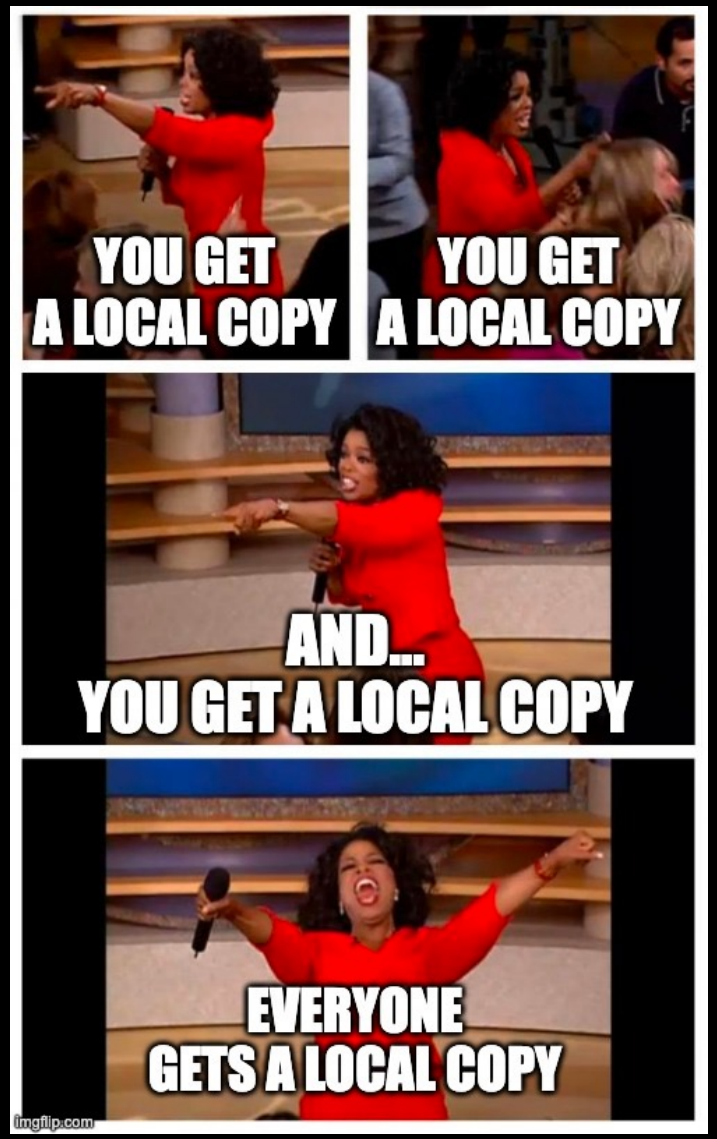
Intro to GitHub

Dr. Theresa Laverty and Dr. Erica Christensen

31 March 2023

Why should you use GitHub?

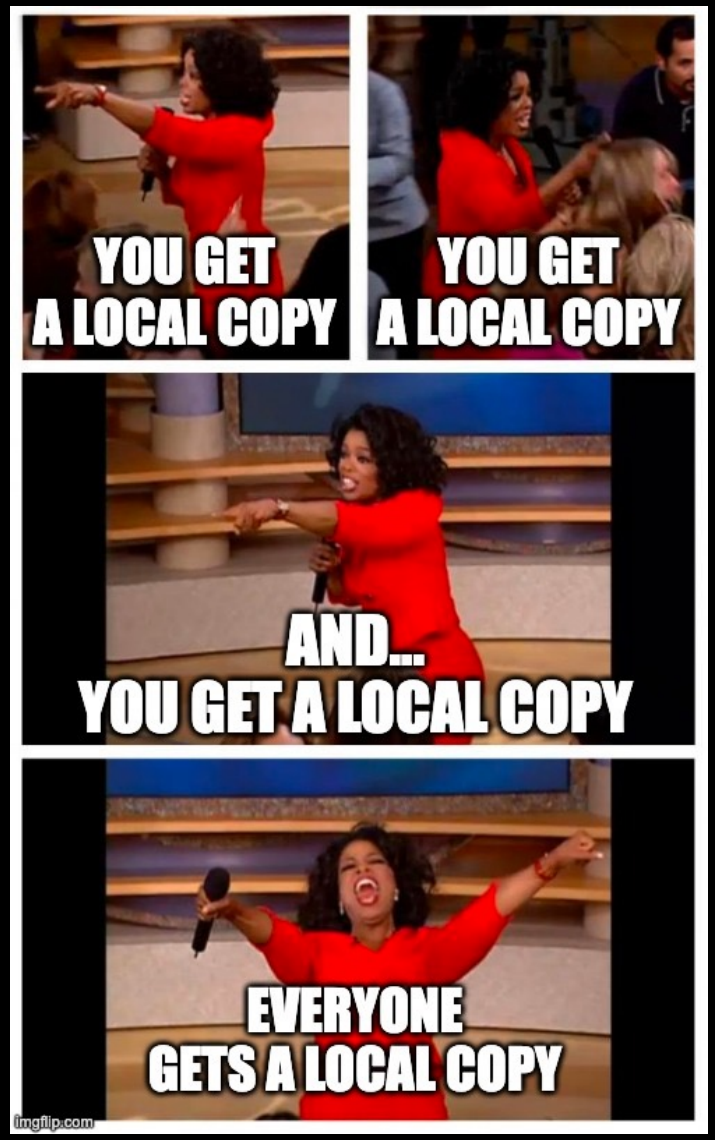
- Version control
  - Like an unlimited ‘undo’
  - Allows many people to work in parallel
  - Like "Track Changes" but better!



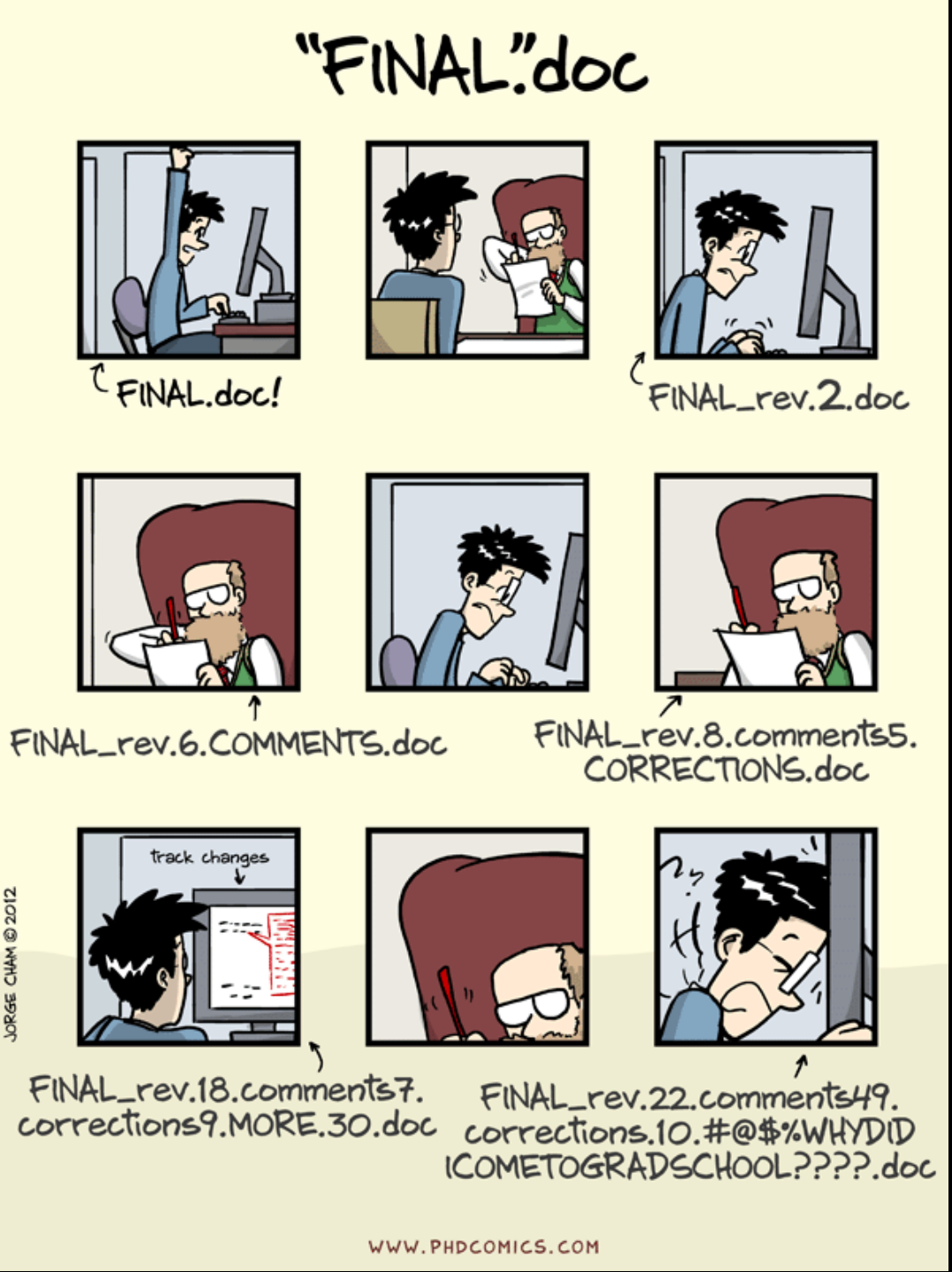
Why should you use GitHub?

- Version control
  - Like an unlimited ‘undo’
  - Allows many people to work in parallel
  - Like "Track Changes" but better!

… BUT not the best for large file storage
  - There are ways to work around this

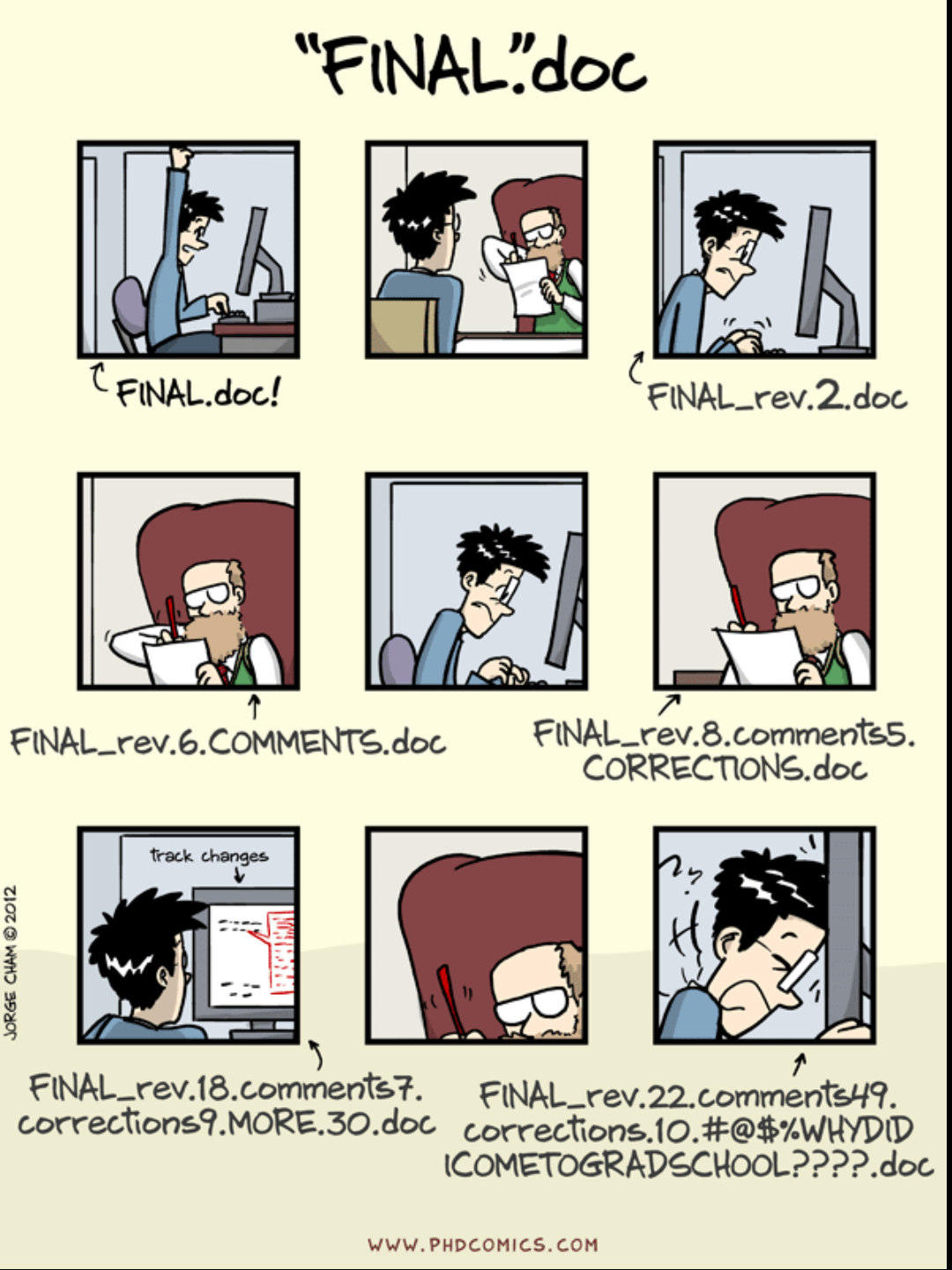


"FINAL".doc



One user:

Icons: https://swcarpentry.github.io/git-novice/01-basics/index.html



One user:

Two users:

Icons: https://swcarpentry.github.io/git-novice/01-basics/index.html

How Git works

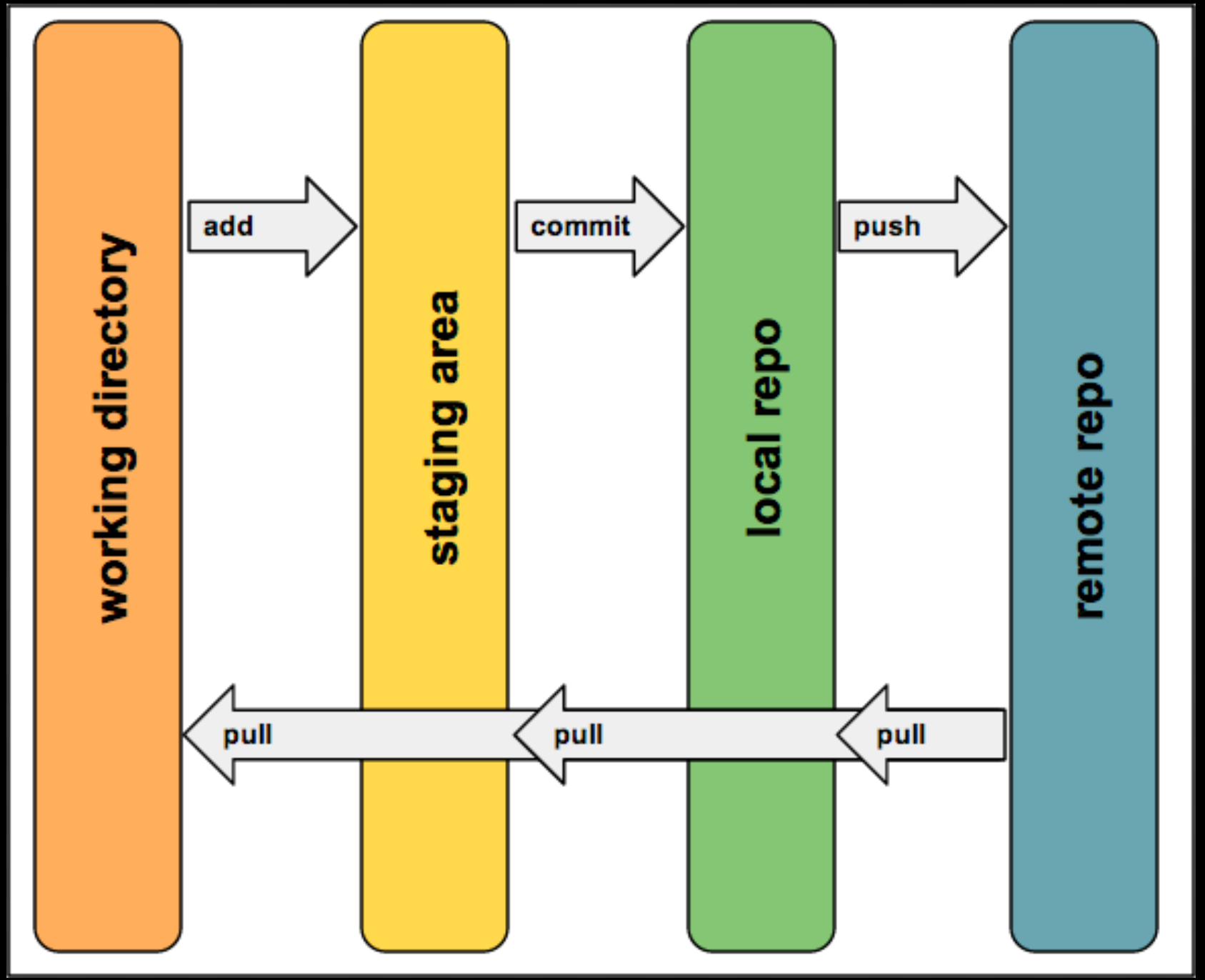


Image: https://hwheeler01.github.io/CompBio/github/

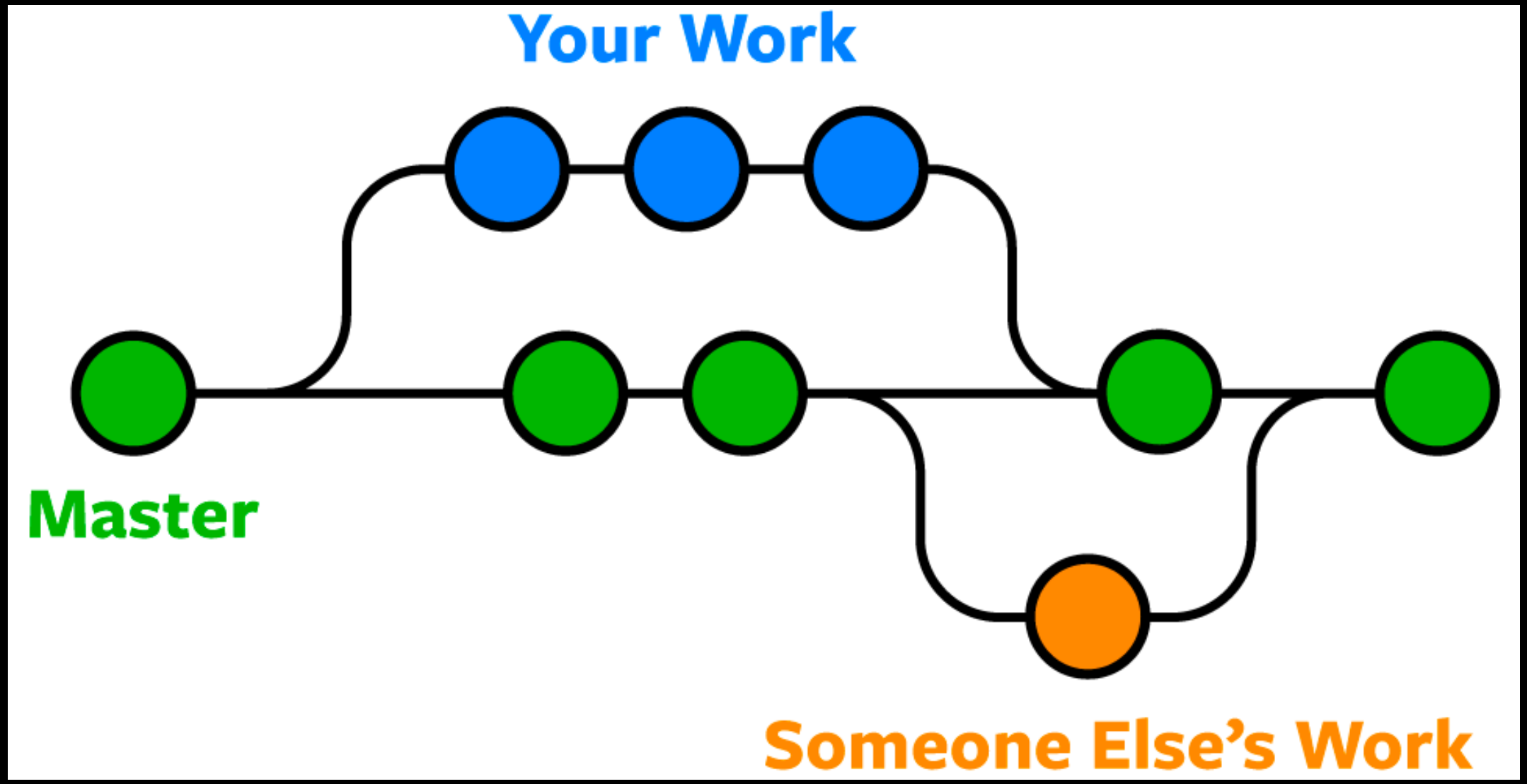


Image: https://www.nobledesktop.com/blog/what-is-git-and-why-should-you-use-it

Git glossary: https://git-scm.com/docs/gitglossary

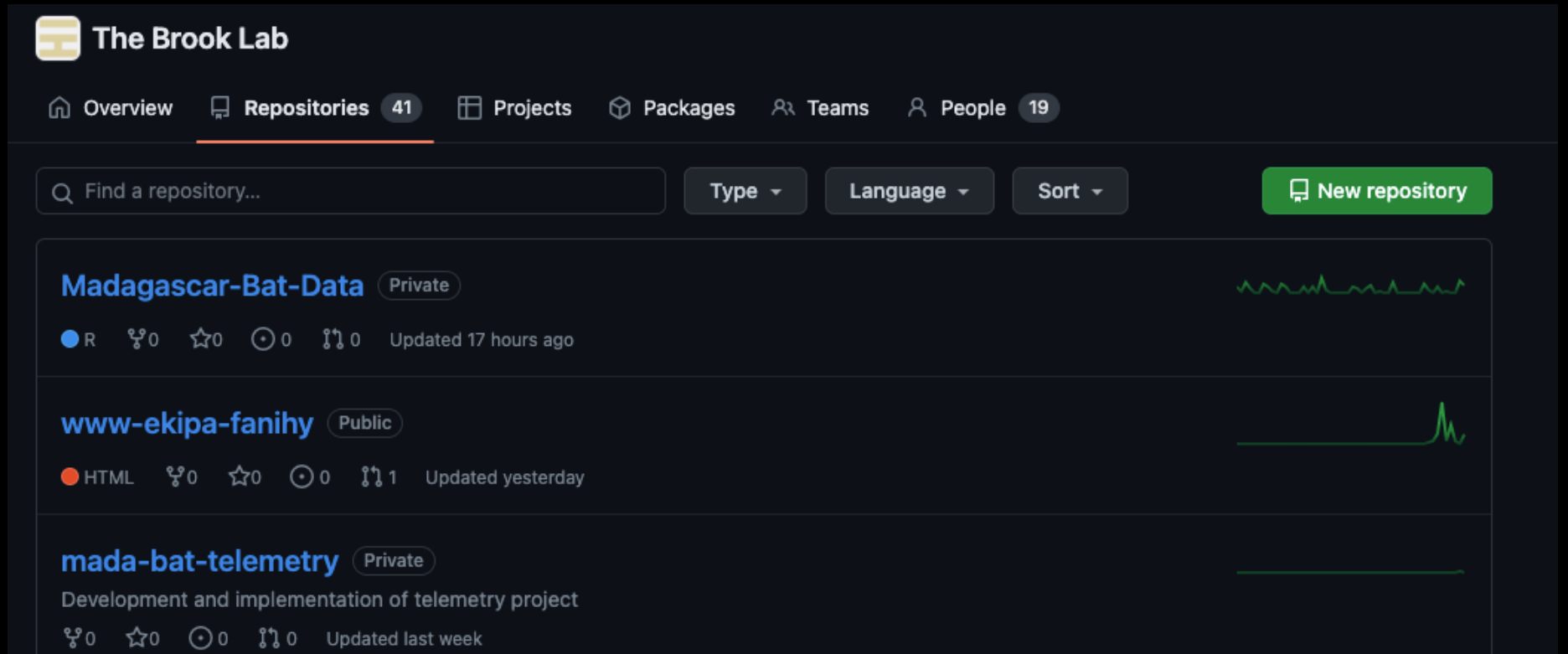
- commit
  - As a noun: A single point in the Git history; the entire history of a project is represented as a set of interrelated commits. The word "commit" is often used by Git in the same places other revision control systems use the words "revision" or "version". Also used as a short hand for commit object.
  - As a verb: The action of storing a new snapshot of the project's state in the Git history, by creating a new commit representing the current state of the index and advancing HEAD to point at the new commit.
- push
  - Pushing a branch means to get the branch's head ref from a remote repository, find out if it is an ancestor to the branch's local head ref, and in that case, putting all objects, which are reachable from the local head ref, and which are missing from the remote repository, into the remote object database, and updating the remote head ref. If the remote head is not an ancestor to the local head, the push fails.
- pull
  - Pulling a branch means to fetch it and merge it.
- pull request
  - An event in Git where a contributor asks a maintainer of a Git repository to review code they want to merge into a project.

Git glossary: https://git-scm.com/docs/gitglossary

- branch
  - A "branch" is a line of development. The most recent commit on a branch is referred to as the tip of that branch. The tip of the branch is referenced by a branch head, which moves forward as additional development is done on the branch. A single Git repository can track an arbitrary number of branches, but your working tree is associated with just one of them (the "current" or "checked out" branch), and HEAD points to that branch.
- master
  - The default development branch. Whenever you create a Git repository, a branch named "master" is created, and becomes the active branch. In most cases, this contains the local development, though that is purely by convention and is not required.
- fork
  - Another way of saying clone or copy. The term fork (in programming) derives from a Unix system call that creates a copy of an existing process. So, unlike a branch, a fork is independent from the original repository. If the original repository is deleted, the fork remains. If you fork a repository, you get that repository and all of its branches.
- origin
  - The default upstream repository. Most projects have at least one upstream project which they track. By default origin is used for that purpose. New upstream updates will be fetched into remote-tracking branches named origin/name-of-upstream-branch, which you can see using git branch -r.

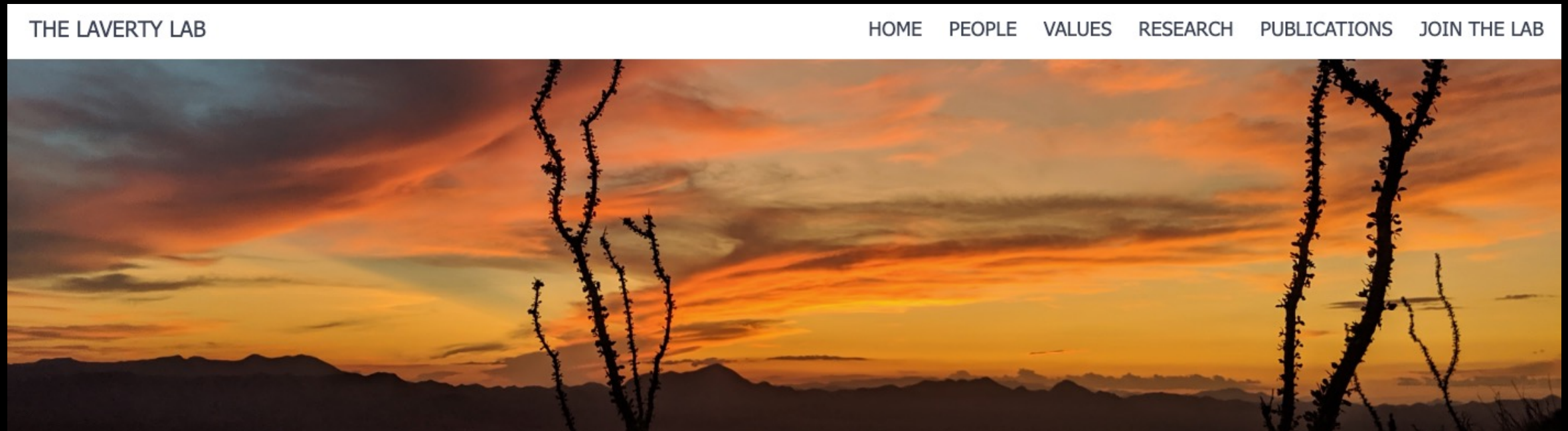
Theresa's experience with GitHub

- Working collaboratively and independently on code largely within repositories hosted on an organization's GitHub page
  - Brook Lab and Laverty Lab



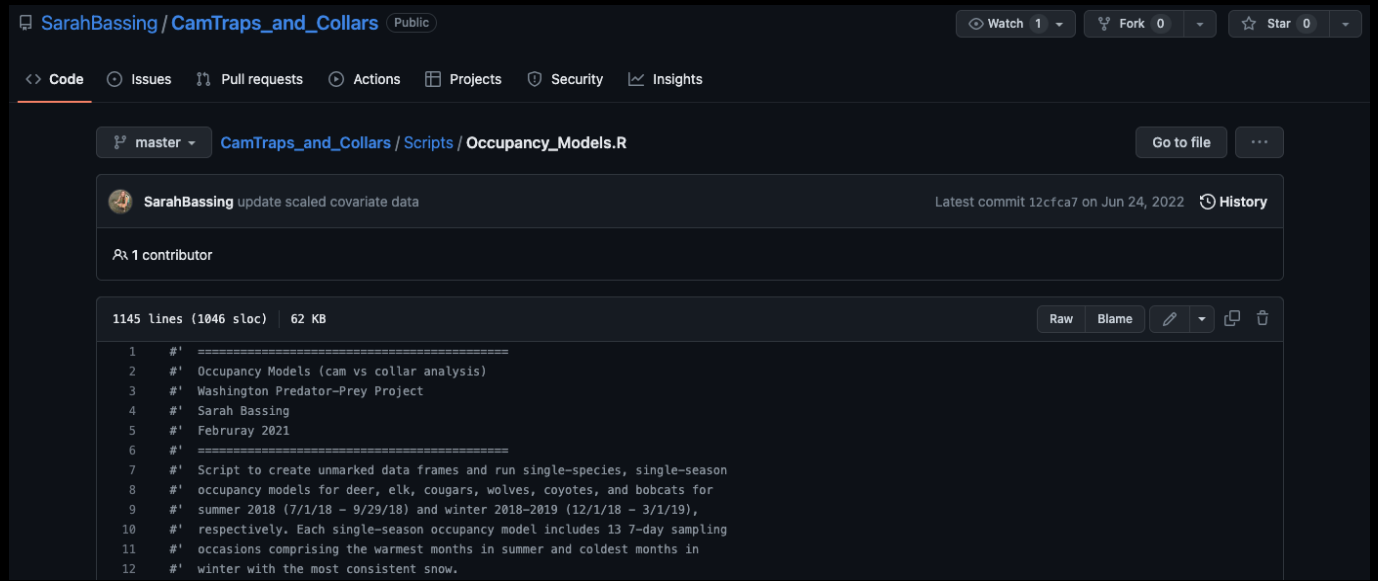
Theresa's experience with GitHub

- Working collaboratively and independently on code largely within repositories hosted on an organization's GitHub page
  - Brook Lab and Laverty Lab
- Hosting our lab's website: https://github.com/LavertyLab/lab-website



Theresa's experience with GitHub

- Working collaboratively and independently on code largely within repositories hosted on an organization's GitHub page
  - Brook Lab and Laverty Lab
- Hosting our lab's website
- Creeping on other people's code and lecture materials



Erica's experience with GitHub

- Code storage
  - Easy to send a link to collaborators
  - Easy to move between work and home computers

Erica's experience with GitHub

- Code storage
  - Easy to send a link to collaborators
  - Easy to move between work and home computers
- Using GitHub to house collective lab data sets
  - Prevents "accidents" when adding to a dataset
  - Any changes to dataset have to be looked at and approved by another person

Erica's experience with GitHub

- Code storage
  - Easy to send a link to collaborators
  - Easy to move between work and home computers
- Using GitHub to house collective lab data sets
  - Prevents "accidents" when adding to a dataset
  - Any changes to dataset have to be looked at and approved by another person
- Versioning code and archiving on Zenodo to get a DOI to cite in papers

Interactive Tasks

1. Create a new repository and clone to your computer
2. Make changes, commit changes, and push to GitHub
3. Review/change .gitignore
4. Revert changes https://intro2r.com/use_git.html
5. Fork an existing repository and clone to your computer
6. Make a pull request
7. Resolve a merge conflict
8. Add a collaborator to a repository/organization
9. Publish repository: DOI link generation https://coderefinery.github.io/github-without-command-line/doi/

Some best practices

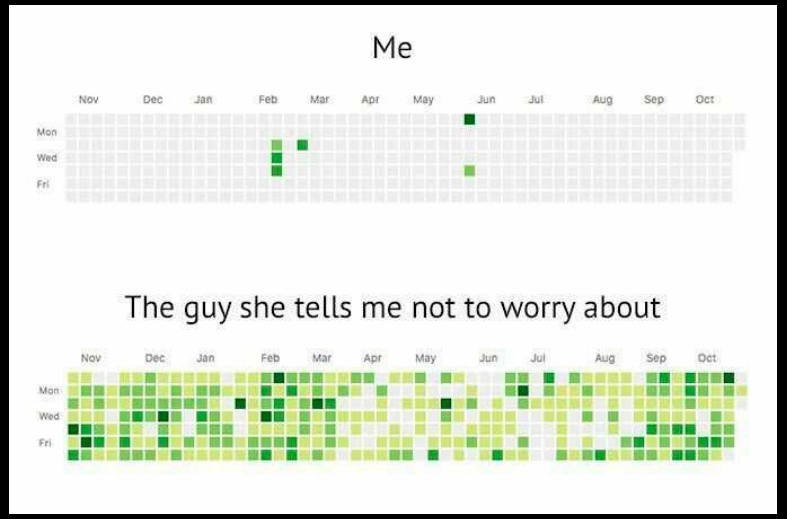
- ALWAYS pull before you push
- Designate one member to maintain the primary repository
- All other members ***fork*** the primary repository and ***clone*** it to their local machine/home directory
- Any changes made by other members can be submitted as pull requests to the primary repository owner
- Primary owner can decide whether to accept the changes or not
- This way, multiple copies of your code will be floating around in case one member does something stupid and deletes their entire repository

Guidance: https://hwheeler01.github.io/CompBio/github/

Git going!


What naming convention you use for git branches?
What's a branch? I only work in master.
imgflip.com


git push --force
origin master


Me
Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct
Mon
Wed
Fri
The guy she tells me not to worry about
Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct
Mon
Wed
Fri


I'M A SNEAKY FOX
SNEAKED IN YOUR HOUSE
SAT ON YOUR COUCH
ATE YOUR PASTA
PUSHED TO GOOGLE MASTER BRANCH
©Li Chen
cxocomics.com
ProgrammerHumor.io

Contact us

tlaverty@nmsu.edu; echriste@nmsu.edu

Helpful references:

- https://happygitwithr.com
- https://swcarpentry.github.io/git-novice/
- https://docs.github.com/en