



# FWE 458

## Environmental Data Science

Lecture 2 git and github

Spring 2024  
Instructor: Min Chen

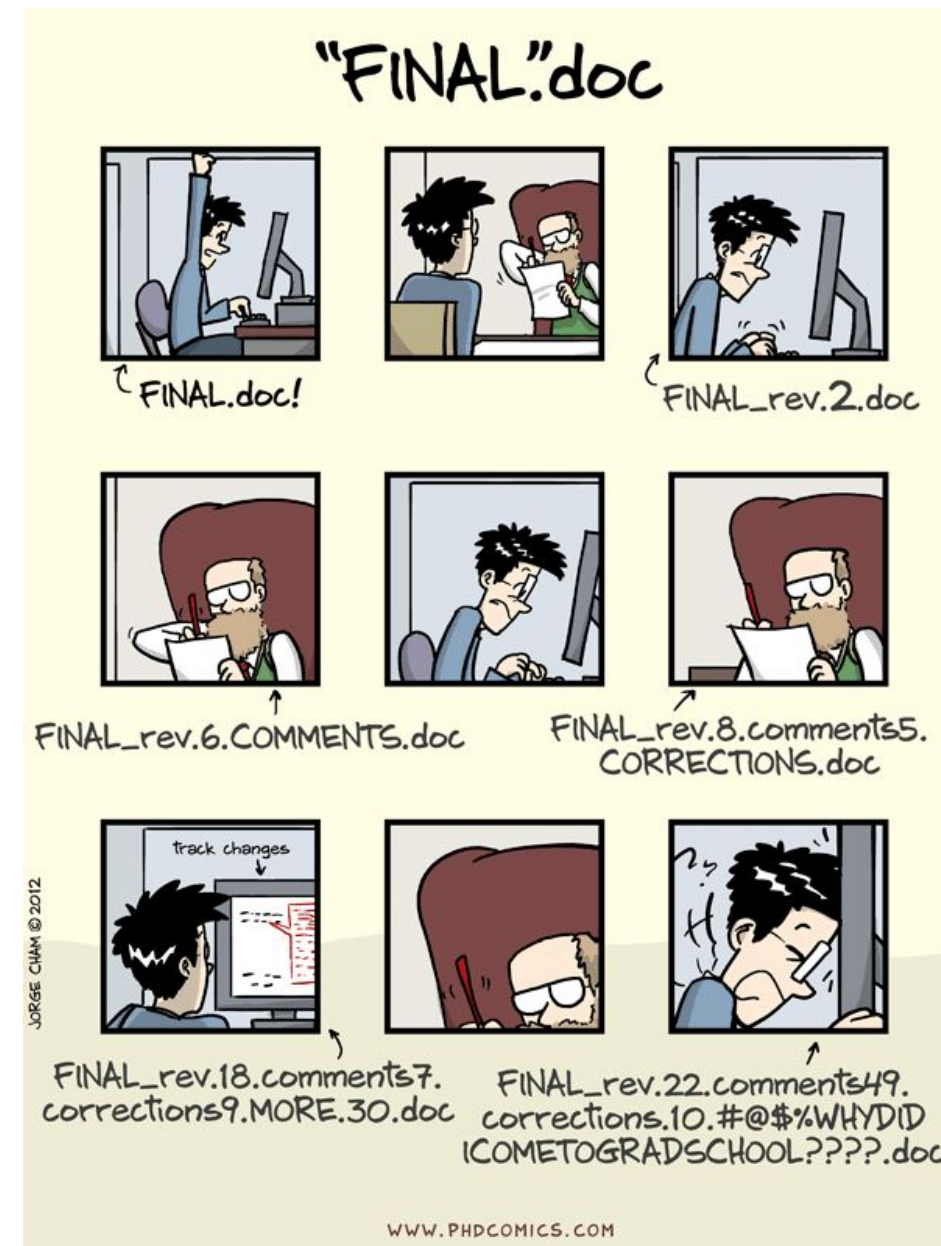
Jan 25, 2024



# Version control

Have you ever met these issues:

- overwriting each other's work
- losing track of the latest version
- not being able to revert to a previous state if something goes wrong





# Version Control

---

- a system that records changes to a file or set of files over time so that you can recall specific versions later
- "repository" (a storage location for your project), "commit" (saving changes), and "revision" (each change or set of changes).
- think about 'Google Doc'





# Benefit of Version Control

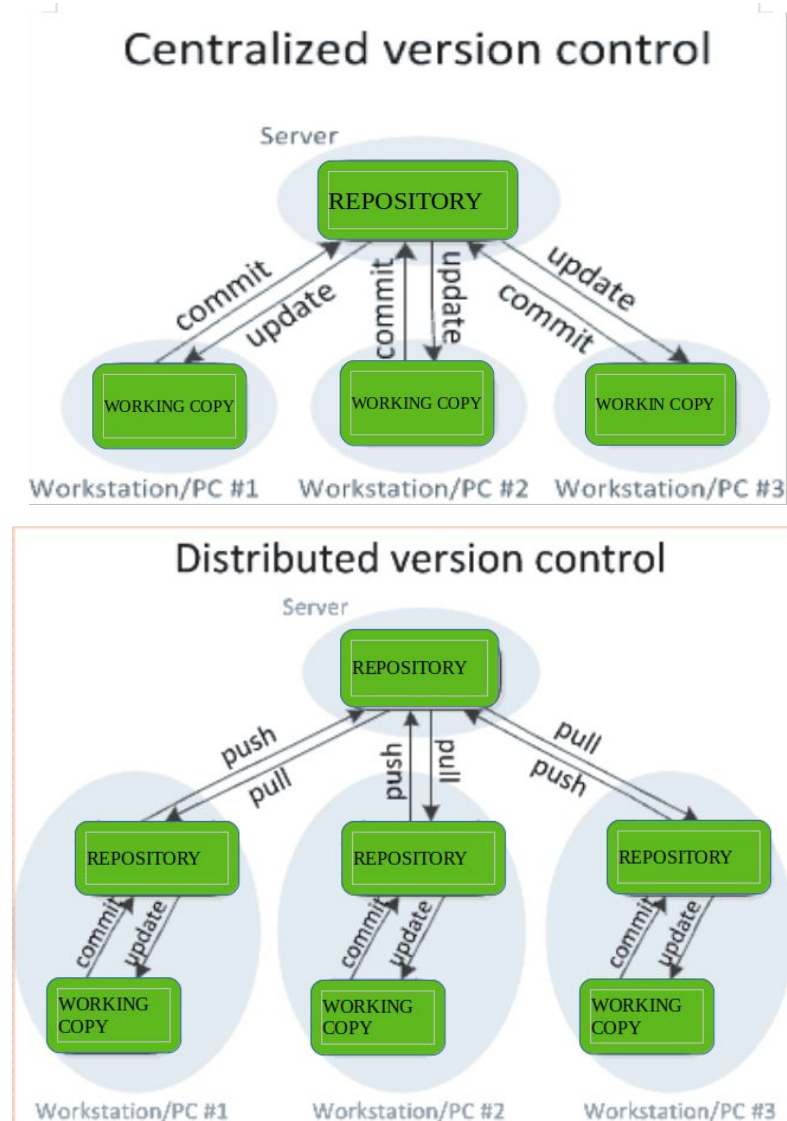
---

- **Collaboration:** Allows multiple people to work on the same project without interfering with each other's changes.
- **Tracking Changes:** Keeps a history of who changed what and when, which is crucial for understanding the evolution of a project and for debugging issues.
- **Reverting:** The ability to revert to a previous state if a mistake is made or if a project direction changes.
- **Experimentation:** Facilitates trying new ideas without the fear of losing the original work, as changes can be made in separate branches.



# Types of Version Control Systems

- Local Version Control Systems: Simple database that keeps all the changes to files under revision control.
- Centralized Version Control Systems (CVCS): Systems like SVN, where a single server contains all the versioned files, and various clients check out files from this central place.
- Distributed Version Control Systems (DVCS): clients don't just check out the latest snapshot of the files; they fully mirror the repository including its full history.





# Git

---

A popular distributed version control system

Git was originally authored by Linus Torvalds in 2005 for development of the Linux kernel, with other kernel developers contributing to its initial development.



<https://git-scm.com/>



# GitHub

---

- A web-based platform used for version control and collaboration
- The GitHub service started in February 2008. The company, GitHub, Inc., has existed as of 2007 and is located in San Francisco. On February 24, 2009, GitHub announced that within the first year of being online, GitHub had accumulated over 46,000 public repositories. The number became 372 million in 2023.





# Git vs GitHub

- **Git is the tool, GitHub is the service:**

- Git: Explain that Git is a distributed version control system, a tool that tracks changes in source code during software development. It's designed for coordinating work among programmers, but it can be used to track changes in any set of files.
- GitHub: Describe GitHub as a hosting service for Git repositories. It provides a web-based graphical interface. It also provides access control, collaboration features, task management, and more.

- **Local vs. Remote:**

- Git: Operates on your local computer. You can use Git without GitHub (or any other central repository) to manage your version control.
- GitHub: Acts as a remote repository. It hosts your code and offers a place for collaboration on that code. It's accessible from anywhere via the internet.

- **Collaboration:**

- Git: Handles the version control part but doesn't offer built-in mechanisms for team collaboration like commenting on changes or approving them.
- GitHub: Provides a platform for collaborative features like pull requests (a way to merge changes from one branch to another after peer review), issues, code review, and team management.





# Download and install git

- Windows users: <https://git-scm.com/download/win>
- Mac users: typically git is already installed
- Open git bash or terminal, type: `git --version`

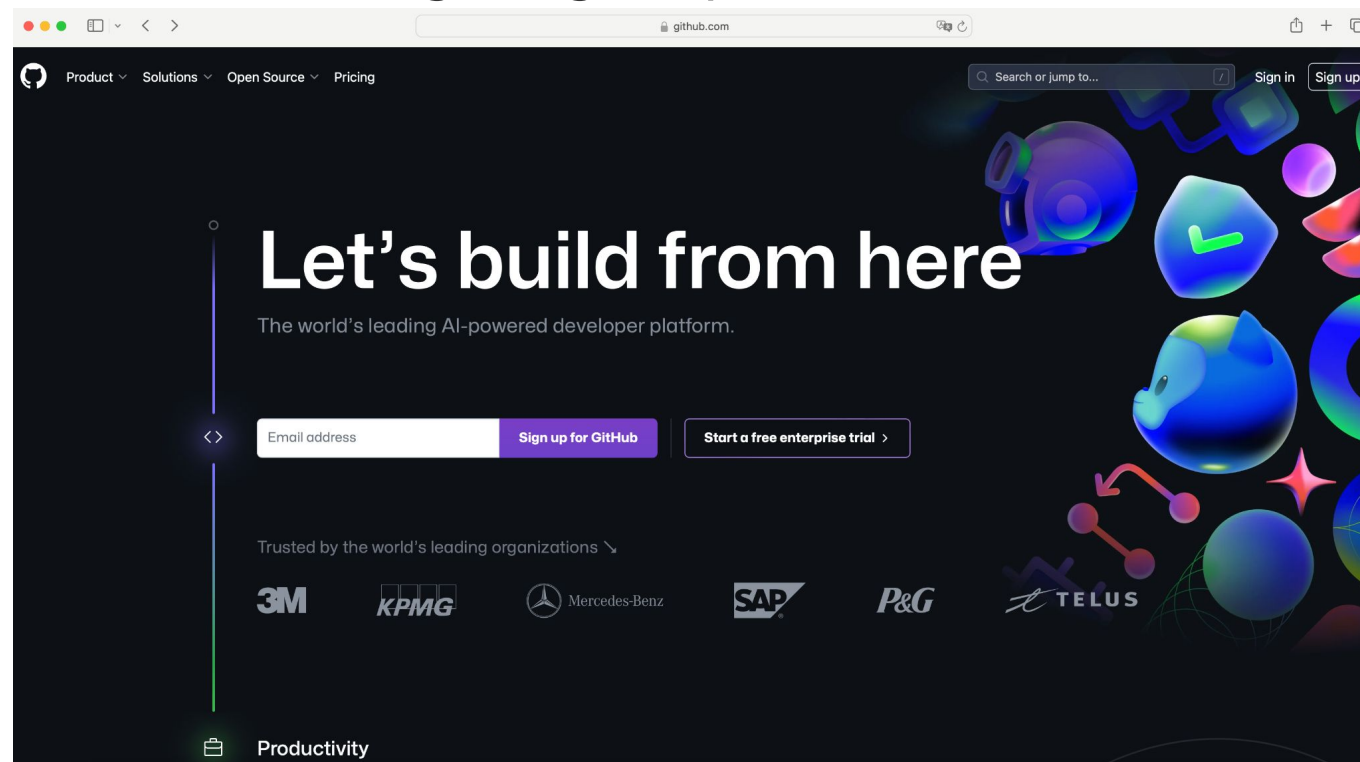
```
MINGW64:/c/Users/Russell IT Spare
Russell IT Spare@RLIT-SPARE01 MINGW64 ~
$ git --version
git version 2.35.0.windows.1
Russell IT Spare@RLIT-SPARE01 MINGW64 ~
$
```



# Create your GitHub account

Navigate to <https://github.com>

Follow the steps after clicking “Sign Up”





# Create your repository on your GitHub

## **Sign In to Your Account:**

- Log in to your GitHub account.

## **Go to Repository Creation Page:**

- Once logged in, you can create a new repository by clicking the "+" icon in the top right corner of the GitHub interface, then selecting "New repository".

## **Repository Details:**

- Repository Name: Give your repository a name. This name should be descriptive and indicate what the repository is for.
- Description (Optional): Provide a brief description of your repository. This helps others understand what your repository is about.
- Visibility: Choose whether the repository is Public (anyone can see this repository) or Private (only you and collaborators you choose can see it).

## **Initialize the Repository (Optional):**

- You can choose to initialize the repository with a README, which is a document explaining what your repository is about.
- You might also want to add a .gitignore file, which tells Git which files or folders to ignore in a project.
- Optionally, you can choose a license for your repository. If you're creating an open-source project, this is important to specify.

## **Create Repository:**

- Click the "Create repository" button to create your new repository.



# Work with Git and GitHub

## Command line tools

windows: git bash

mac: terminal

## What we will do

Config git

Clone a repository

Making Changes locally

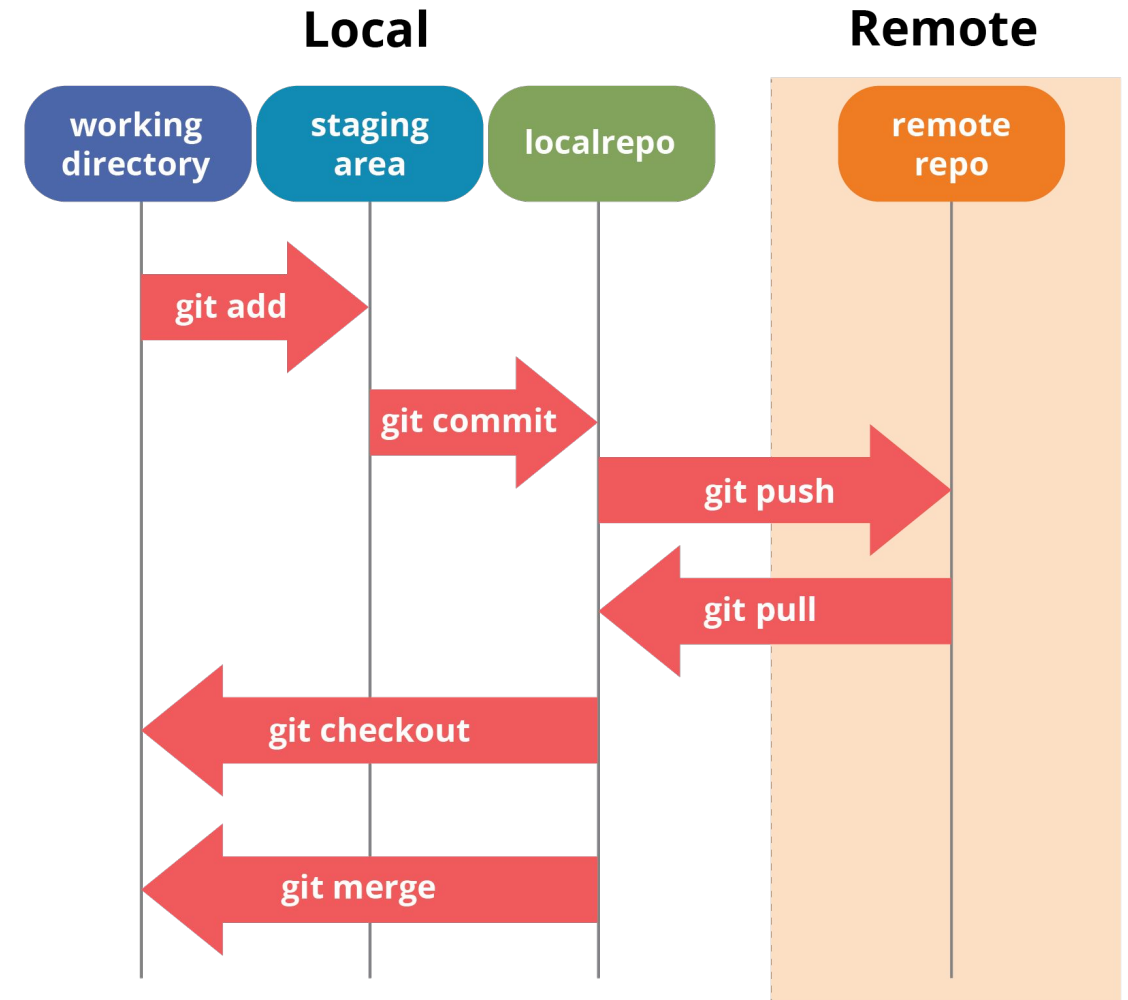
Pushing Changes to GitHub

Pull Changes and Manage Upstream

Handle Merges and Conflicts

<https://education.github.com/git-cheat-sheet-education.pdf>

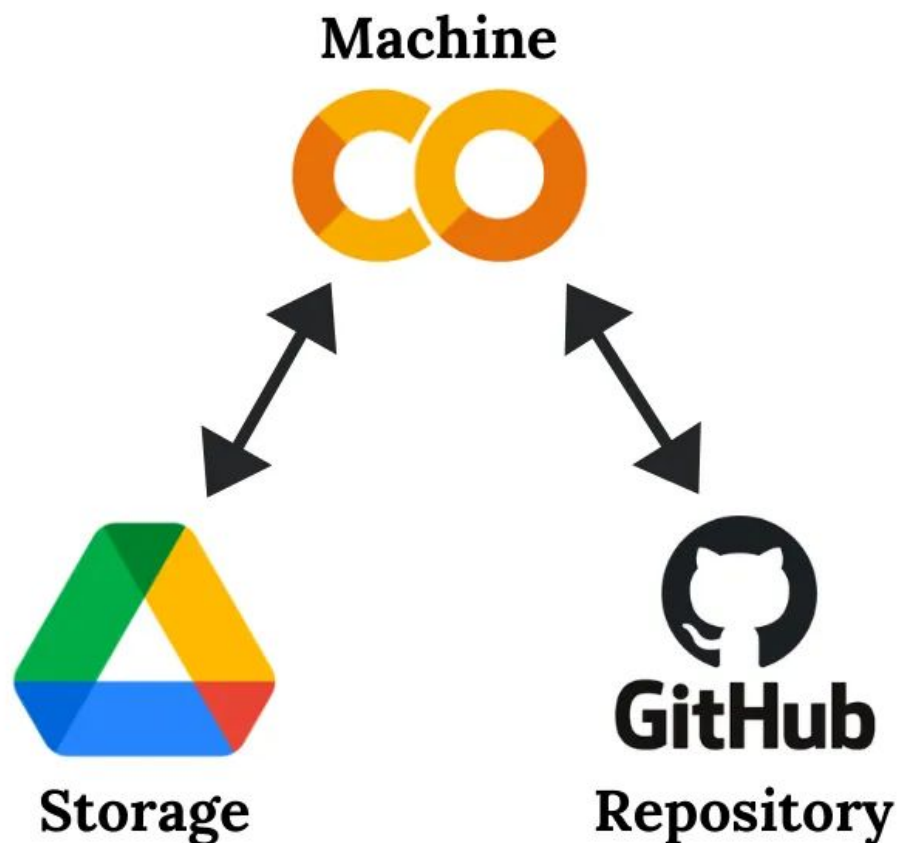
Git Cheat Sheet



<https://www.edureka.co/blog/git-tutorial/>



# Use Google Colab with GitHub via Google Drive



- **Colab** enables anyone to write and run Python code of any kind directly in a web browser.
- **GitHub** serves as a platform for hosting code, facilitating both version control and collaborative work. It provides a space where individuals can collectively contribute to projects from any location, ensuring seamless teamwork while maintaining the integrity of the original project.
- **Google Drive** is a service for storing and synchronizing files. It allows users to save their files on cloud servers, keep files updated across various devices, and share them with others.

<https://medium.com/analytics-vidhya/how-to-use-google-colab-with-github-via-google-drive-68efb23a42d>





# Homework assignment

See Canvas

# Next Lecture: Fundamentals of Python

