

Hosted by quantitative Cell & Molecular Biology, Colorado State University

Open Source Software

Development with  **GitHub**

Open Source Software Development with GitHub

Outline

- . Open source software and Github
- . User Testimonials
- . How to – user interface (Kristin)
- . How to – commandline (Reed)

Open Source Software

Source code: the technical blueprint that tells the computer program how to function.

```
<!DOCTYPE html>
<html id="home-layout">
  <head>
    <meta http-equiv="content-type" content="text/html; charset=utf-8">
    <title>Source Code Pro</title>
    <!-- made with <3 and AFDKO -->
    <meta name="keywords" content="sans, monospace, open source, coding, for">
    <link rel="stylesheet" type="text/css" href="https://sourcecodepro.com/css/main.css">
  </head>
  <body>
    <div id="main">
```

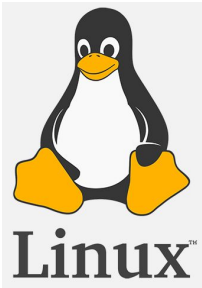


```
0011111100000011111001001100011011100101001111111111101
1110111111000011101011111101100011110111111010101100110
1101111100001110001101010010111010011111000110001000110
11000111011110111111011111111010001111110111011111110
111110011011101111000111101110101110100011111110111110
110011011111011000100100000010111011101110111111111111
1111100011111100011111111111110011110011111100101111111
01111110110011101111011111110101111101111111101100111010
1111111110001111111001010010100011111011110110100111101
0001111111111011000101000011110010000000111100100100011
1011111001111111101010111111000101110111000001001111110
0111000011110111011111001111110011111111001100000110111
1011101101000010011001100011101110000110010000001100111
1101100110001010111101101000111110100011010111110010111
111111110011000111100111101010000110100111000100100010
1110100001011100111111100000111110111111011110101111100
111001001101011111111110011111111111111111111001010111111
0000000001111111100001100111101100100011011011010110001
01000100111111110011111111111111111001111100011111001110101
0100011111001001111111100001011110111001101111110110011111
```

Open Source Software

Open Source: source code that is accessible to inspect, modify, copy and use.

```
<!DOCTYPE html>
<html id="home-layout">
  <head>
    <meta http-equiv="content-type" content="text/html; charset=utf-8">
    <title>Source Code Pro</title>
    <!-- made with <3 and AFDKO -->
    <meta name="keywords" content="sans, monospace, open source, coding, for">
    <link rel="stylesheet" type="text/css" href="https://sourcecodepro.com/css/main.css">
  </head>
  <body>
    <div id="main">
```



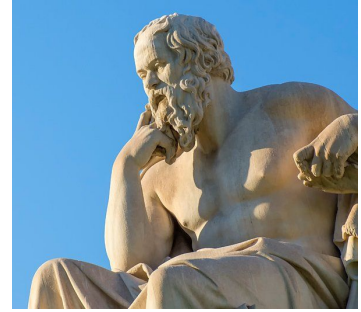
Closed Source: proprietary source code that can not be accessed or change by the public.

[illegible]

iTunes

Open Source Software

Its a Philosophy!



- Proponents believe people should have the right to use their software in whatever way they see fit without licensing restrictions.
- Allows software to evolve through open contributions from many users.
- Critics worry if open source leads to difficulty in managing programs and wonder if software can be reliably developed without profit intensive.

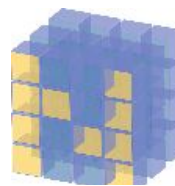
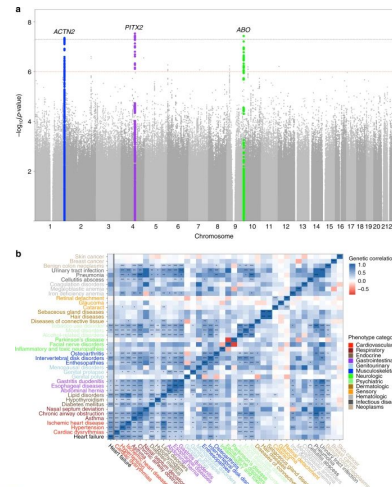
Open Source Software

FOR YOU!

- Transparency & reproducibility in data analysis
- Collaborative development of and easy access to community tools
- Coding Portfolio

SCIENTIFIC
REPORTS

nature



NumPy



How do you develop open source software?

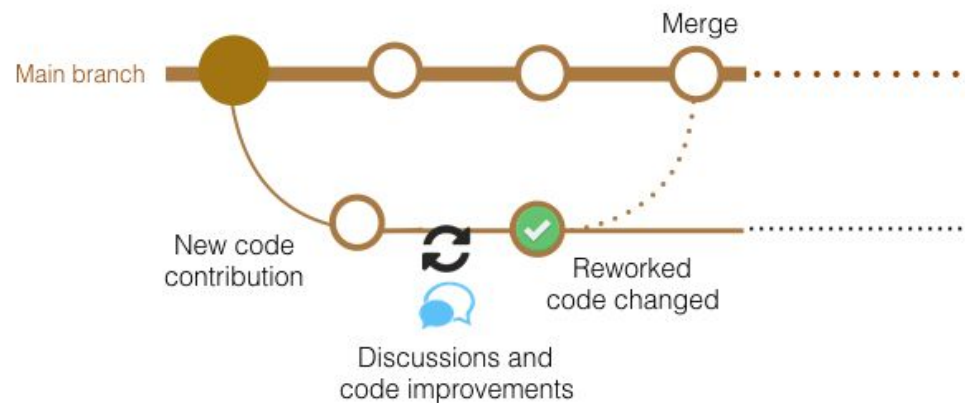
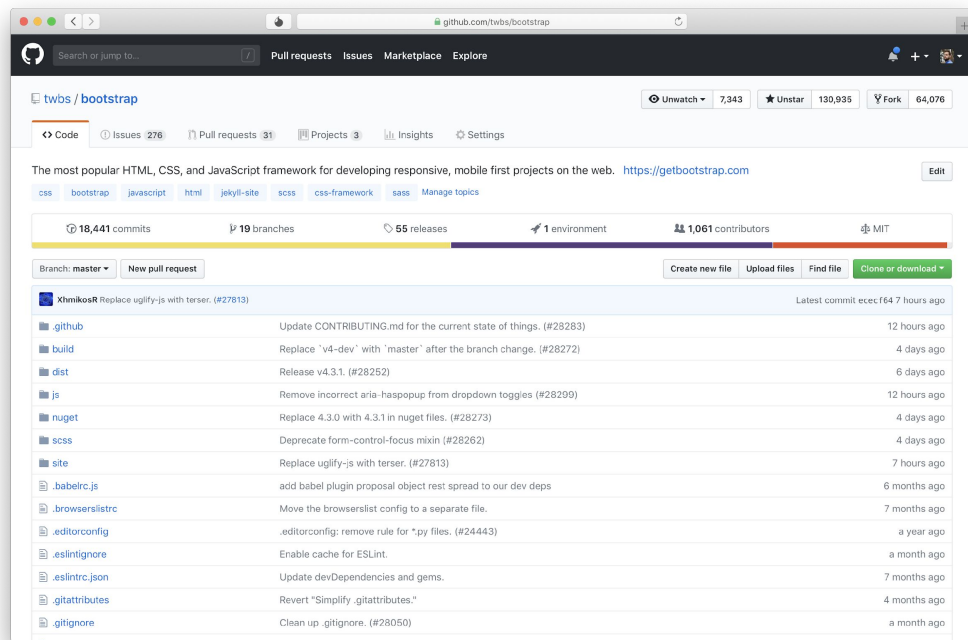
1) Write some code



2) Publish on Internet



Developing Open Source Software with GitHub



Developing Open Source Software with GitHub

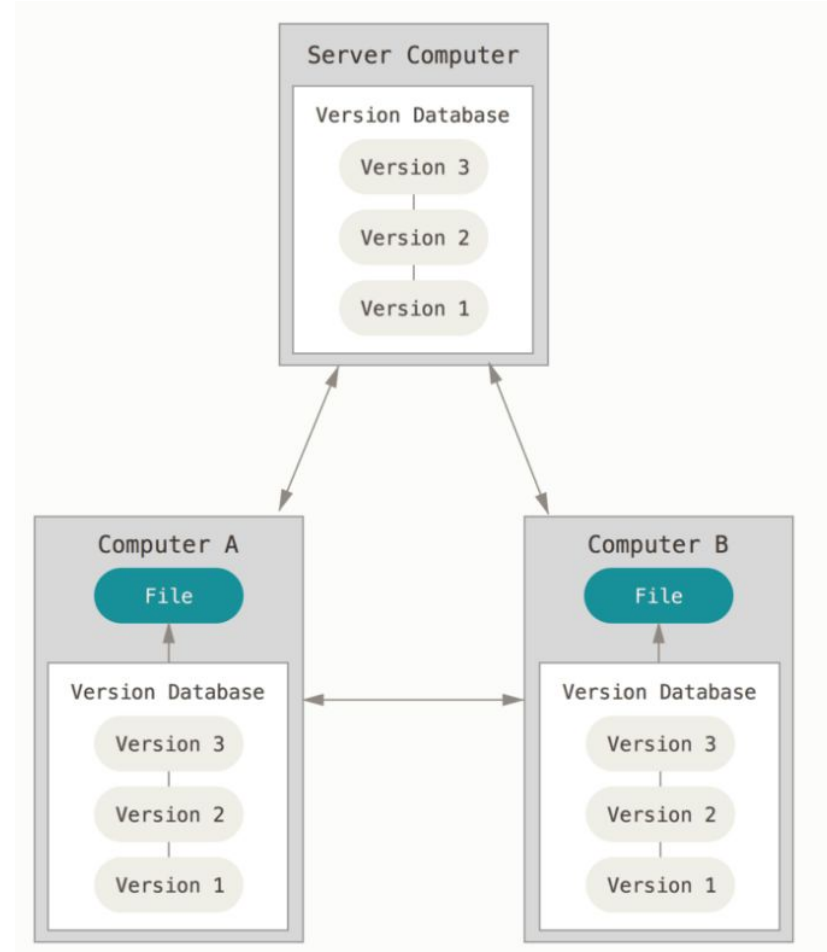


Resources:

- .GitHub's User manual
- .Community support
- .YouTube
- .Google!

Version Control

Version control is a system that records changes to a file or set of files over time so that you can recall specific versions later.



Version Control



20 BEST VERSION CONTROL SYSTEMS OF 2020 - 1/2

01

AWS CodeCommit

It is a fully-managed source control service that empowers your team to collaborate on code. It offers a secure storage solution for your binary files, documents, and source code in the cloud.

02

Team Foundation Server

It is an integrated server suite to share code, track and monitor workflows, and ship software. Its version control solution offers unlimited private repositories for storage and collaboration on the code.

03

GitHub

It is a software development platform built explicitly for developers. It empowers you to manage projects, build software, and host and review code.

04

JEDI VCS

It is an open-source VCS platform. It offers features such as project and line history, time tracker, project-based rights management, an integrated ToDo Manager, and customizable DiffTool.

Protecting Your Open Source Projects

Choose an open source license

An open source license protects contributors and users. Businesses and savvy developers won't touch a project without this protection.

Which of the following best describes your situation?



I need to work in a community.

Use the [license preferred by the community](#) you're contributing to or depending on. Your project will fit right in.

If you have a dependency that doesn't have a license, ask its maintainers to [add a license](#).



I want it simple and permissive.

The [MIT License](#) is short and to the point. It lets people do almost anything they want with your project, like making and distributing closed source versions.

[Babel](#), [.NET Core](#), and [Rails](#) use the MIT License.



I care about sharing improvements.

The [GNU GPLv3](#) also lets people do almost anything they want with your project, *except* distributing closed source versions.

[Ansible](#), [Bash](#), and [GIMP](#) use the GNU GPLv3.

choosealicense.com

Protecting Your Open Source Projects

choosealicense.com

What if none of these work for me?

**My project isn't
software.**

[There are licenses for that.](#)

**I want more
choices.**

[More licenses are available.](#)

**I don't want to
choose a license.**

[Here's what happens if you don't.](#)

The content of this site is licensed under the Creative Commons Attribution 3.0 Unported License.


[About](#)

[Terms of Service](#)

[Help improve this page](#)

Curated with ❤️ by GitHub, Inc. and You!

Protecting Your Open Source Projects

 BioRRW Add files via upload ...	bd204c1 on Oct 19	🕒 182 commits
📁 containers	Update R2R-0.0.1-Fasta-QC-Ann-Only.nf	3 months ago
📁 docs	Update usage.md	last month
📄 .gitignore	Update .gitignore	7 months ago
📄 LICENSE	Create LICENSE	9 months ago
📄 R2R-0.0.1.nf	Add files via upload	last month
📄 README.md	Update README.md	3 months ago
📄 nextflow.config	Update nextflow.config	3 months ago

Hybrid and non-hybrid whole genome assembler with resistome identification for Illumina, PacBio and Nanopore reads.

📖 [Readme](#)

📜 [MIT License](#)

Releases

🏷️ 1 tags

[Create a new release](#)

Permissions

- Commercial use
- Distribution
- Modification
- Private use

Conditions

- License and copyright notice

Limitations

- Liability
- Warranty

~~Why would anyone want to use my code?!~~

How does my open source code help me?



Bioinformatics Engineer

APPLY FOR THIS JOB

REMOTE / COMPUTATIONAL SCIENCE / FULL-TIME

Qualifications

- MS or PhD in Bioinformatics, Biology or Computer Science, or a related field
- Minimum 2 years experience in building and maintaining production scale bioinformatics pipelines
- Fluency with Python
- Experience with AWS
- Experience with metagenomics
- Proven record of bioinformatics expertise, via a vis publications, publicly-available computational tools packages, or github repos that show thoughtful genomics analysis

**Job posting from Trace
Genomics on 11/30/2020**

~~Why would anyone want to use my code?!~~ How does my open source code help me?




BIOINFORMATICS ANALYST

- **Documentation and Reproducibility:** Writes and maintain tools/scripts with documentation, concept of singularity, version control and reproducibility. **Familiarity with GitHub is preferred.** Maintain transparent open access documentation of project goals, analytical approaches and milestones for each project.

Job posting from
Cincinnati Children's
Hospital on 11/30/2020

Who else uses GitHub?



Netflix, Inc.
Netflix Open Source Platform
Los Gatos, California <http://netflix.github.io/> netflixoss@netflix.com

Repositories 187

Packages

People 22

Projects

Type: All

Language: All

vmaf

Perceptual video quality assessment based on multi-method fusion.

Python


458

2,007

49 (11 issues need help)

6

Updated 2 minutes ago



titus-control-plane

Titus is the Netflix Container Management Platform that manages containers and provides integrations to the infrastructure ecosystem.

containers

titus

Java

Apache-2.0


64

271

2

2

Updated 40 minutes ago



genie

Distributed Big Data Orchestration Service

java

distributed-systems

cloud

microservices

big-data

spring-boot

microservice

Java

Apache-2.0


324

1,413

2

1

Updated 1 hour ago



conductor

Conductor is a microservices orchestration engine -

<https://netflix.github.io/conductor/>

Java

Apache-2.0


962

3,103

208 (27 issues need help)

31

Updated 2 hours ago



Top languages

Java

JavaScript

Python

Go

Ruby

Most used topics

security

python


aws

machine-learning

monitoring

People

22 >



Report abuse

Who else uses GitHub?



Amazon

<https://amazon.com>

Verified



Repositories 117



Packages



People 16



Projects

Pinned repositories



ion-java

Java streaming parser/serializer for Ion.



Java

☆ 780

🔗 76



style-dictionary

A build system for creating cross-platform styles.



JavaScript

☆ 1.4k

🔗 203



computer-vision-basics-in-microsoft-excel

Computer Vision Basics in Microsoft Excel (using just formulas)



2.2k

🔗 212



sketch-constructor

Read/write/manipulate Sketch files in Node without Sketch plugins!



JavaScript

☆ 454

🔗 28



smoke-http

Specialised HTTP Client for service operations abstracted from the HTTP protocol.



Swift

☆ 48

🔗 8



milan

Milan is a Scala API and runtime infrastructure for building data-oriented systems, built on top of Apache Flink.




Scala

☆ 24

🔗 8

How else can I utilize GitHub?



Monica Powell
M0nica

Building tech to elevate people.
Founder of React Ladies a community for React JS developers.

[Follow](#) [Sponsor](#) [...](#)


1.6k followers · 27 following · 312

New York, New York
github@aboutmonica.com
https://www.aboutmonica.com
@waterproofheart

Highlights

- Arctic Code Vault Contributor
- GitHub Star
- PRO

Organizations

-  **React-Ladies**
We're a group of women and non-binary ReactJS enthusiasts in New York City (and beyond).


Overview Repositories 158 Projects Packages

M0nica / README.md

Hi, I'm Monica 🖐️ 👤

MONICA POWELL

software engineer, content creator & community organizer




I'm a software engineer who is passionate about making open-source more accessible, creating technology to elevate people, and building community. Some technologies I enjoy working with include ReactJS, Jamstack (JavaScript, APIs + Markup) and GraphQL. I recently was selected to be an inaugural [GitHub Star](#) 🌟 based on my involvement in the tech community. My interest in the React ecosystem led me to launch [React Ladies](#), a community for women and non-binary ReactJS developers.


Find me around the web 🌐

- Learning in public on [Twitch](#) or [monica.dev](#) 📺 📄
- Tinkering with interactions on [Codepen](#) 📌
- Sharing updates on [LinkedIn](#) 💼

Pinned

 **React-Ladies**

We're a group of women and non-binary ReactJS enthusiasts in New York City (and beyond).

 **affirmation_generator**

A random affirmation generator written with ReactJS :-)

How else can I utilize GitHub?

Lucia (Lucy) Williams

Research

Teaching

Outreach

CV



Lucia Williams

Computer science graduate student studying algorithms, both theory and application.

📍 Bozeman, Montana

✉ Email

🐙 Github

I am a third-year PhD student in computer science at Montana State University in Bozeman, Montana. My advisor is [Brendan Mumey](#). I am interested in the design and analysis of algorithms and data structures for problems in computational biology. I have worked on flow network based methods for RNA sequencing and on methods for finding conserved genomic regions in the presence of missing data or without a single reference genome. I have also worked with the [Computational Topology and Geometry](#) group at MSU to develop the theory behind a new method for comparing shapes without manual landmarking. See my [research page](#) for more detail on my research pursuits.

Before starting my PhD, I worked as a data scientist at an online advertising company in Seattle, after earning undergraduate degrees in applied math and urban planning. Outside of work, my main hobby is playing and coaching Ultimate Frisbee. I also enjoy human-powered outdoor adventures.

Recap so far

- Open Source Software
 - Allows software to evolve through open contributions from many users
- Version Control
 - a system that records changes to a file or set of files over time
- Protecting your work with licences
- Ability to centralize and showcase your work and collaborative projects

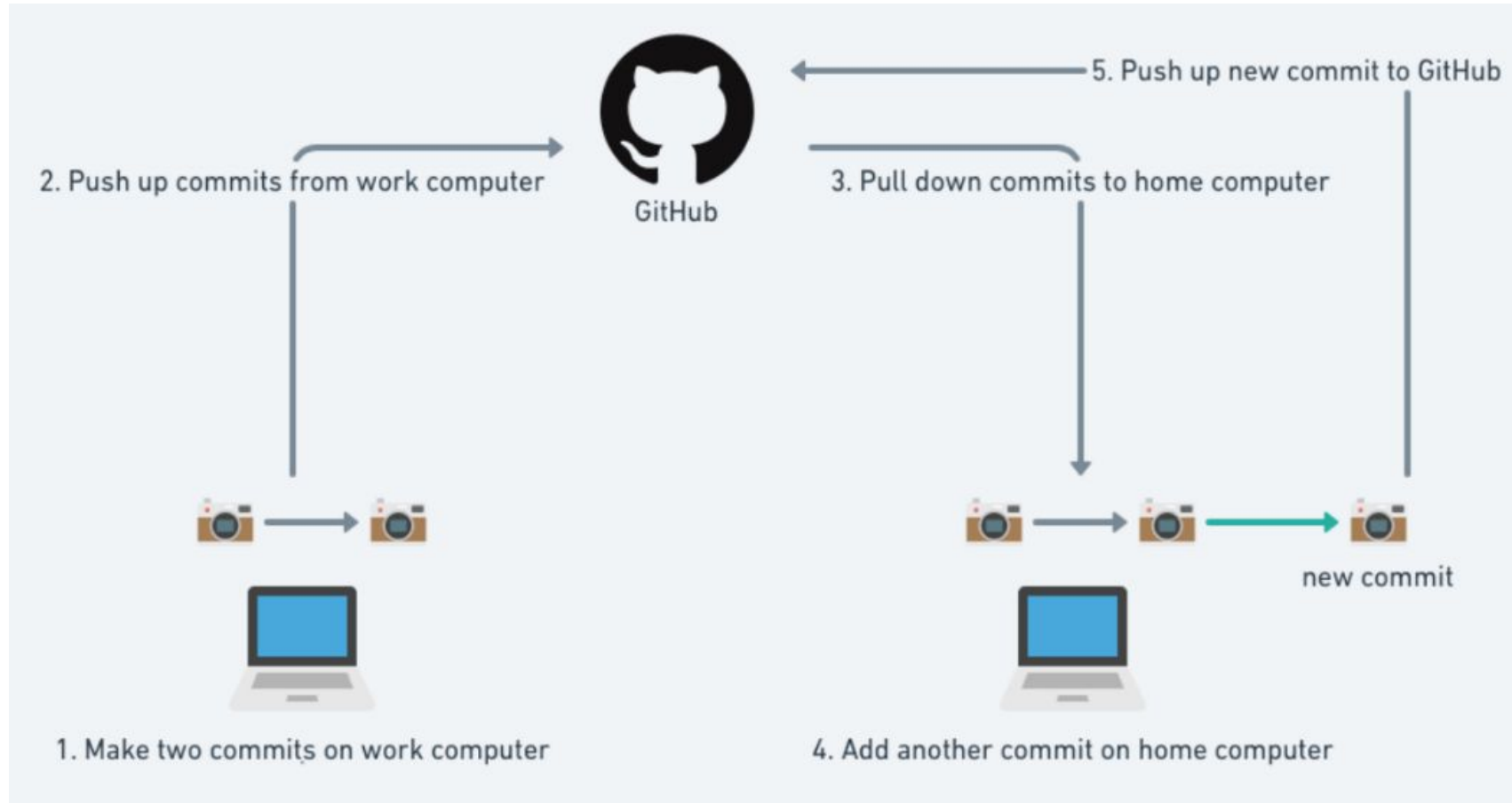
How is GitHub used in the biological sciences?

Examples from CMB Faculty

Erin Osborne Nishimura

Department of Biochemistry and Molecular Biology

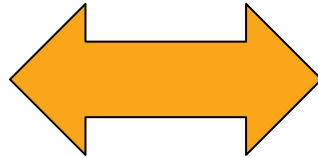
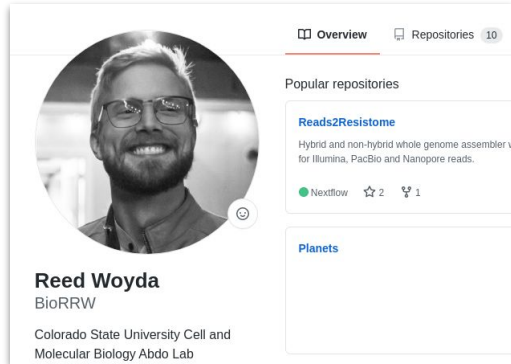
Interacting with GitHub via the command line



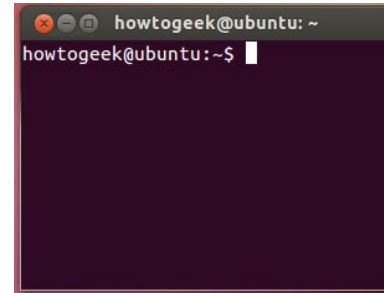
Interacting with GitHub via the command line

What is Git?

- Git is the command line version control system (VCS) software which works on your local computer.
- Git was created by Linus Torvalds in 2005 for development of Linux
- You need Git to use GitHub. You can use Git locally without GitHub.



git



Interacting with GitHub via the command line

Brief introduction on git command line usage:

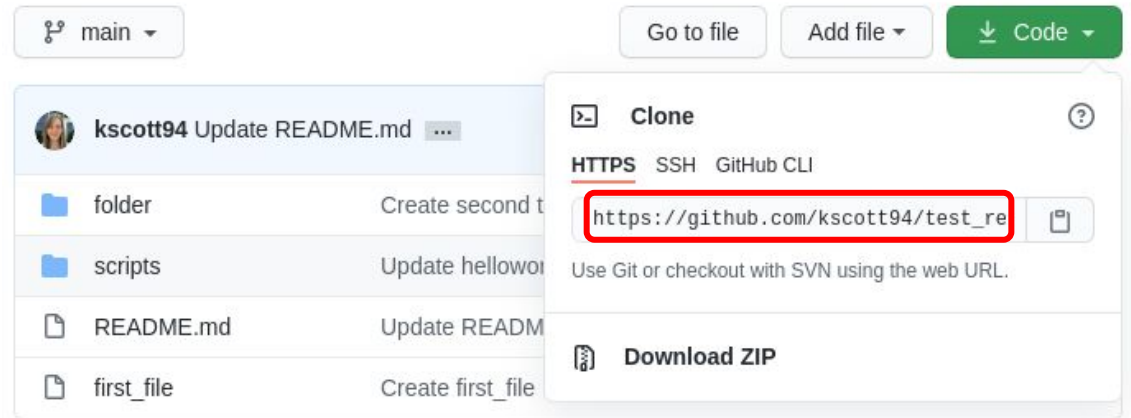
- **Cloning** existing repositories
- **Editing** code from repositories
- **Commit** modified code to existing repositories



Interacting with GitHub via the command line

How to obtain Open Source code from GitHub

1. Open terminal on local computer
2. Navigate to directory of your choosing
3. `git clone https://github.com/kscott94/test_repository.git`



Interacting with GitHub via the command line

Collaboration

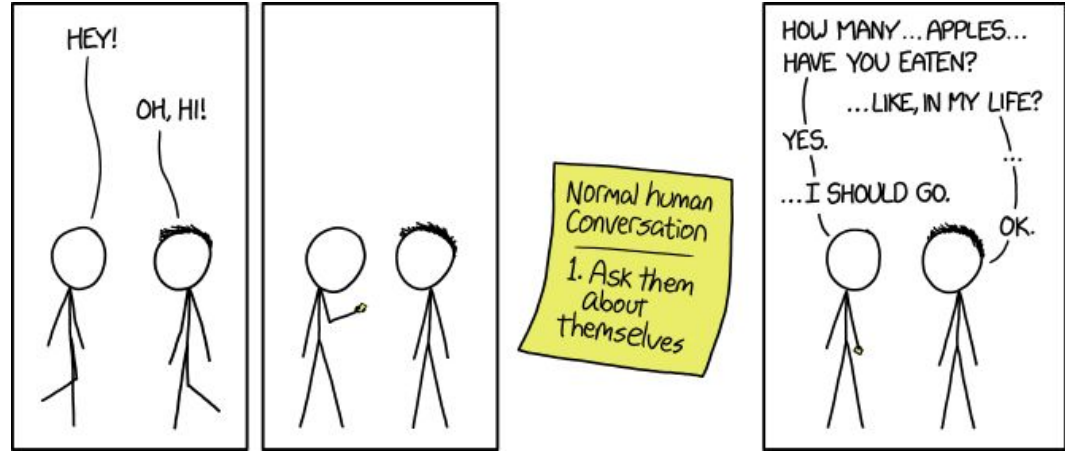
List of command I used to edit a script from a collaborators GitHub repository:

1. In desired directory, tell Git who we are:
 - a. `git config --global user.name "BioRRW"`
 - b. `git config --global user.email reedwoyda@gmail.com`
2. Initialize current directory as a Git repository:
 - a. `git init`
3. Connect to the desired repository:
 - a. `git remote add qcmb_github https://github.com/kscott94/test_repository`
4. Make edits to file of choice (here I edited the helloworld.py script)
5. Prepare the edited file for upload to GitHub:
 - a. `git add helloworld.py`
6. Lastly, **commit** the edited file for final upload to the repository:
 - a. `git commit helloworld.py -m "Updated helloworld.py script"`

Recap: How to use GitHub/Git

- How to create a GitHub repository
 - Add content
 - Add collaborators
- How to download a repository from GitHub (command-line)
- How to 'connect' to an existing repository for collaboration (command-line)
 - Make edits
 - Commit those edits to the GitHub repository
- See edits reflected in GitHub repository

Open Discussion/Questions



Resources

- Basic Git usage
 - <https://git-scm.com/docs/gittutorial>
- Basic GitHub Guide
 - <https://guides.github.com/activities/hello-world/>
 - <https://pages.github.com/>
- GitHub documentation
 - <https://docs.github.com/en>
- Git/GitHub usage (Youtube)
 - https://www.youtube.com/watch?v=SWYqp7iY_Tc&ab_channel=TraversyMedia
- This slide show
 - https://github.com/BioRRW/qCMB_OpenSource_Webinar