [Before FOSS Starts](#) section.

<https://squidfunk.github.io/mkdocs-material/>



Home CV Gallery Projects Blog

Tomasz Włodarczyk, MSc

PhD Student in Environmental Science

- Tucson, AZ, USA
- The University of Arizona
- Email
- ResearchGate
- Twitter
- LinkedIn
- GitHub
- Google Scholar

I am a PhD student pursuing a degree in [Environmental Science](#) at the University of Arizona. My research focuses on the identification and study of native plant species that accumulate and tolerate heavy metals. In my PhD program, I am collaborating with researchers from The French National Centre for Scientific Research (CNRS) at the intersection of phytoremediation and green chemistry. I also conduct research within the Center for Environmentally Sustainable Mining at the University of Arizona ([CESM](#)).

Tomasz Włodarczyk - 2nd Place ENViSion

Surveys and technologies

Watch on YouTube

<https://jekyllcodex.org/>

<https://themes.jekyllrc.org/>

<https://github.com/TRwłodarczyk/TW-Science>



Open Data – Where to store?

Private



Public



CyVerse Platform



Analyze & Share

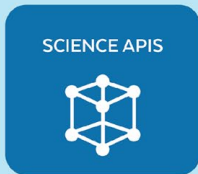


Build Your Skills



Bring Your Own

PRODUCTS



CyVerse platform <http://www.cyverse.org> provides data management tools such as:

- Data storage (upload 100GB, request 10 TB)
- Cloud-based analysis using VICE (Visual and Interactive Computing Environment)
- Enhances FAIR data principles (DOI request and data publishing directly on CyVerse Data Commons).



AI tools

<https://github.com/ua-data7/LearningResources/wiki/AI-Tools-Landscape>



COGNOSYS

AI Knowledge Discovery / Research Assistants

- [chatPDF](#)
- [Consensus](#)
- [Cradle](#). (Proteins design).
- [Elicit](#)
- [Genei](#)
- [Research Rabbit](#)
- [SciSummary](#)
- [SciSpace - typeset.io](#)
- [scite](#)
- [Semantic Scholar](#)
- [TextLayer](#)
- [Wizdom.ai](#)

General Writing Assistants

- [Jasper](#)
- [Lex](#)
- [Notion.so](#)
- [Rytr](#)

AI Generated Presentation Templates

- [SlidesPilot](#)

Code assistants

- [AlphaCode](#). DeepMind.
- Included in [Bard](#). Google.
- Included in [GPT 3.5 | GPT 4.0](#). OpenAI.
- Included in [Claude](#). Anthropic.
- [Code Whisperer](#). Amazon.
- [CoPilot X](#). Github.
- [Ghostwriter](#). Replit.
- [Tabnine](#)

Popular

General AI Chat bots

- [Bard](#). Google.
- [Bing Chat](#). (Use [Microsoft Edge](#)).
- [GPT 3.5 | GPT 4.0](#). OpenAI.
- [Claude](#). Anthropic.
- [Perplexity.ai](#). (Try Perplexity Labs to play with different LLM)
- [Poe](#). (Offers a wide collection of bots).

other options

- [glean](#)
- [mem](#)
- [OpenIndex.ai](#)
- [Rewind](#)

Chat bots as Chrome Extensions

- [WebChatGPT](#). ChatGPT/Claude/Bard/Bing.
- [Google Search Labs - Generative](#)

Text-to-image

- [DALL-E 3.0](#). OpenAI.
- [Deep Dream Generator](#)
- [Firefly](#). Adobe
- [Image Creator](#). (Use [Microsoft Edge](#))
- [Leonardo](#)
- [Midjourney](#)(via Discord channel)
- [StableDiffusion](#)

Text-to-video

- [Runway](#)
- [Fliki](#)
- [Wonder](#)

Reproducibility - containers



- Docker is the virtualization software
- Docker does packaging the application with all dependencies, configuration, system tools and runtime into a container
- You no longer have to install exact versions and dependencies to reproduce a code
- Run a container with its own isolated environment
- Focus on the development, instead of on installing and configuring services on your machine

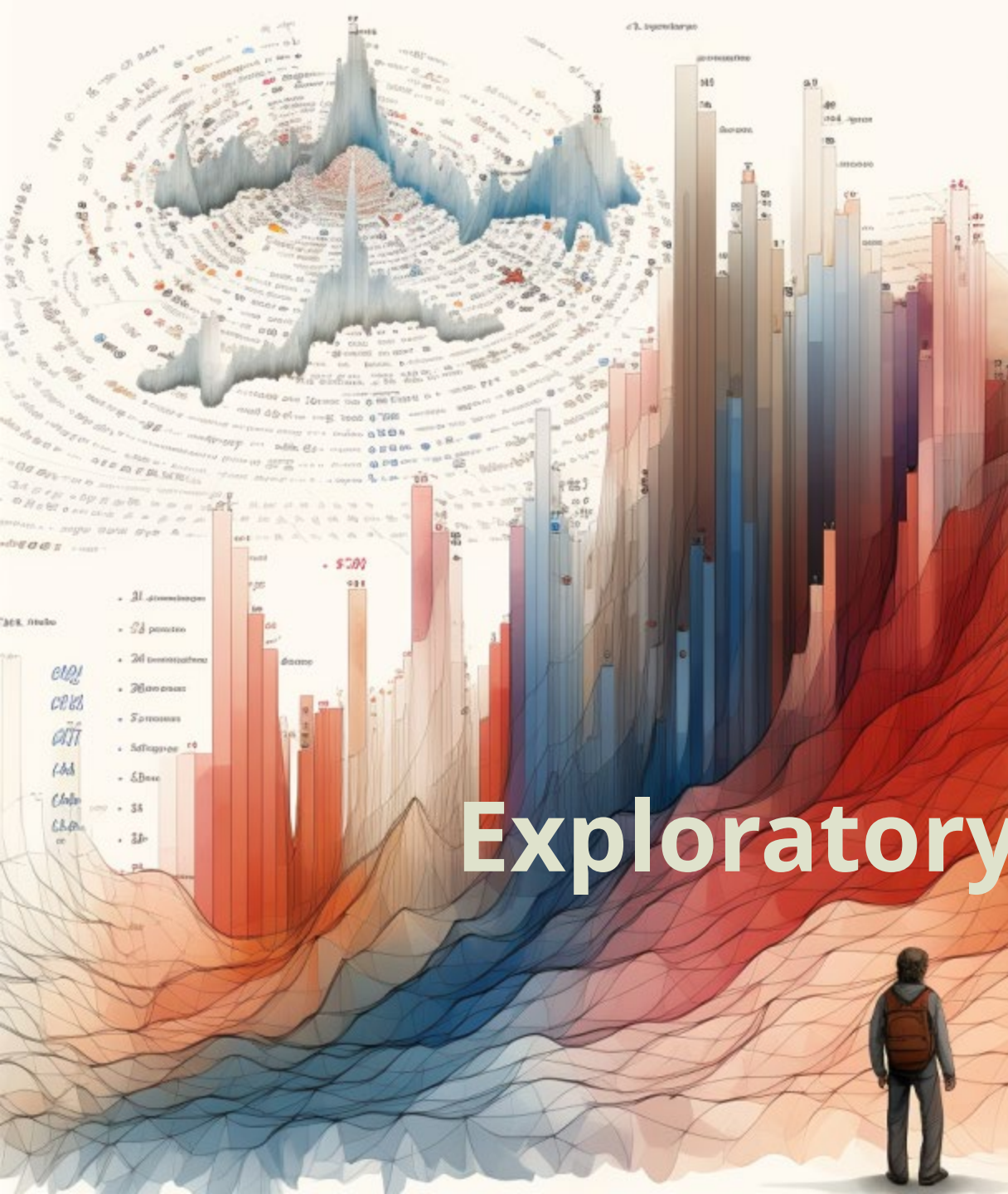
```
bash
```

```
docker pull username/my-custom-environment:tag
```

```
bash
```

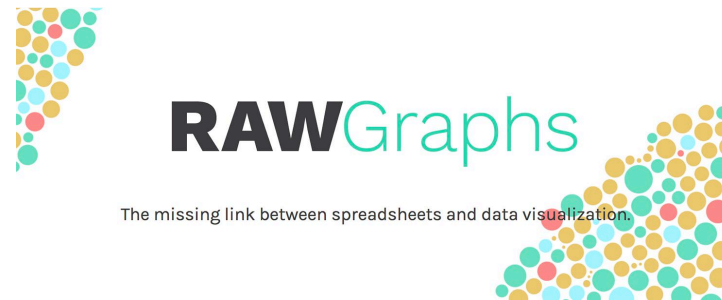
```
docker run -it username/my-custom-environment:tag
```





Exploratory Data Analysis

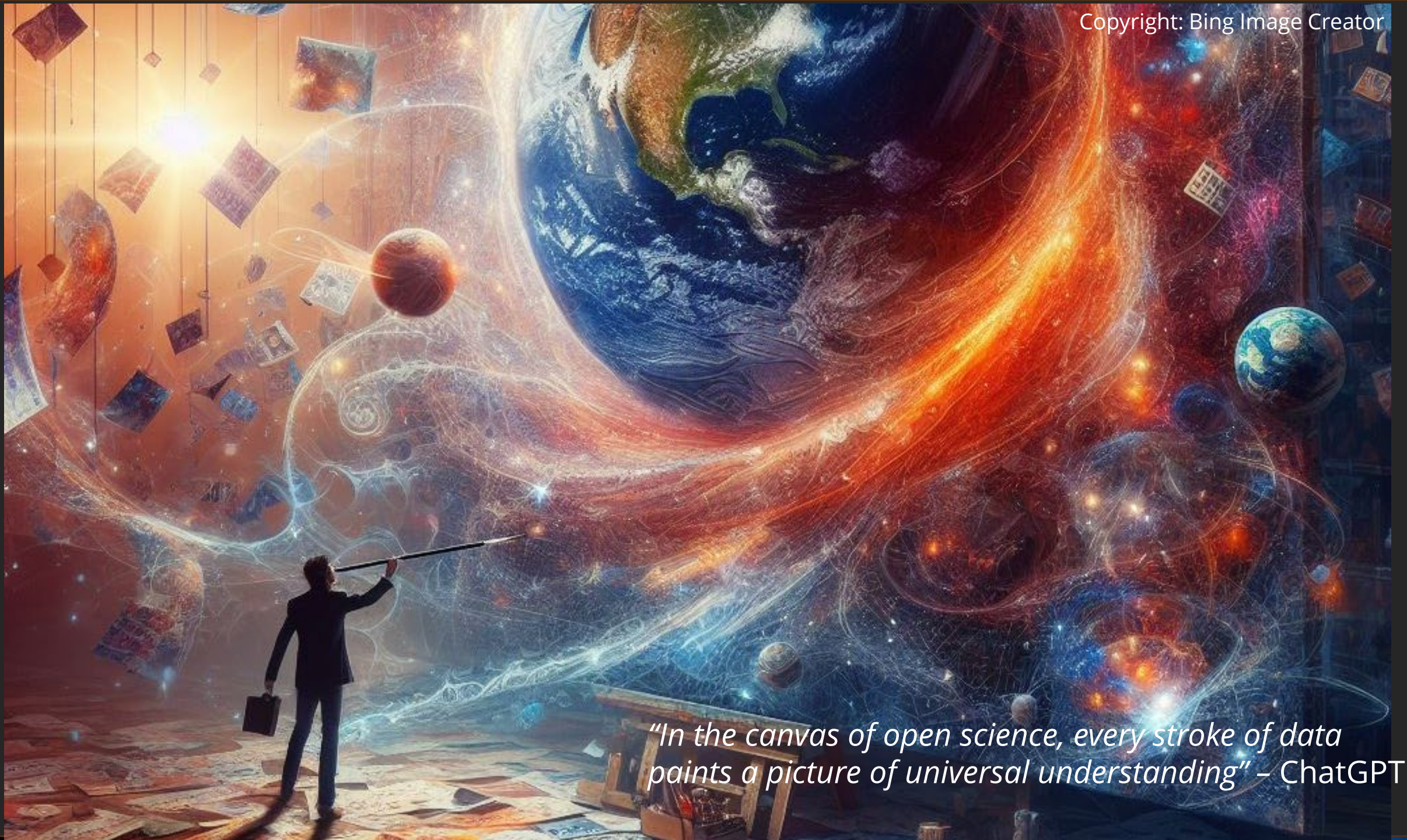
Explore, refine and visualize data (no coding)



OpenRefine



data**voyager**



"In the canvas of open science, every stroke of data paints a picture of universal understanding" – ChatGPT

Access Presentation Links:
<https://trwlodarczyk.github.io/OpenScience/>

*"In the canvas of open science, every stroke of data
paints a picture of universal understanding" – ChatGPT*