

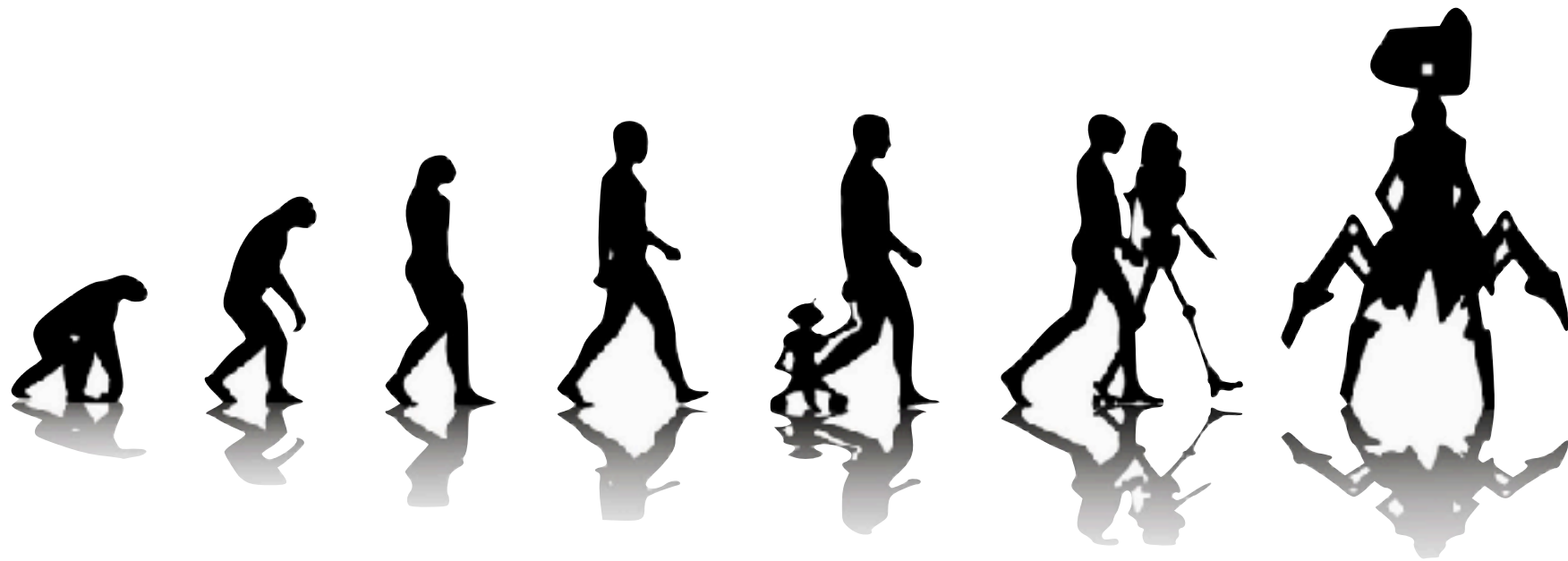
Version Control: The Gain You Get for Your Pain

Jennifer Bryan
RStudio, UBC

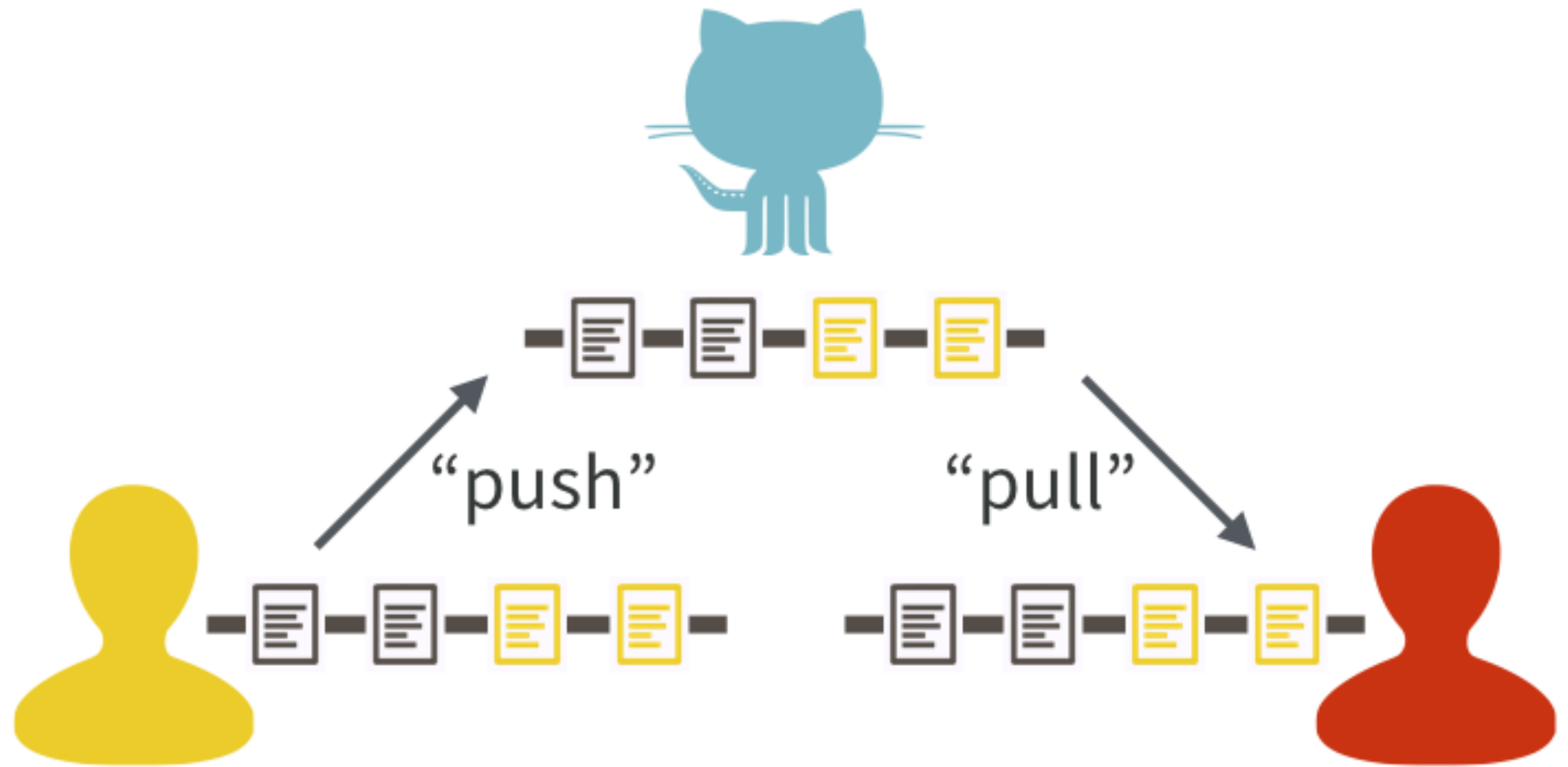
 @JennyBryan

 @jennybc

 jennybryan.org



version control manages
the evolution a set of files



hosted version control lets you do this with friends (and strangers)

Why version control in your teaching life?

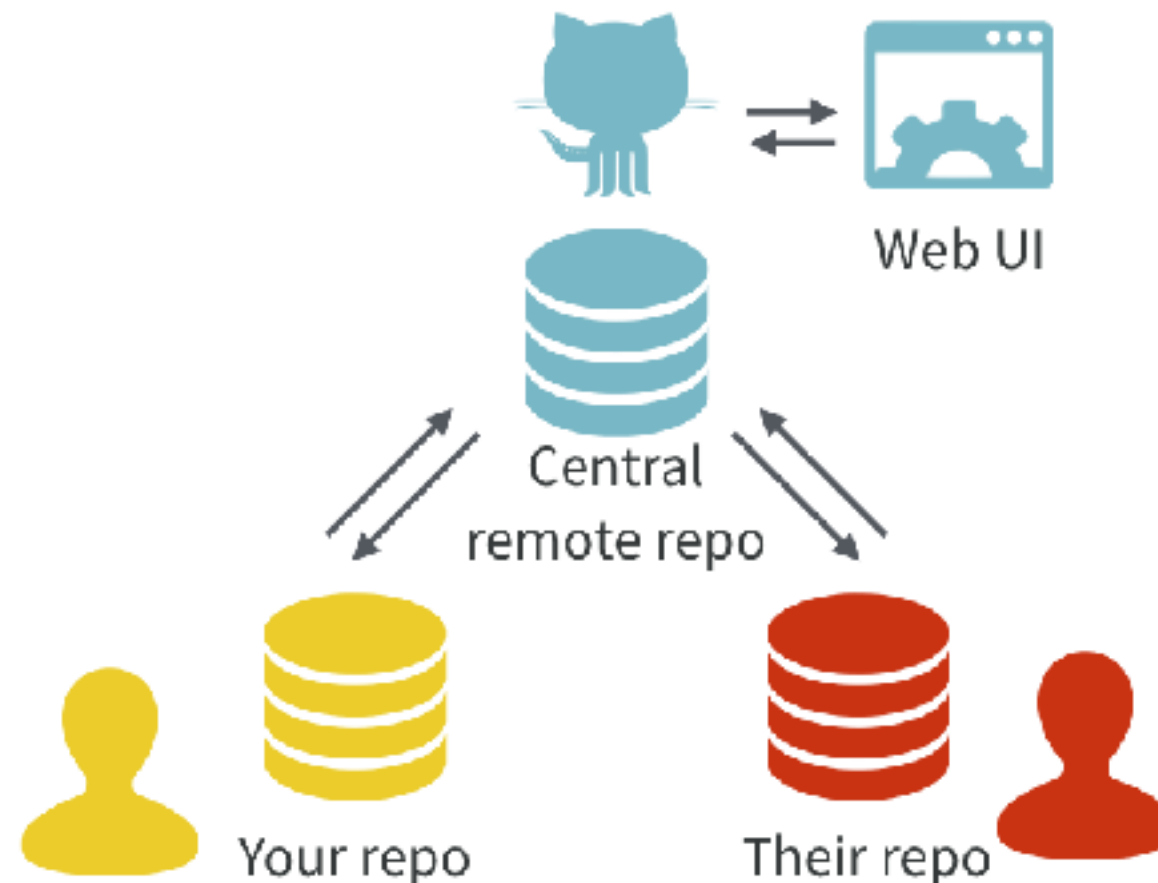
1. Useful for your course mechanics.
2. Students should have basic familiarity.
3. Supports adoption in rest of your life.

Excuse me, do you have a moment to talk about version control?

available in:

PeerJ Collection "Practical Data Science for Stats"

The American Statistician "Special Issue on Data Science"



Practical Data Science for Stats

<http://bit.ly/practical-data-sci>



[Submit an article](#)[New content alerts](#)[RSS](#)[Citation search](#)[Current issue](#) [Browse list of issues](#)

< Issue 2 **Issue 1** >

Special Issue on Data Science

Editorial

Editorial

[Introduction: Special Issue on Data Science](#) >

Hadley Wickham, Jennifer Bryan & Nicole Lazar

Page: 1

Published online: 24 Apr 2018

[Citation](#) | [Full Text](#) | [PDF \(192 KB\)](#) |

1045

Views

0

CrossRef citations

3

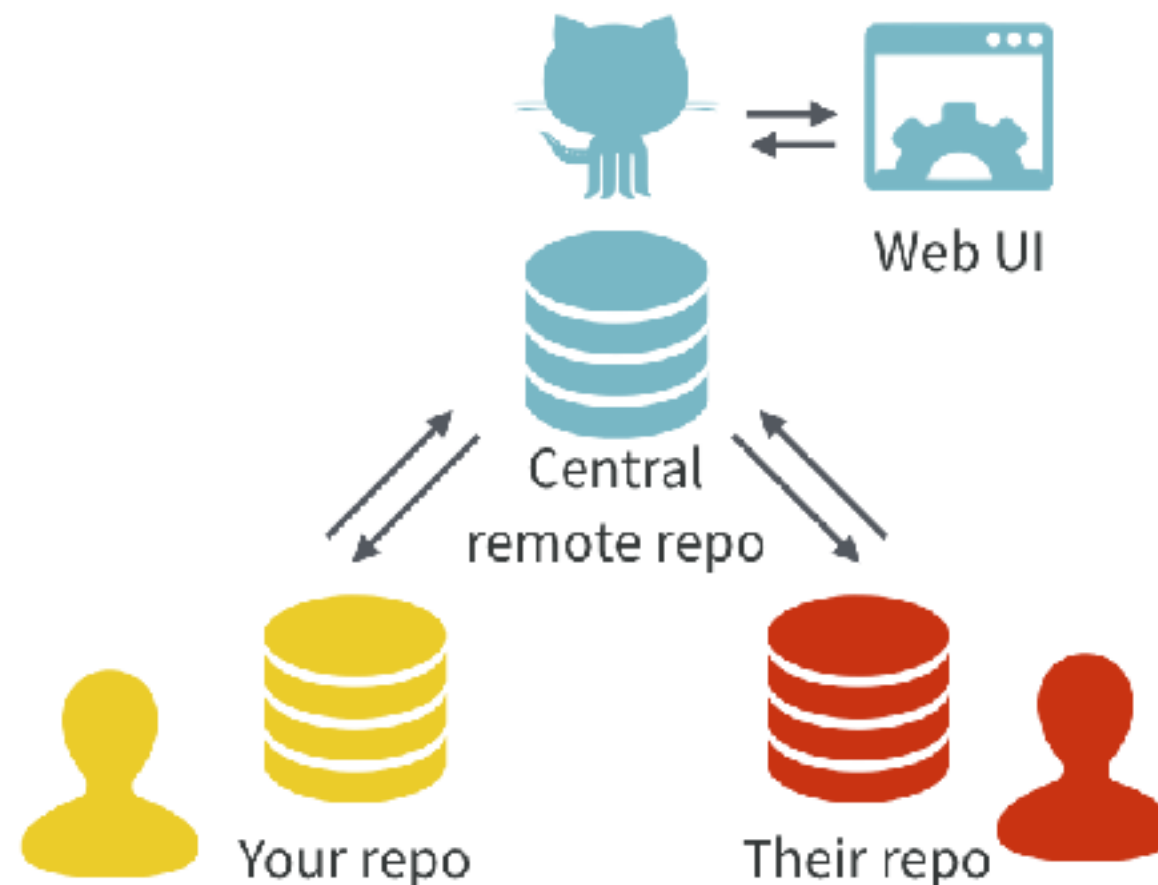
Altmetric

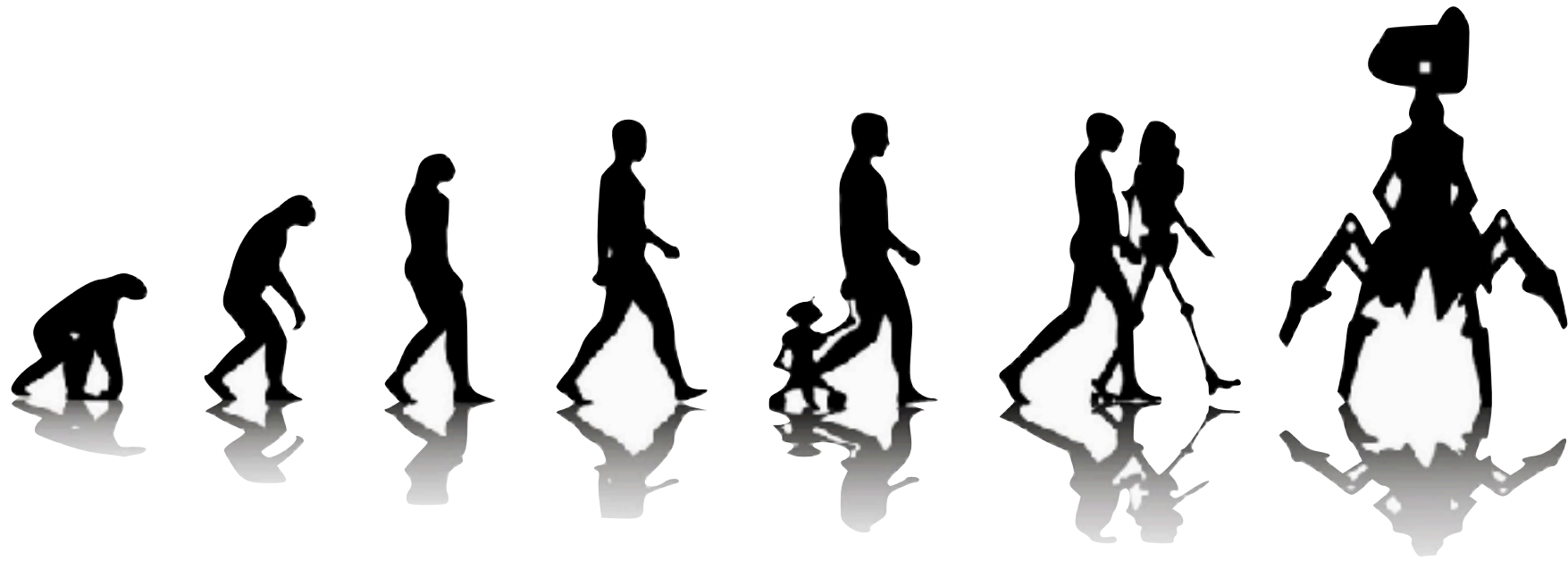


Excuse me, do you have a moment to talk about version control?

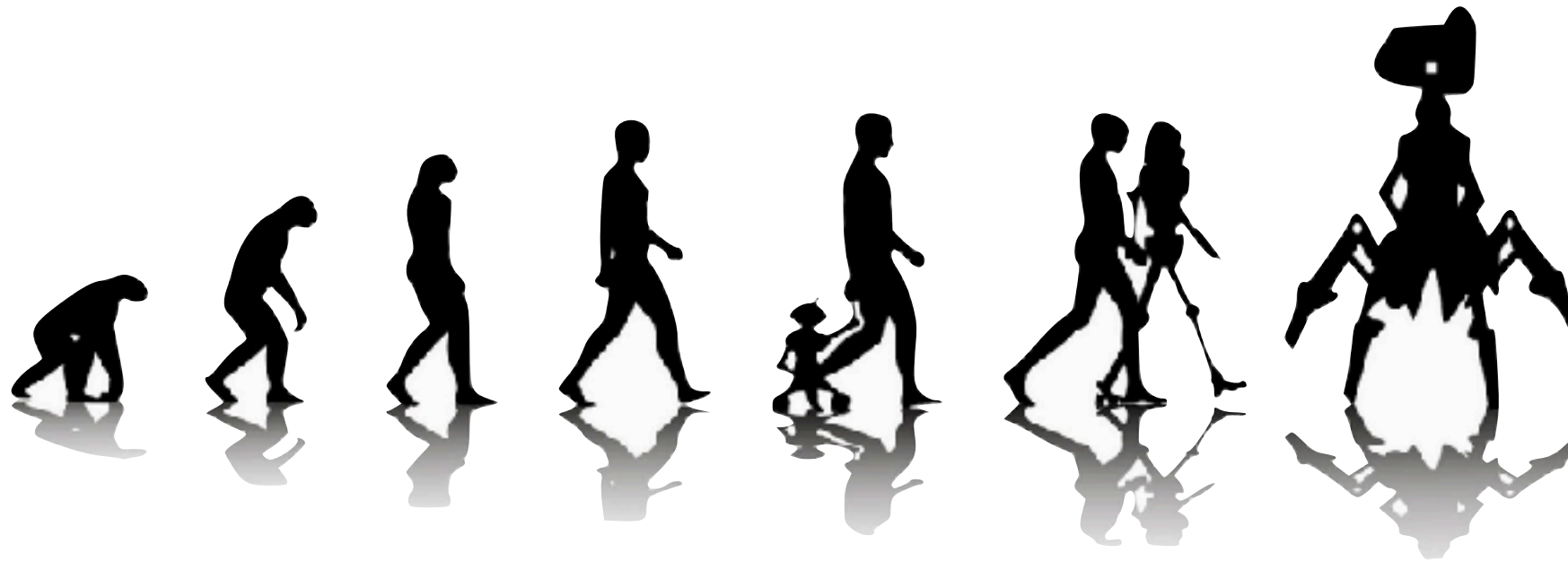
<https://doi.org/10.7287/peerj.preprints.3159v2>

<https://doi.org/10.1080/00031305.2017.1399928>

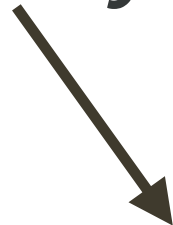




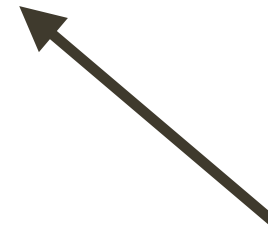
I use **Git + GitHub**



version control system



I use **Git + GitHub**



"hosted version control"

"the remote"

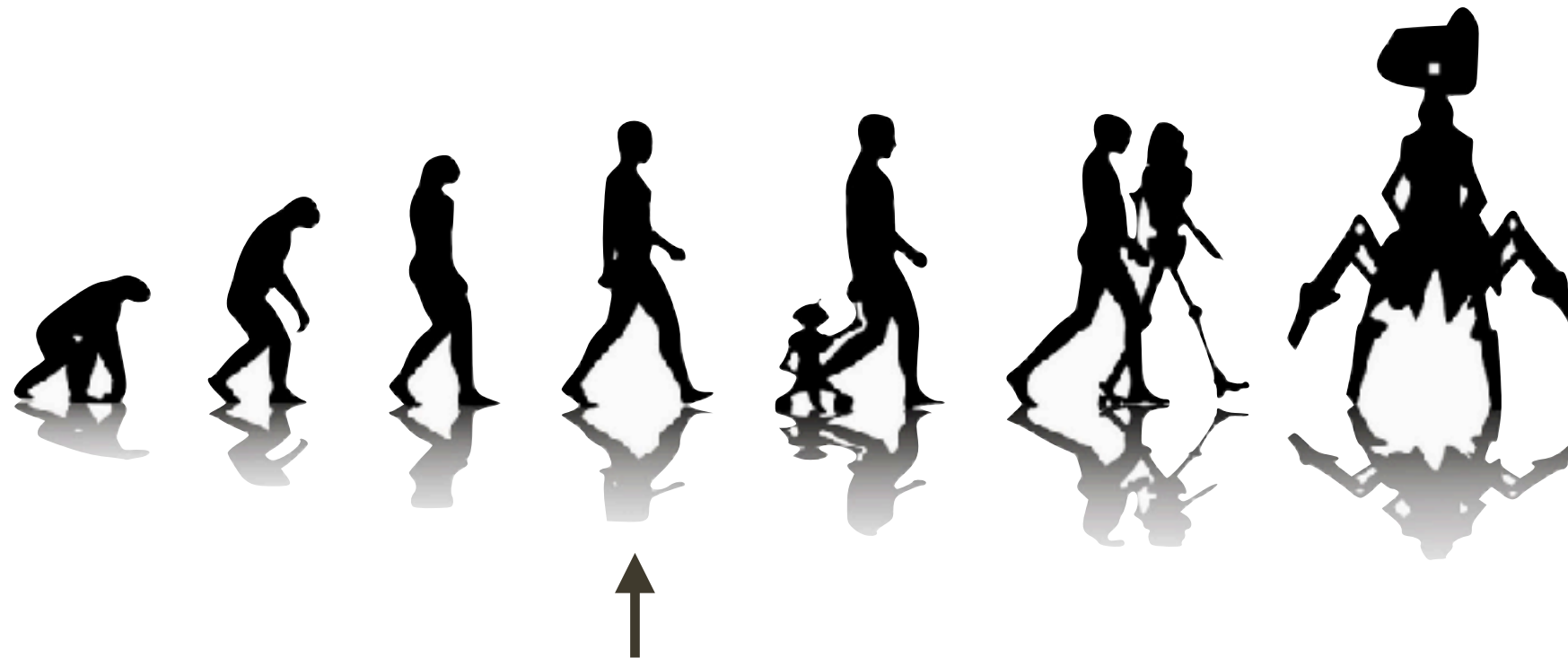
What does version control feel like?

Define a set of files to manage holistically.

A **repository**, a.k.a. "repo"

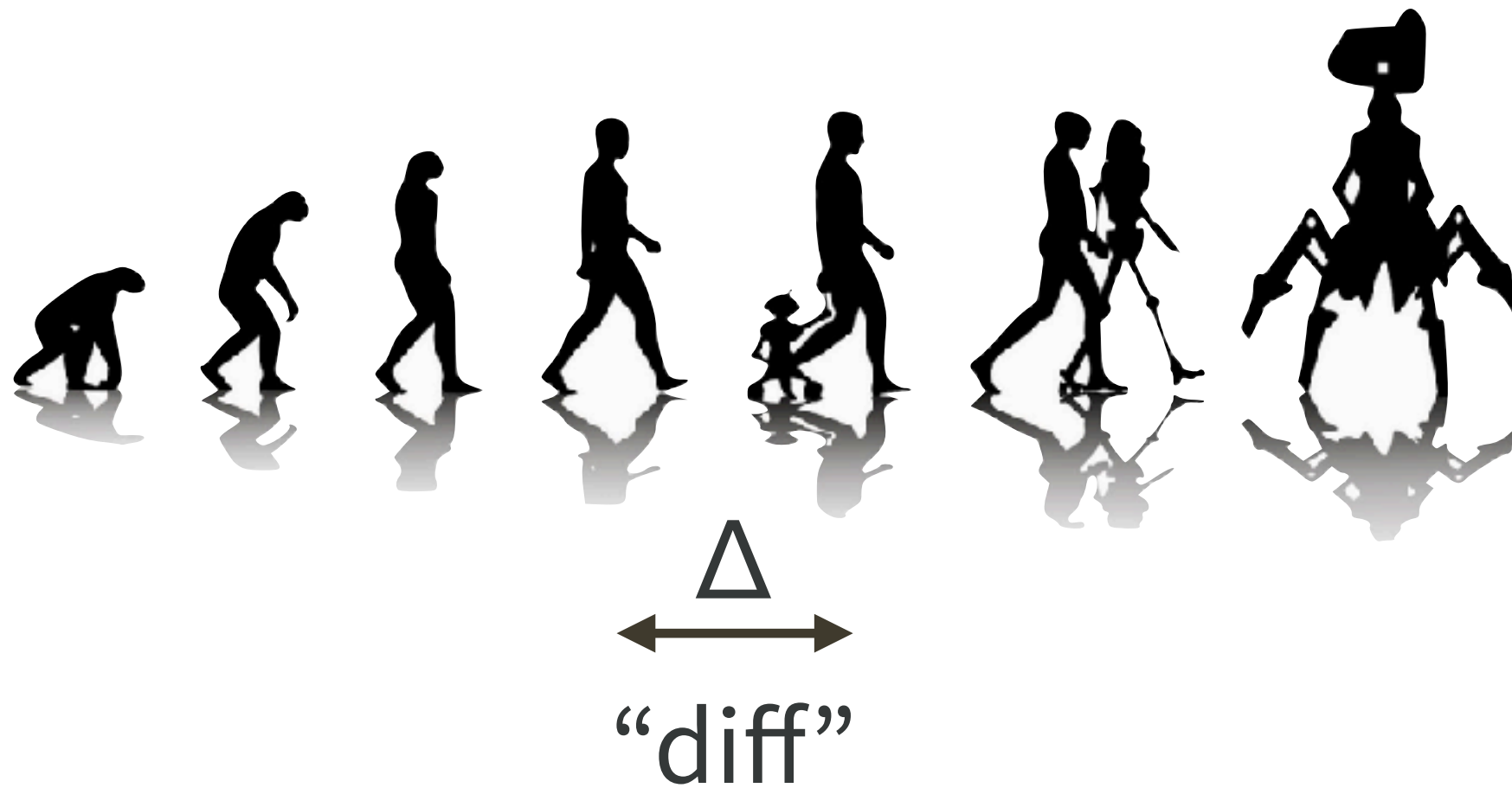
A **folder** on your computer

Do at start or *post hoc*

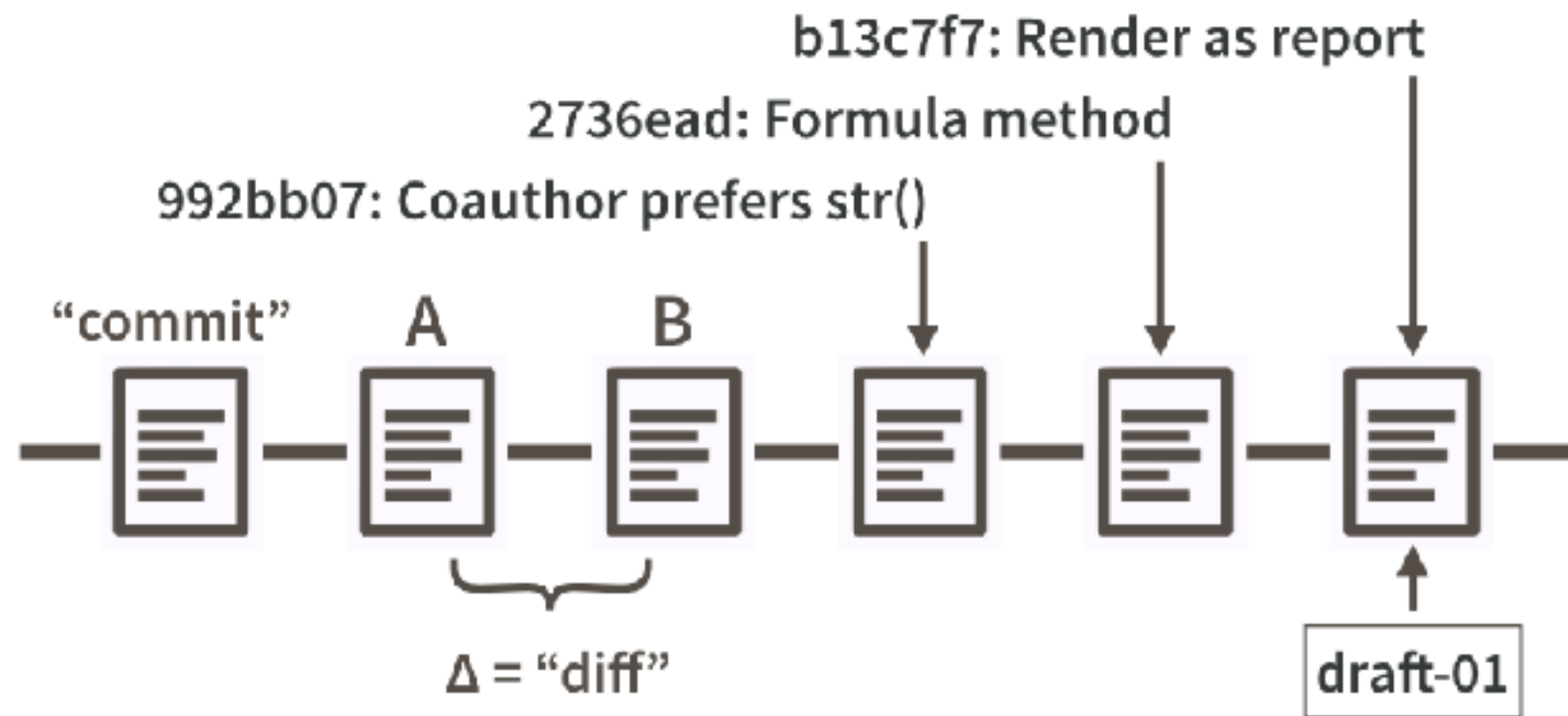


“commit”

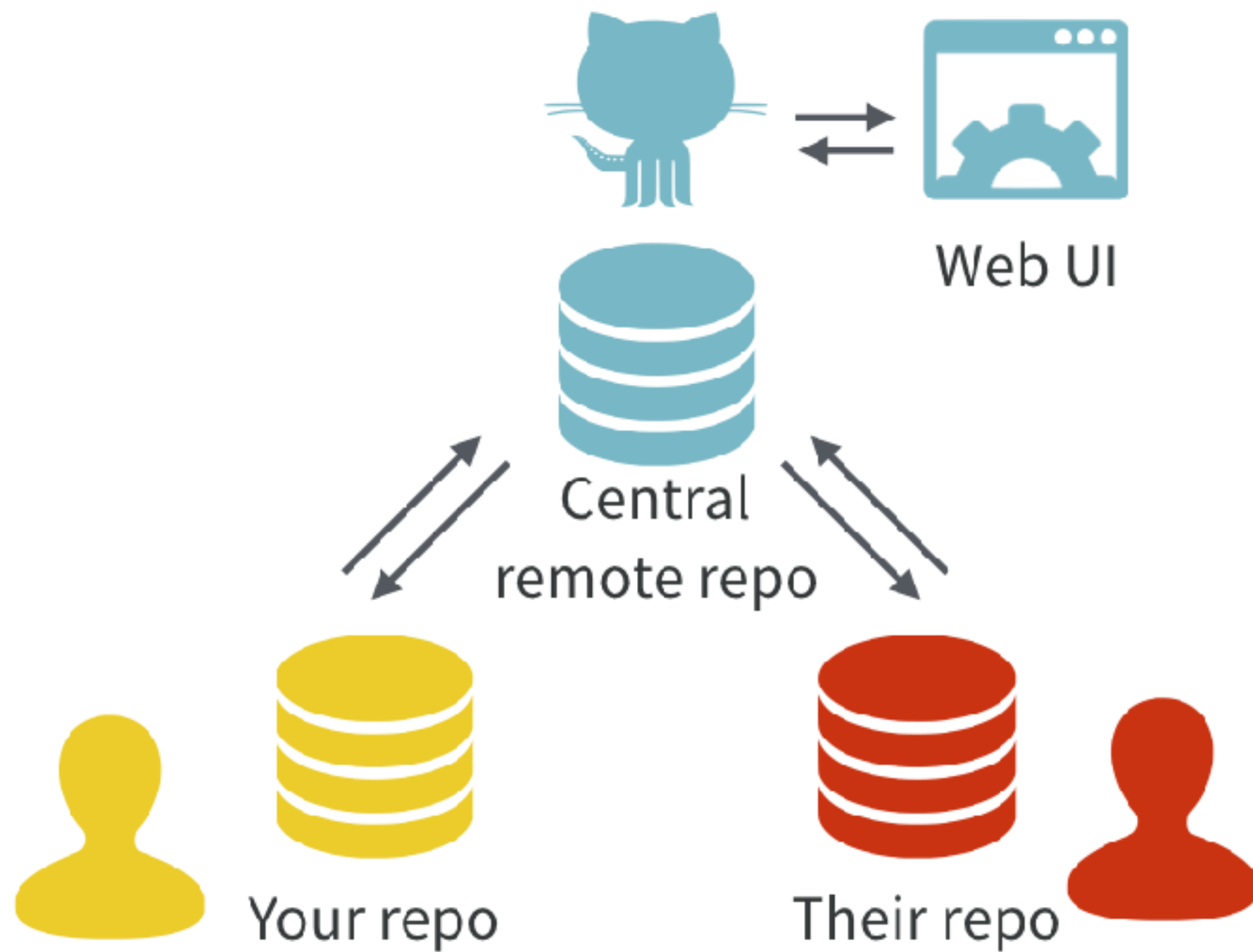
a file or project state that is **meaningful to you**
for inspection, comparison, restoration



What changed here?
Why?

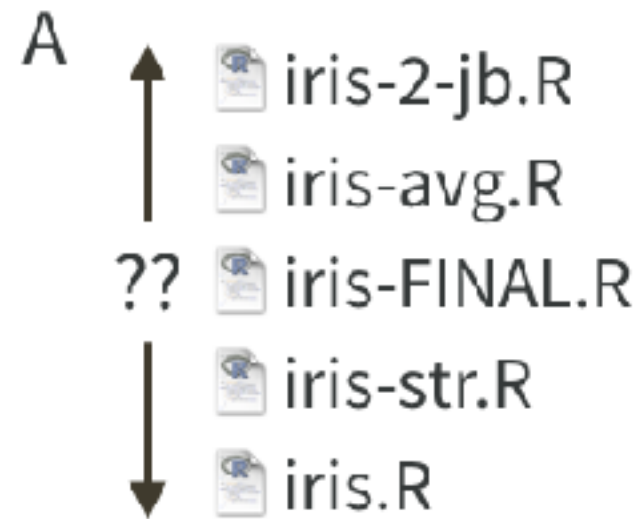


Work as you normally do, but
periodically make a commit

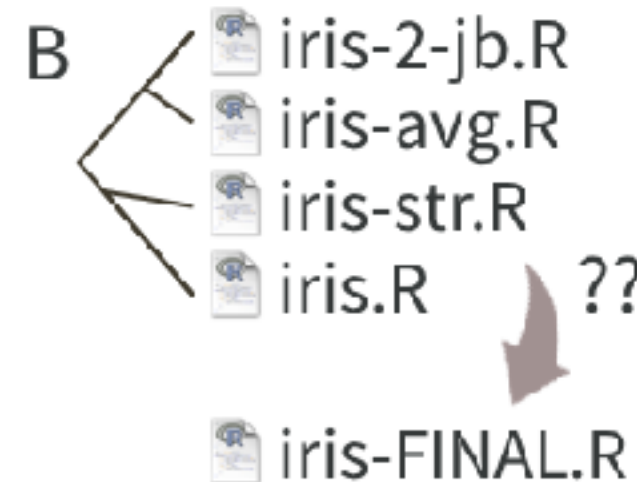


Periodically sync up with a
"remote" copy of this repo

How different from other workflows?



draft-01 Render as report
Formula method
Coauthor prefers str()
Avg by species
Obligatory iris example

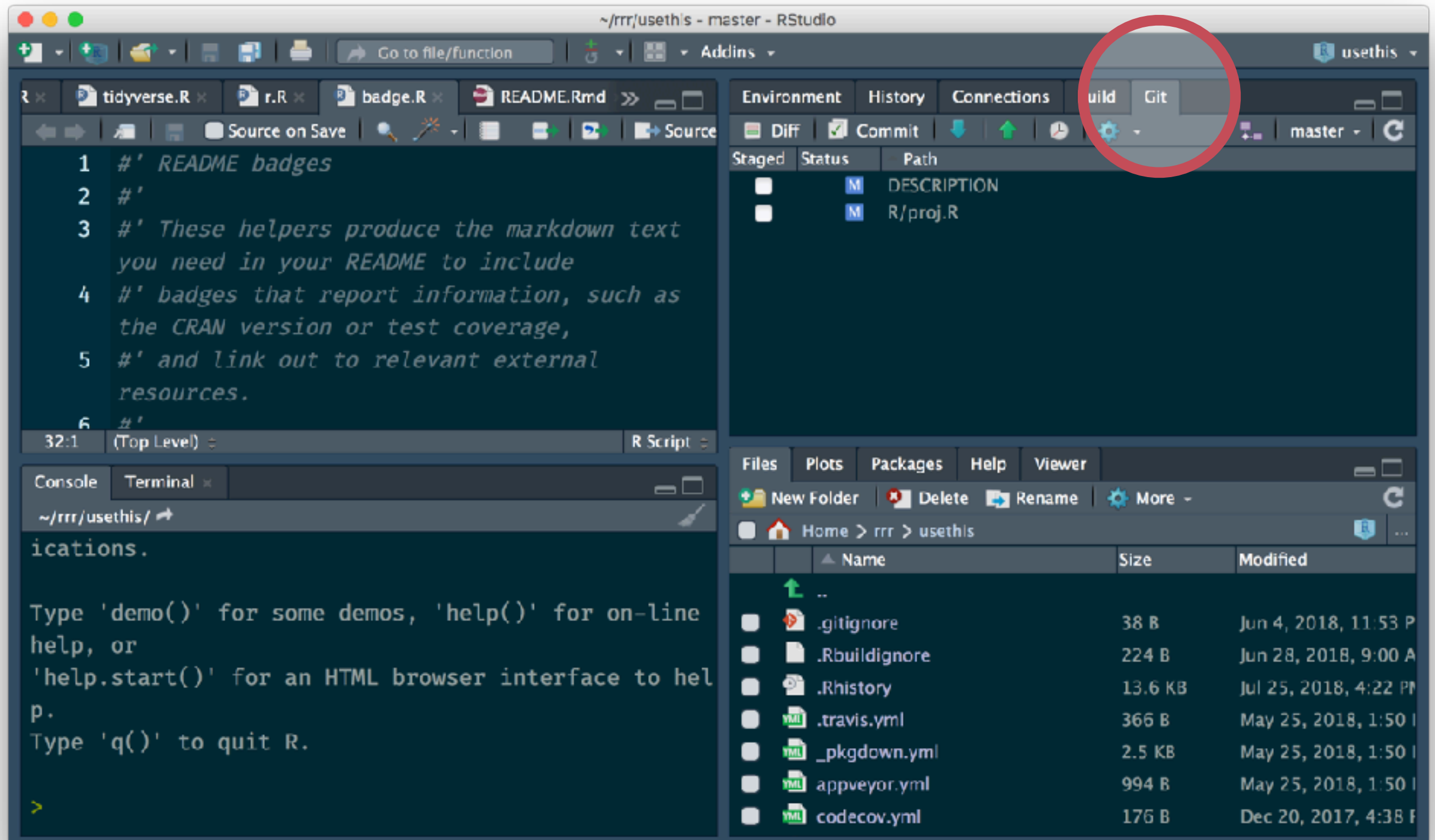


draft-01 Render as report
Merge branch 'formula'
 formula Formula method
Coauthor prefers str()
Merge branch 'species'
 species Avg by species
Obligatory iris example

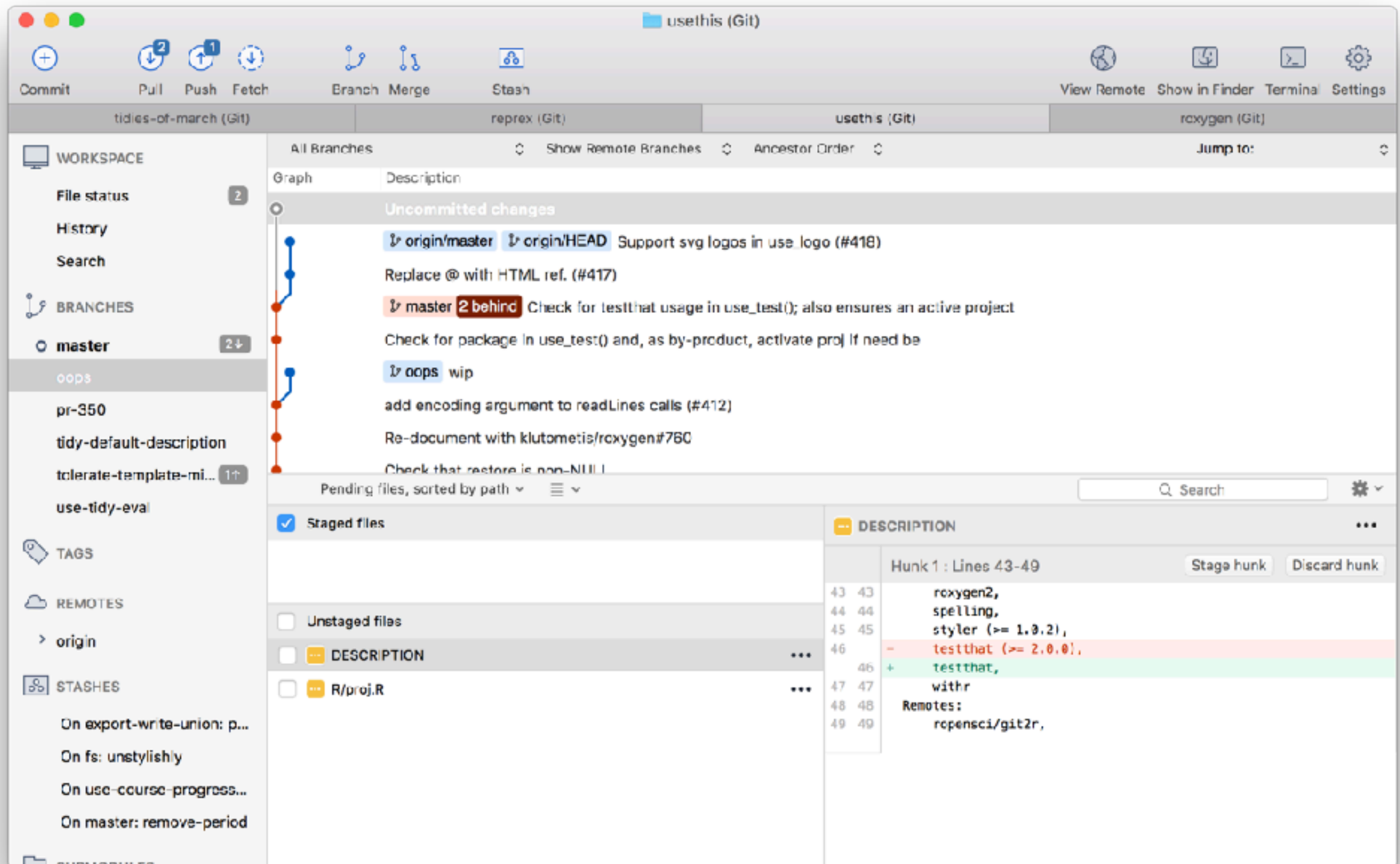
Tips and gotchas

- Don't use Git without also using GitHub (or similar)
- Use in a project with an audience or collaborators (not solo)
- De-emphasize Word docs, Excel sheets, PDFs
- Emphasize Markdown, HTML, R code, csv files
- Have no shame about using modern Git front-ends, i.e. a nice GUI locally and GitHub remotely
- It's OK to ignore branching, rebasing, and other intermediate/advanced features until needed.

RStudio knows about Git!



SourceTree: my favorite Git GUI



Same repo, as it looks on GitHub

The screenshot shows the GitHub interface for the `r-lib / usethis` repository. The browser address bar shows `github.com/r-lib/usethis`. The repository page includes a search bar, navigation links for Pull requests, Issues, Marketplace, and Explore, and a header with repository statistics: 12 Unwatch, 240 Unstar, and 73 Fork. Below the header, there are tabs for Code, Issues (75), Pull requests (5), Projects (0), Insights, and Settings. A section titled "Set up commonly used components" with a link to `http://usethis.r-lib.org` and an "Edit" button is present. Below this, a bar shows repository statistics: 1,311 commits, 8 branches, 6 releases, and 64 contributors. A row of buttons includes "Branch: master", "New pull request", "Create new file", "Upload files", "Find file", and "Clone or download". The main content area displays a list of commits, with the latest commit by `Jimhester` and `Jennybc` titled "Support svg logos in use_logo (#418)" dated 3 days ago. Below the commit list, a table shows the repository structure with folders and their associated commit messages and dates.

Folder	Commit Message	Time Ago
<code>.github</code>	Mod CODEOWNERS to stop auto-notification (#400)	a month ago
<code>R</code>	Support svg logos in use_logo (#418)	3 days ago
<code>docs</code>	Switch to fs (#385)	a month ago
<code>inst</code>	Update wording around CoC to match the README snippet	2 months ago
<code>man</code>	Re-document with klutometis/roxygen#760	14 days ago

happygitwithr.com

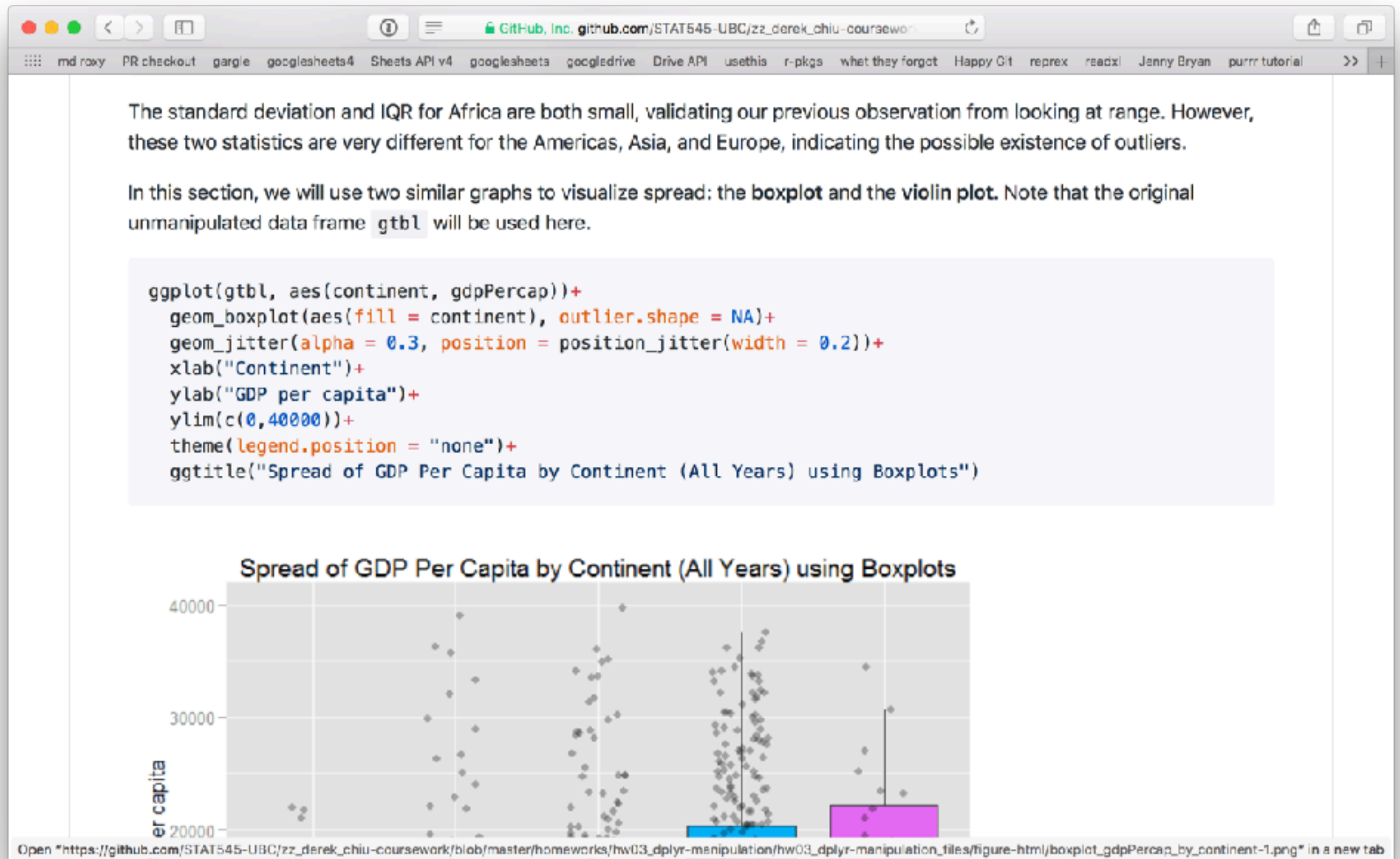


1 current instructor, 1 past instructor, 1 kind stranger,
all editing the STAT545 website

The screenshot shows a GitHub commit history page for the repository `STAT545-UBC/STAT545-UBC.github.io`. The page lists several commits, with four specific ones highlighted by red circles:

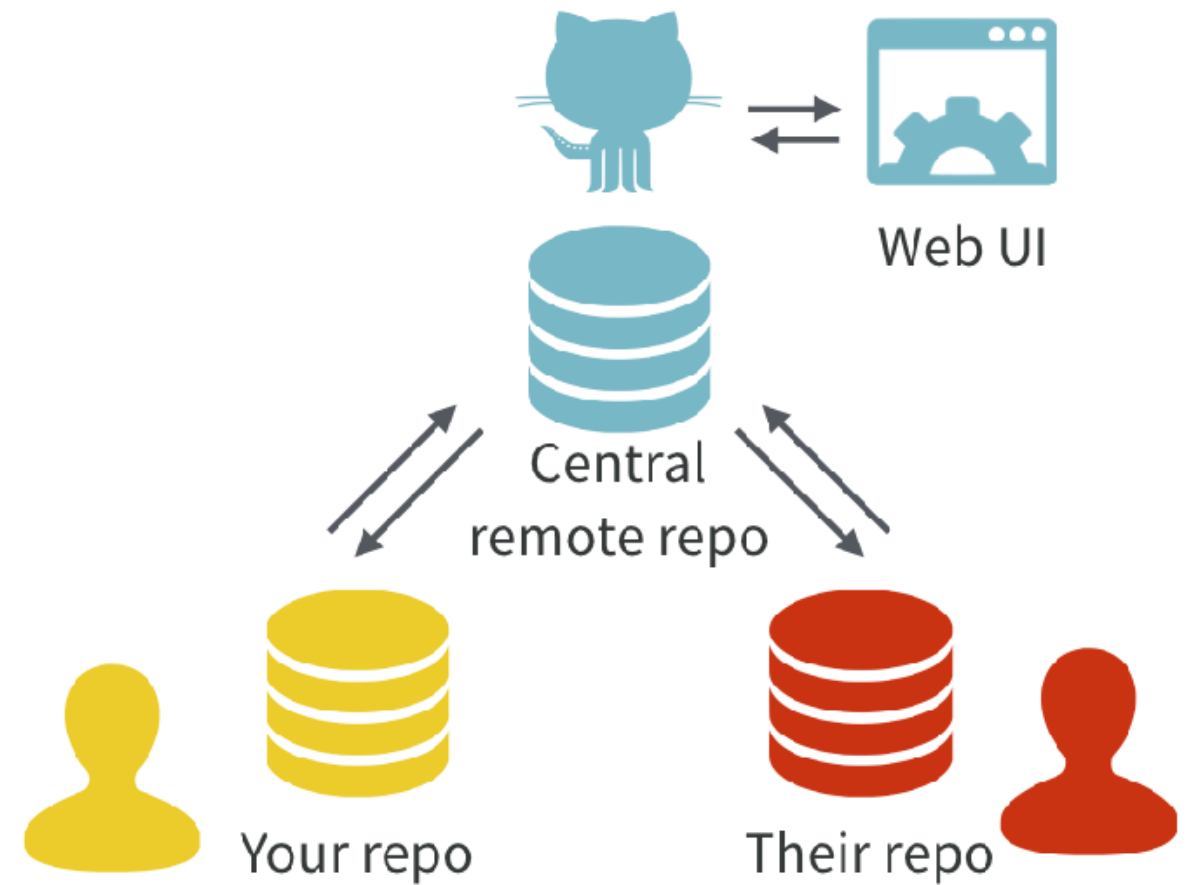
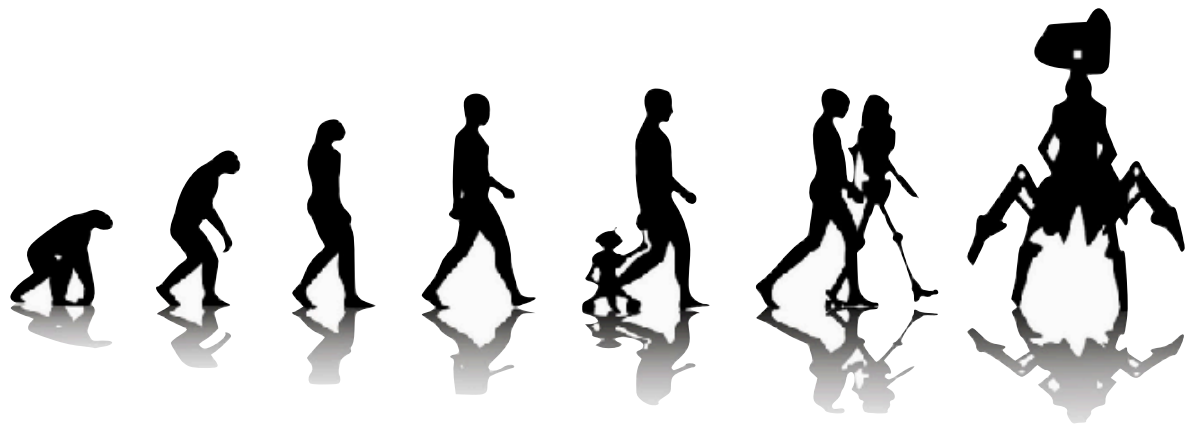
- scraping after class** by `gvdr` committed on Nov 30, 2017. Commit hash: `5455613`.
- hw09 peerreview** by `gvdr` committed on Nov 30, 2017. Commit hash: `bd30379`.
- Merge branch 'markdly-markdly-patch-1'** by `jennybc` committed on Nov 28, 2017. Commit hash: `1e5da0d`.
- Re-render the dplyr join cheatsheet** by `jennybc` committed on Nov 28, 2017. Commit hash: `7fd03c4`.
- new syllabus and everything** by `gvdr` committed on Nov 28, 2017. Commit hash: `38aef32`.
- new 111lm after lecture** by `gvdr` committed on Nov 28, 2017. Commit hash: `9885c13`.
- fixing overwidth issue** by `markdly` committed on Nov 28, 2017. Commit hash: `d22d1f2`.
- remove unnecessary tr tag** by `markdly` committed on Nov 28, 2017. Commit hash: `6d7b51b`.
- html cm111** by `gvdr` committed on Nov 28, 2017. Commit hash: `708f5d8`.

Each STAT545 student gets a repo to hold their coursework, all done in R and R Markdown



Sampler of resources re: Git(Hub) in the classroom

- Excuse Me, ...
- [Teaching with GitHub](#), blog post re: UBC Master of Data Science Program
- [Run a course with GitHub](#) in [happygitwithr.com](#)
- Cetinkaya-Rundel M, Rundel CW. (2017) [Infrastructure and tools for teaching computing throughout the statistical curriculum](#). In PeerJ Collection "Practical Data Science for Stats" and The American Statistician "Special Issue on Data Science"
- [education.github.com](#)
- [The Plain Person's Guide to Plain Text Social Science](#), Kieran Healy
- [Why you should use GitHub: Lessons for the classroom and newsroom](#), guest blog post by Alexey Zagalsky for [Storybench](#)



 @JennyBryan

 @jennybc

 jennybryan.org