Git for Scientific Research and Collaboration

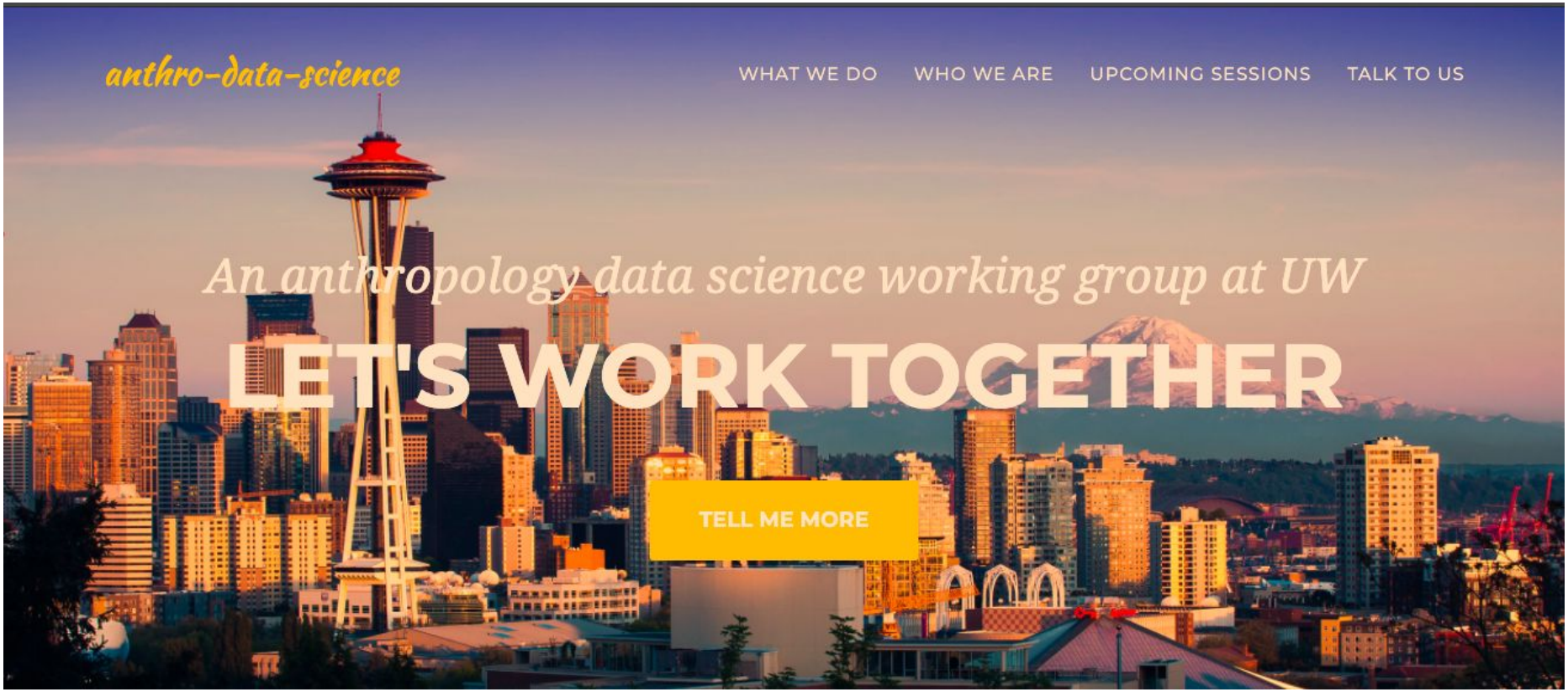
Anthro Data Science group
Department of Anthropology
University of Washington
August 2020

Workshop Schedule

times are approximate

Start time	End time	Topic
09:00	09:15	Introduction to data science group by Delaney Glass; Motivation for using Git and Github by Ben Marwick
09:15	09:45	Define key tools & concepts such as Git and GitHub, remote, local, commit, push, pull by Liying Wang Activity 1: log into your GitHub account, follow some people, and make a repository
09:45	10:20	Define concepts of fork, clone by Gayoung Park Activity 2 (work in pairs): learn to fork, commit, and pull request on GitHub. Add a new file, add text to that file using RStudio. Look at commit history and blame view on GitHub
10:20	10:55	Collaborating and resolving merge conflicts by Anwesha Pan Activity 3 (work in pairs): learn to collaborate with Git & RStudio. Do the full cycle of fork, clone, commit, identify and resolve merge conflicts

Who We Are



https://anthro-data-science.github.io/

Motivation for using Git and GitHub

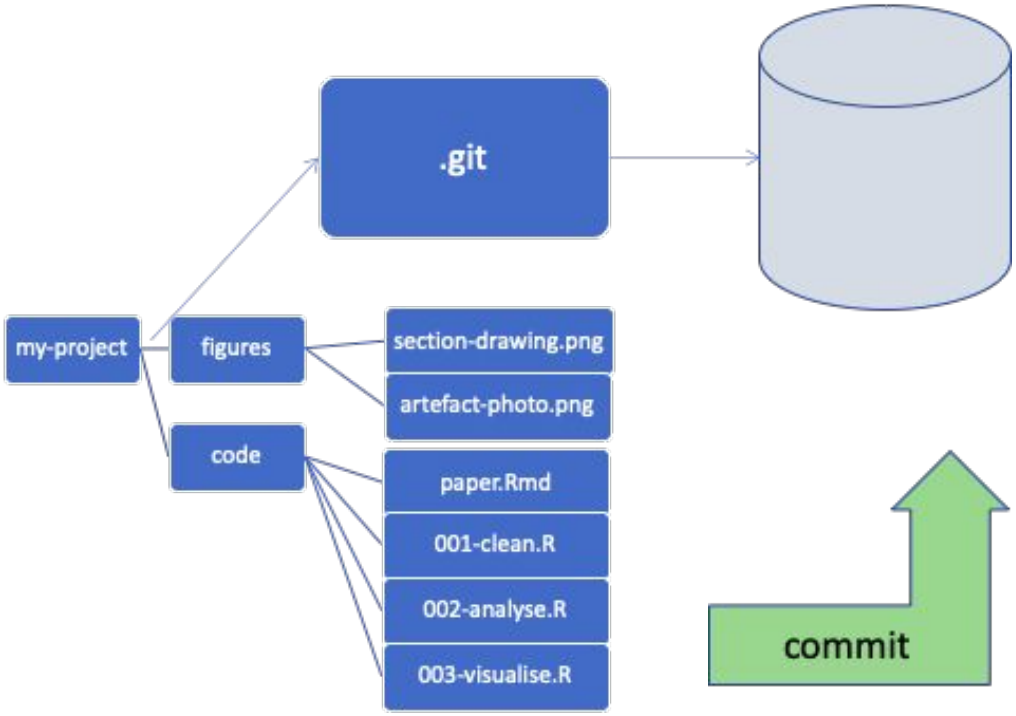
git GitHub

The version control software on our computer

Git can be used from the terminal. Here are the most commonly used commands

command	description
`git clone` ***`url [dir]`***	download a git repository to your computer
`git add` ***`files`***	adds file contents to the staging area
`git commit`	records a snapshot of the staging area
`git push`	upload your data or changes in your files to a remote repository
`git pull`	fetch from a remote repo and try to merge into the current branch
others: `status, help, init, reset, branch, checkout, merge, log, tag`	

A Commit



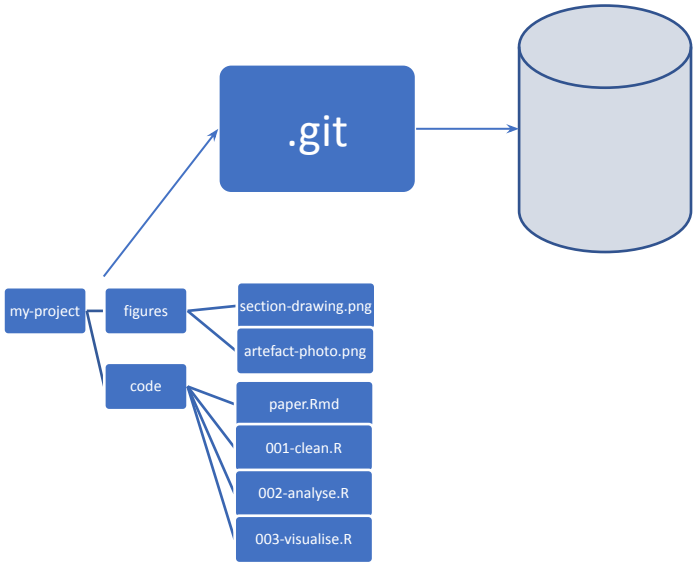
When you do a “commit”, you record all your local changes into the Git database.

The database is “append-only”. Nothing is ever over-written there, so everything you ever commit can be recovered.

Uploading to a remote server

your local computer

a remote server, somewhere on the internet, eg. github.com

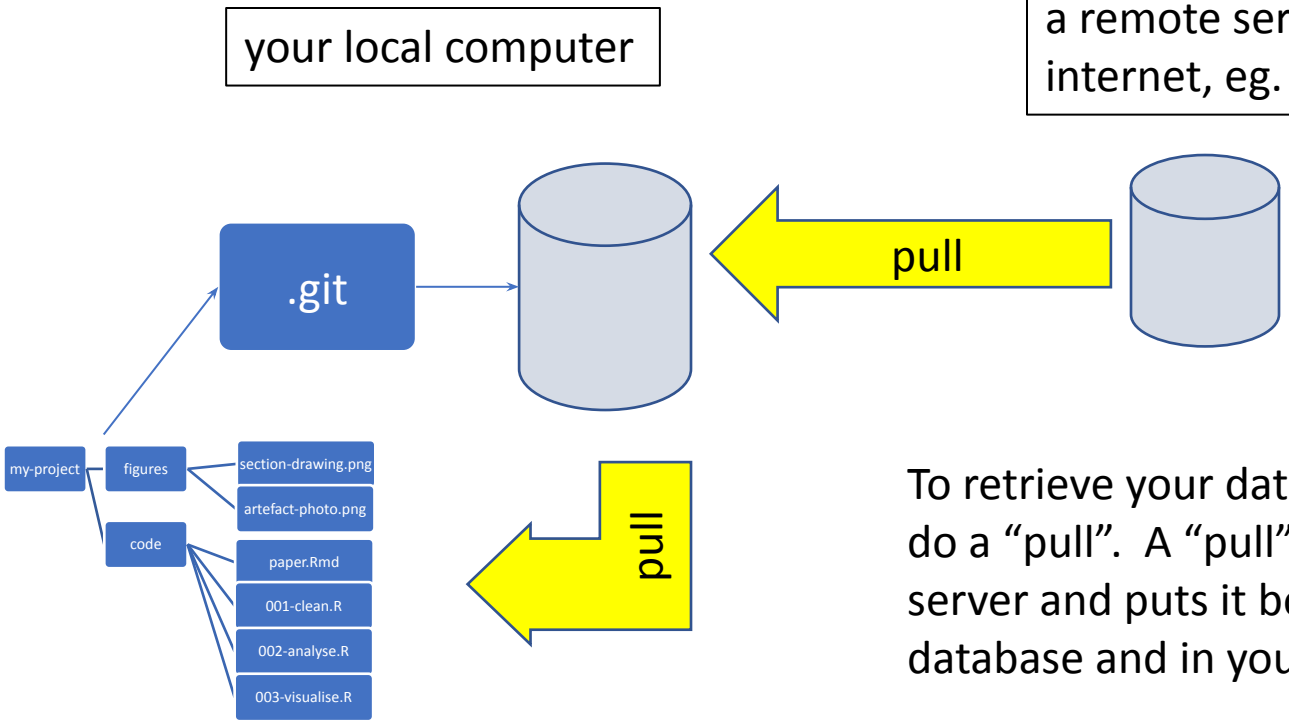


At the end of each work session, you need to save your changes on the server. This is called a "push".

Now all your data is backed up.

- You can retrieve it, on your machine or some other machine.
- We can retrieve it

Downloading from a remote server



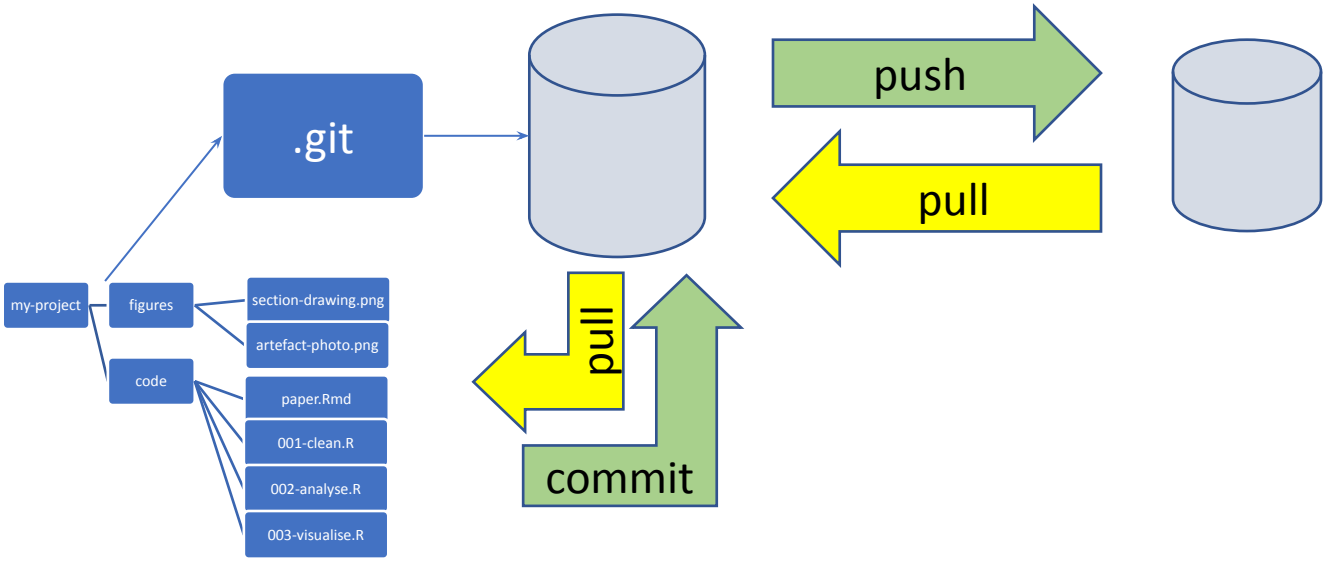
To retrieve your data from the server, you do a “pull”. A “pull” takes the data from the server and puts it both in your local database and in your ordinary files.

If your local file has changed, git will merge the changes if possible. If it can’t figure out how to merge it, you will get an error message. We'll learn how to deal with these later.

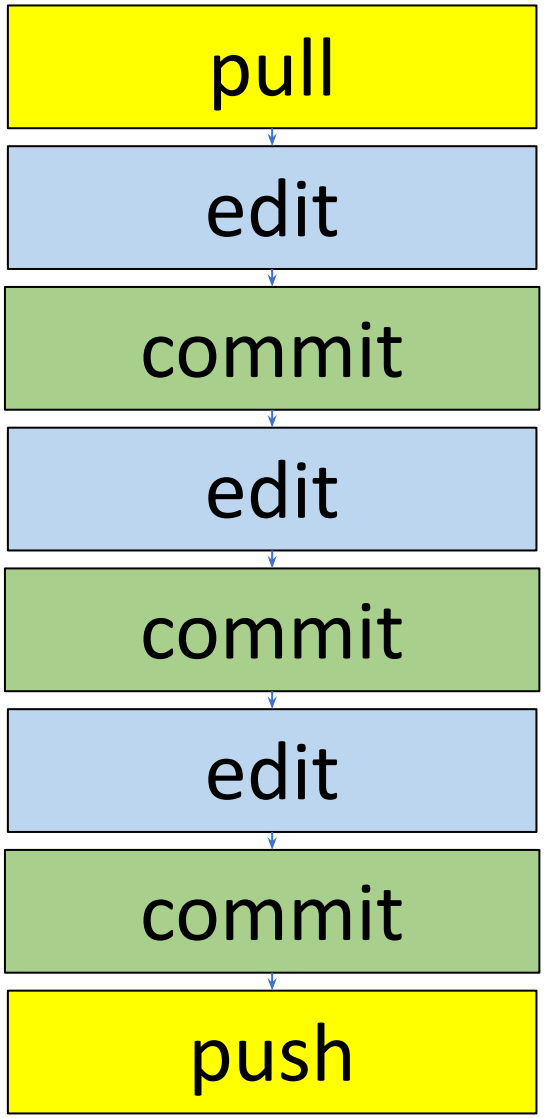
The whole picture: commit, push pull

your local machine

a remote server, somewhere on the internet, eg. github.com



Our typical workflow



Best practice: commit your work whenever you've gotten one part of your problem working, or before trying something that might fail.

If your new stuff is screwed up, you can always "revert" to your last good commit. (Remember: always "revert", never "roll back")

git

GitHub

The version control software on our computer

A company that provides a web service for people using Git

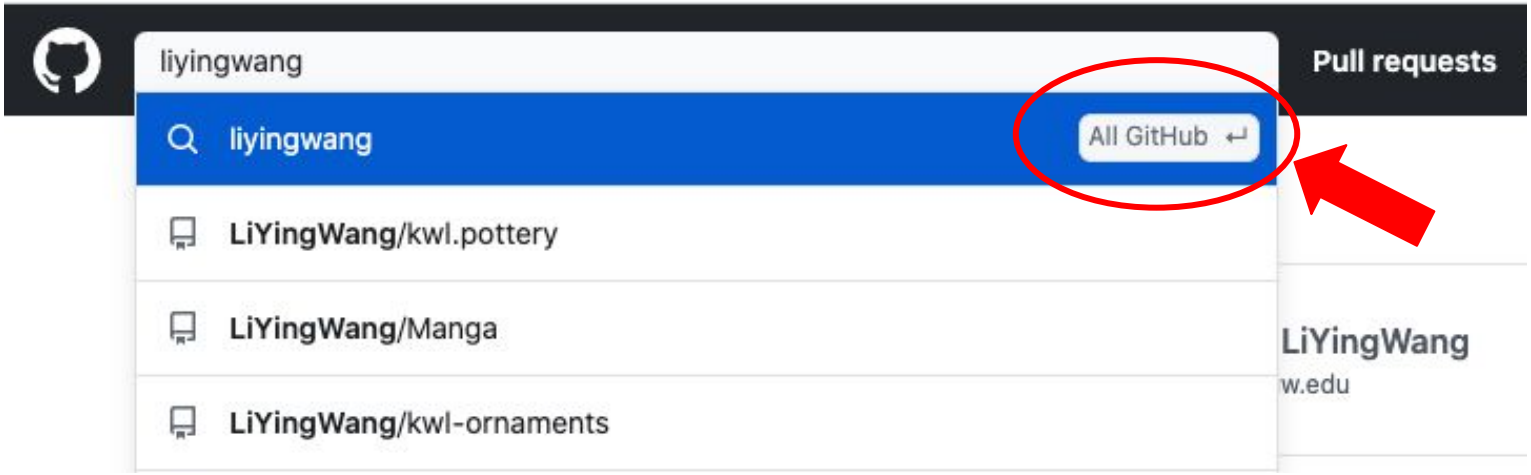
Let’s take a quick look at GitHub

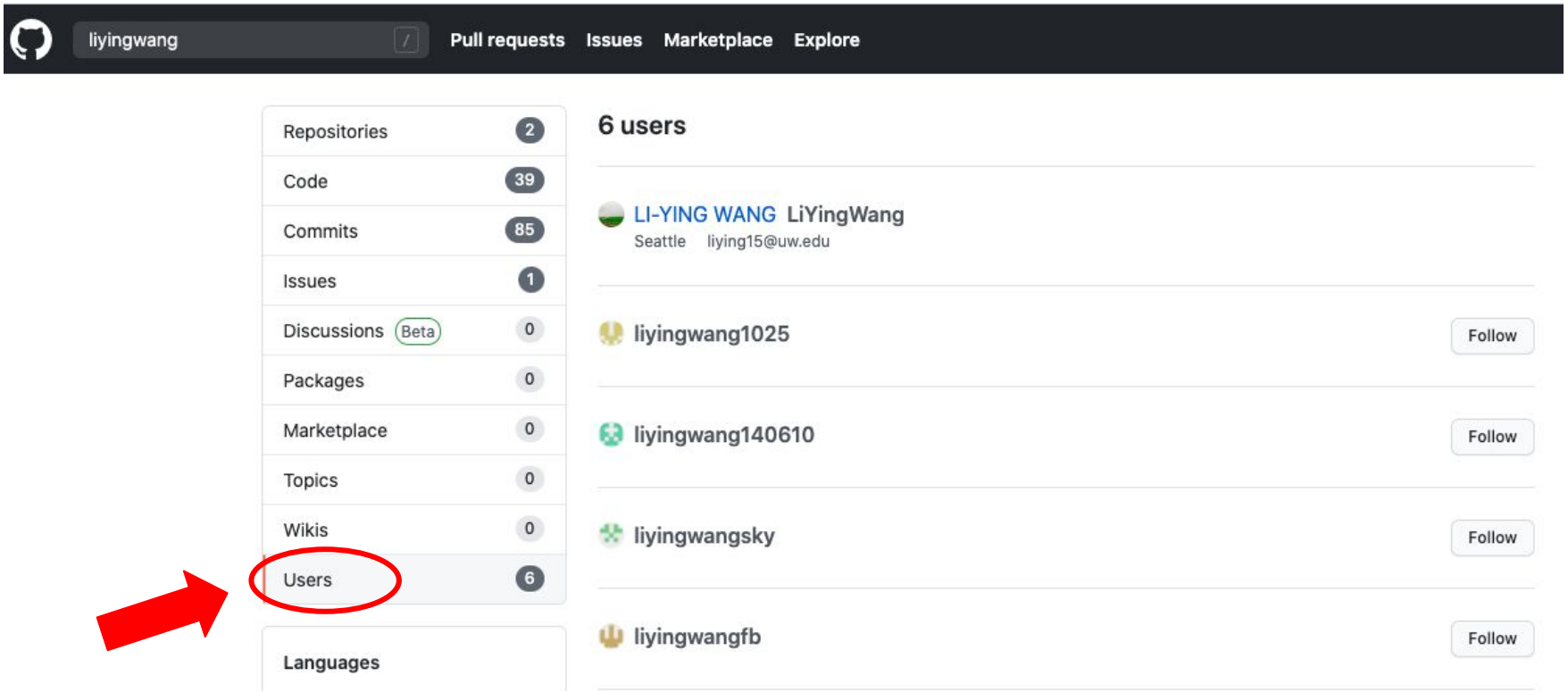
Activity(10 min):

-log into your GitHub account

-follow some people (share GitHub usernames in the zoom chat)

-create a repository


liyingwang
Pull requests
liyingwang
All GitHub
LiYingWang/kwl.pottery
LiYingWang/Manga
LiYingWang/kwl-ornaments
LiYingWang
w.edu


liyingwang
Pull requests
Issues
Marketplace
Explore
Repositories 2
Code 39
Commits 85
Issues 1
Discussions Beta 0
Packages 0
Marketplace 0
Topics 0
Wikis 0
Users 6
Languages
6 users
LI-YING WANG LiYingWang
Seattle liying15@uw.edu
liyingwang1025
Follow
liyingwang140610
Follow
liyingwangsky
Follow
liyingwangfb
Follow

Contact us if you have any questions!

Delaney Glass (dglass1@uw.edu)

Liying Wang (liying15@uw.edu)

Gayoung Park (gayoungp@uw.edu)

Anwesha Pan (anweshap@uw.edu)

Ben Marwick (bmarwick@uw.edu)

Thank you!