

Version Control with Git

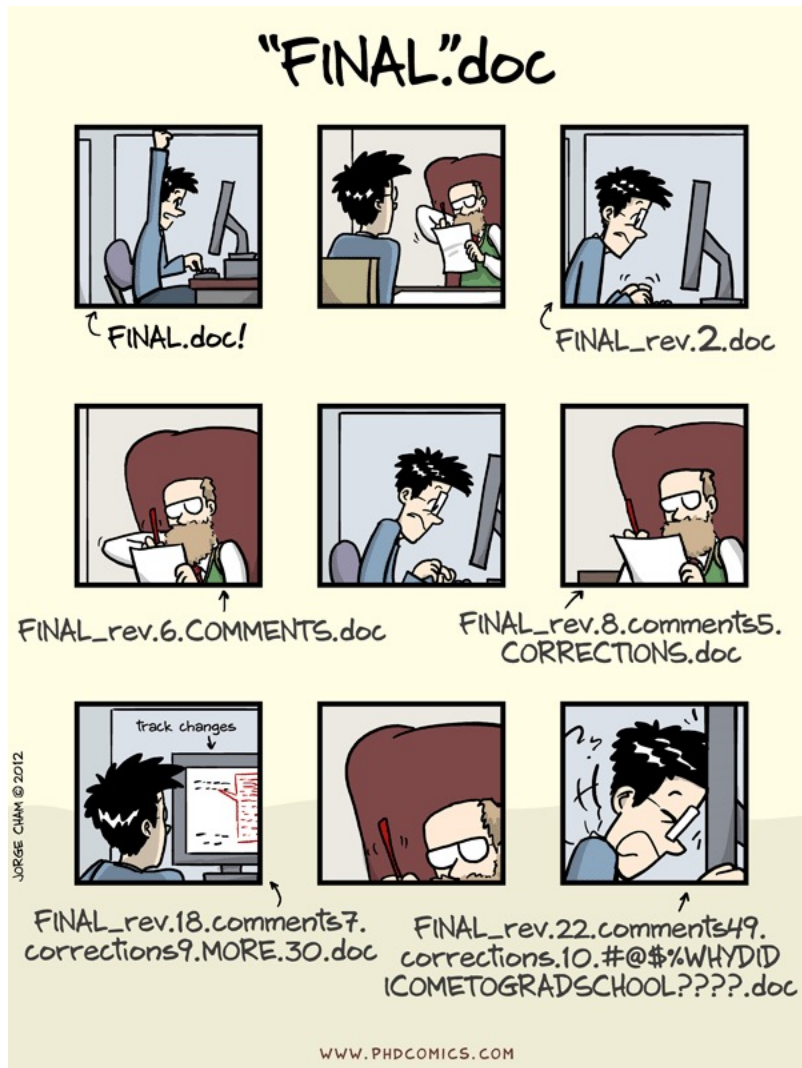
Data Science Initiative Workshop Series

Fall 2021



UNIVERSITY OF
OREGON

What is version control?



- Version control is an organized way of maintaining a record of changes
- Git is a system for distributed version control – not the only one, but popular among scientists
- Some advantages to using Git/version control:
 - Enhance reproducibility
 - Fix mistakes by reverting to earlier versions
 - Improve project structure
 - Backup versions in remote repositories**
 - Facilitate collaboration**

The mechanics of version control

Single user:



Multiple users:



Some Git vocabulary

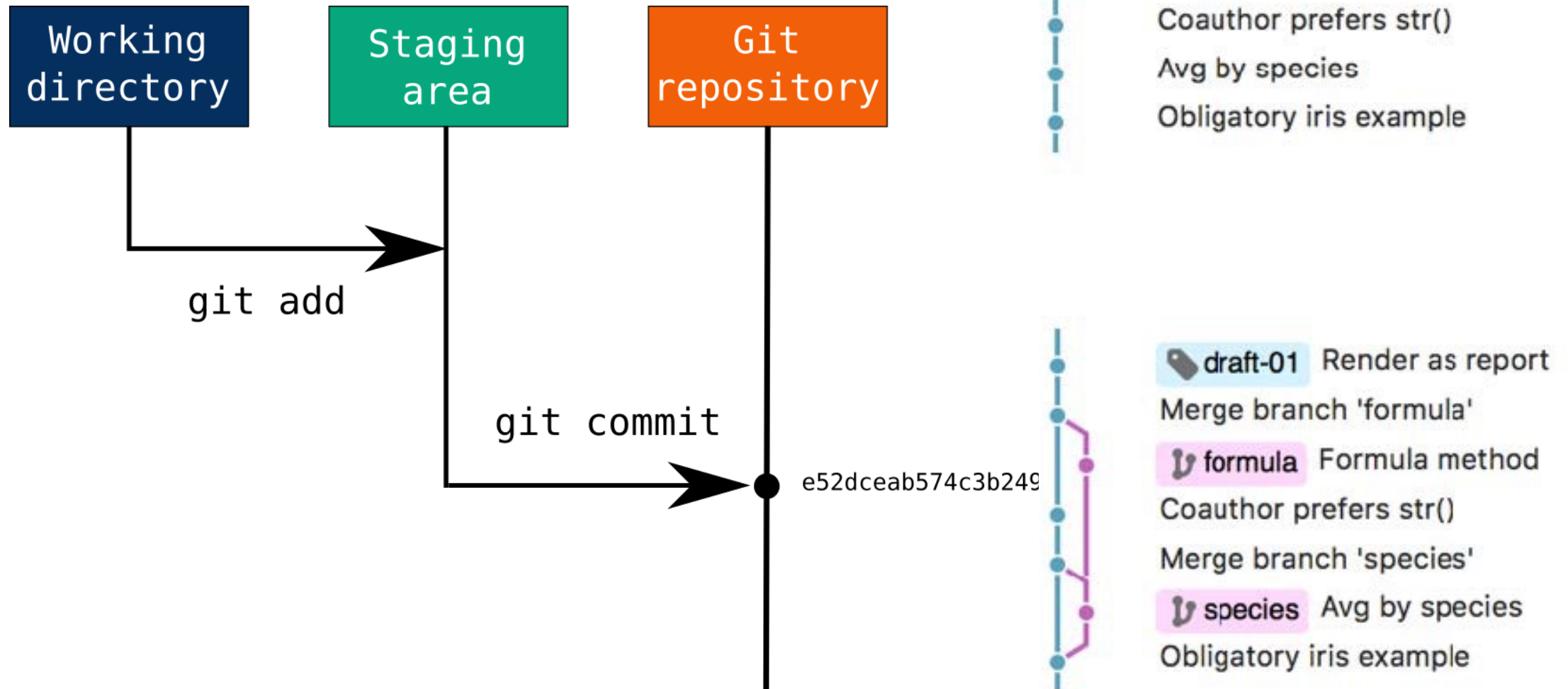
Repository/repo – the collection of files and directories associated with a project and tracked with version control

Commit – a snapshot of a repository's history that is recorded by Git

Diff – Changes in the repository's content associated with the commit

Branches – Concurrent work (changes to file content) can occur in parallel branches, so that you can focus on developing one aspect of the repository/project independently

The basic Git workflow



Using Git

- In the shell
- With a Git client