

Version control for reproducible science

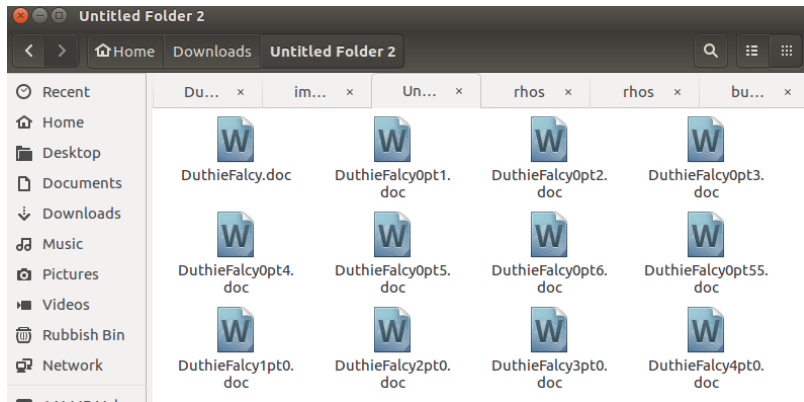
Brad Duthie

4 December 2019


Rough outline of version control workshop


1. **14:00** What is version control, and why use it?
2. **14:20** Getting set up – good file management
3. **14:30** The [GitKraken](#) interface and simple commits
4. **15:00** Setting up [GitHub](#), pushing and pulling
5. **15:30** Branching using [GitKraken](#)
6. **16:00** Merging and merge conflicts
7. **16:30** Independent work using version control



What is version control, and why use it?






What is version control, and why use it?


 Commits on Mar 1, 2019


Some work on the SI
 bradduthie committed on 1 Mar 2019 ✓



 5d29331 


Major restructure and revision of the Discussion to compare to Gibbs
 bradduthie committed on 1 Mar 2019 ✓
...et al.


 a40ede5 



 Commits on Feb 27, 2019


Edit the manuscript up to Reviewer 2 specific comment 3 -- these are
 bradduthie committed on 27 Feb 2019 ✓
...next on the list, specifically the new Discussion paragraph



 d411fd5 

 Commits on Feb 22, 2019

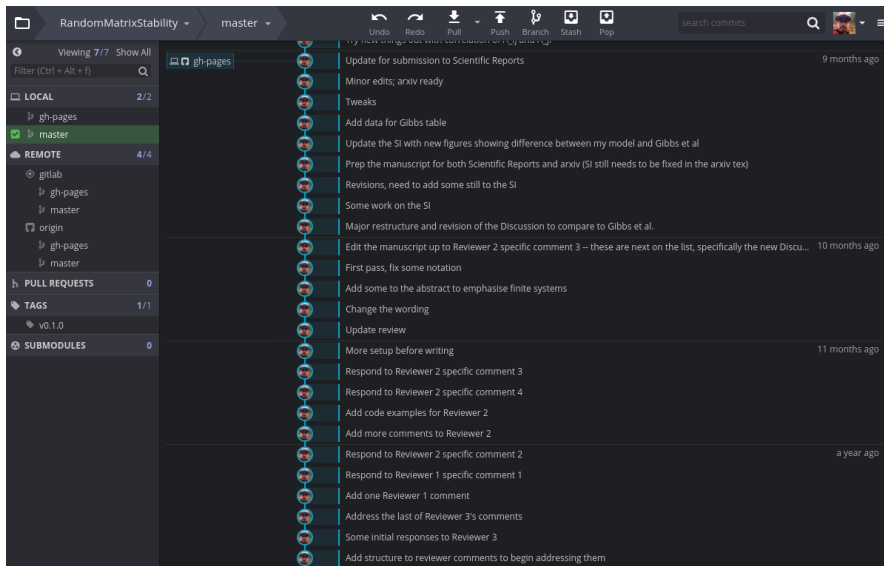
First pass, fix some notation
 bradduthie committed on 22 Feb 2019 ✓

 39f854b 

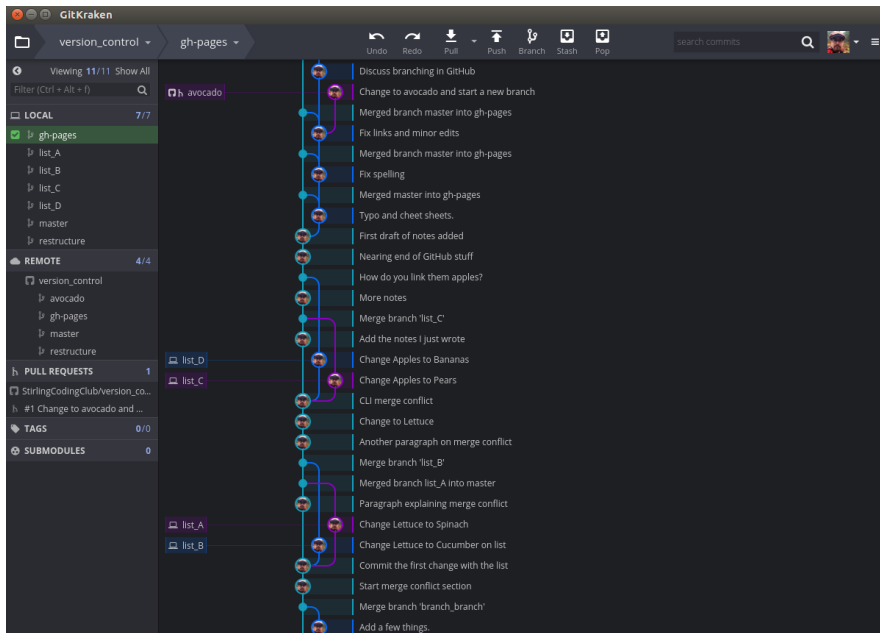
Add some to the abstract to emphasise finite systems
 bradduthie committed on 22 Feb 2019 ✓

 612aa4b 

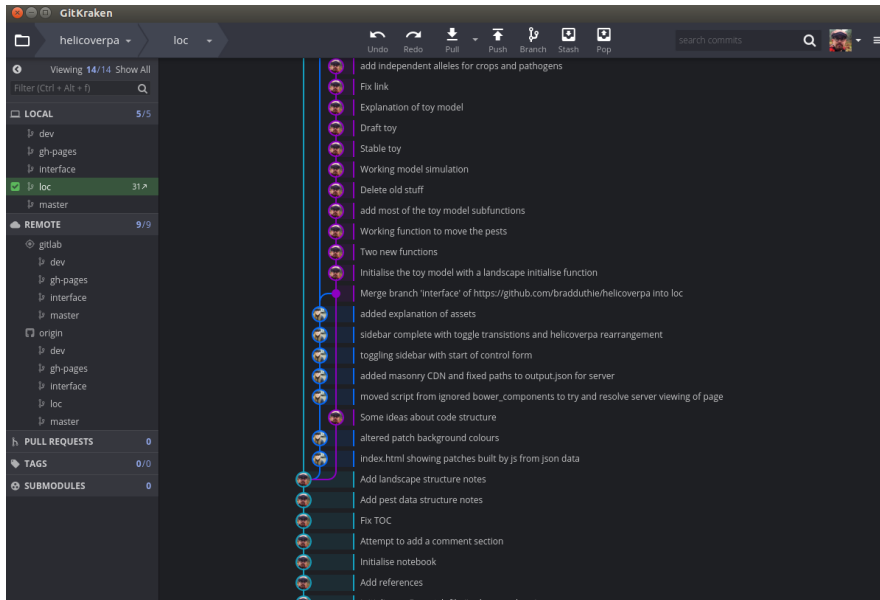
What is version control, and why use it?



What is version control, and why use it?



What is version control, and why use it?



What version control software does

- ▶ Software that records changes you make to files over time
 - ▶ Manage different *versions* of files
 - ▶ Recover old files, keep track of file changes
 - ▶ Collaborate with others on shared files

What version control software does

- ▶ Software that records changes you make to files over time
 - ▶ Manage different *versions* of files
 - ▶ Recover old files, keep track of file changes
 - ▶ Collaborate with others on shared files
-
- ▶ **Put more intuitively**, version control takes a snapshot in time (called a '**commit**') of all the files in one of your folders (called '**repositories**')
 - ▶ Visualise changes to your files over time
 - ▶ Look at the differences between file versions
 - ▶ Record who changed files, and what they changed

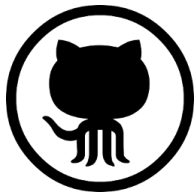
Version control makes science easier

- ▶ Organise files and avoid having to 'save as' multiple versions
 - ▶ analysis_1.R
 - ▶ analysis_2.R
 - ▶ analysis_FINAL.R
 - ▶ analysis_FINAL_no_really_this_time.R

Version control makes science easier

- ▶ Organise files and avoid having to 'save as' multiple versions
 - ▶ analysis_1.R
 - ▶ analysis_2.R
 - ▶ analysis_FINAL.R
 - ▶ analysis_FINAL_no_really_this_time.R
- ▶ Have a clear history of what you have done, when, and why (through commit comments)
- ▶ Never worry about losing your data, analysis, or writing when integrating with [GitHub](#)

Version control can help open science



- ▶ Transparent record of data collection, analysis, and writing
- ▶ Record publicly available on [GitHub](#), [Bitbucket](#), or [GitLab](#)
- ▶ GitHub repository can be copied, reproduced, and discussed
- ▶ [git](#) and GitHub can track individual contributions to a project

Most researchers use git (and GitHub)



- ▶ Free and open-source
- ▶ Separate from [GitHub](#)

Most researchers use git (and GitHub)



- ▶ Free and open-source
- ▶ Separate from [GitHub](#)
- ▶ Works across platforms
 - ▶ Windows
 - ▶ Linux
 - ▶ Mac
- ▶ Invented by [Linus Torvalds](#)

Objectives: using version control

By the end of today you will be able to use git with [GitHub](#) and [GitKraken](#) to manage your projects with version control.

Objectives: using version control

By the end of today you will be able to use git with [GitHub](#) and [GitKraken](#) to manage your projects with version control.

- ▶ Understand key concepts of version control
- ▶ Perform basic tasks in GitHub and GitKraken
 - ▶ Staging and committing files
 - ▶ Pushing to and pulling from GitHub
 - ▶ Branching, merging, & resolving merge conflicts

Objectives: using version control

By the end of today you will be able to use git with [GitHub](#) and [GitKraken](#) to manage your projects with version control.

- ▶ Understand key concepts of version control
- ▶ Perform basic tasks in GitHub and GitKraken
 - ▶ Staging and committing files
 - ▶ Pushing to and pulling from GitHub
 - ▶ Branching, merging, & resolving merge conflicts

Discuss, share, and get additional help by [raising an issue](#) in the [version_control repository](#) on the [Stirling Coding Club](#) GitHub organisation.

Why focus on using GitKraken?



- ▶ Free to download and use
- ▶ Easy GitHub integration
- ▶ Graphical user interface
- ▶ Visualisation of repository

Accompanying notes to these slides are available in the **version_control** repository, and include instructions for using the command line interface, and for editing directly in GitHub.