

Workshop: Version control for more effective collaboration - Introducing Git and GitHub (3 hrs)

Attribution Most of the content in this workshop is from lessons by [Software Carpentry](#) released under [CC BY 4.0-license](#).

Workshop objectives

- Introduce the concept of version control and why it's useful for scientists.
- Introduce git version control tool and it's commandline interface.
- Demonstrate the basics of Git.
- Demonstrate the basics of GitHub web service.

Description

A version control system is a tool for managing changes to a set of files. Version control is better than mailing files back and forth:

- Nothing that is committed to version control is ever lost. Since all old versions of files are saved, it's always possible to go back in time to see exactly who wrote what on a particular day.
- With full change history, we know who to ask if we have questions later on, and, if need be, revert to a previous version, much like the “undo” feature in an editor.
- When several people collaborate in the same project, it's possible to accidentally overlook or overwrite someones changes: the version control system automatically notifies users whenever there's a conflict between one person's work and another's.

Version control is the lab notebook of the digital world: it's what professionals use to keep track of what they've done and to collaborate with other people. Every large software development project relies on it, and most programmers use it for their small jobs as well. And it isn't just for software: books, papers, small data sets, and anything that changes over time or needs to be shared can and should be stored in a version control system.

[Git](#) has quickly become the most popular version control software around. It is very versatile and fast tool, which unfortunately has a bit steep learning curve. However, because of its popularity, there's plenty of good and easy to approach documentation around.

[GitHub](#) is a web-based service that puts a lot of emphasis on the social aspect of code and content sharing. GitHub is a company that hosts Git repositories in the web and provides a web interface to interact with repos they host.

Prerequisites

A laptop (Windows, Mac or Linux) and a working installation of git. More complete installation instructions will be distributed later. Registering to [GitHub](#) before the workshop is highly recommended!

In this hands-on workshop we will be using git on command line. There are [several graphical user interfaces](#) available for git, but the command line interface is far better for learning the conceptual underpinnings.

Recommended reading

- Hampton S.E., Anderson S.S., Bagby S.C., Gries C., Han X., Hart E.M., Jones M.B., Lenhardt C., MacDonald A., Michener W.K., Mudge J., Pourmokhtarian A., Schildhauer M.P., Woo K.H., & Zimmerman N. (2015) [The Tao of open science for ecology](#). Ecosphere, 6, 1–13.
- Ram, K. (2013): [Git can facilitate greater reproducibility and increased transparency in science](#). Source Code for Biology and Medicine, 8(1), 1–14.