Workshop Series

Open Science Skills in R

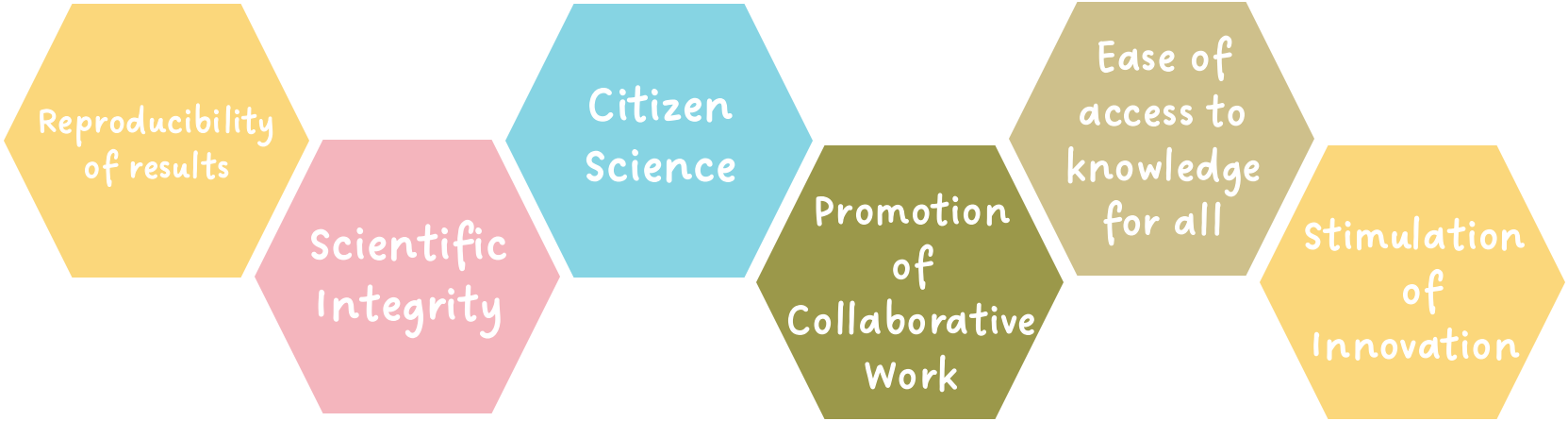
Brought to you by

Christelinda Laureijs
Julia Riley
Elizabeth Stregger

What is Open Science?

The process of making the content and process of producing evidence and claims transparent and accessible to others.

SUPPORTS:



Munafo et al. 2017

About this Workshop Series

THE LEADERS

Christelinda Laurejis

M.Sc. Candidate
in Biology

Elizabeth Stregger

Data and Digital
Services Librarian

Dr. Julia Riley

Assistant
Professor

We all love coding in R and open science!!!

About this Workshop Series

THE SERIES

Wednesdays from 5:00-6:30 PM in AVDX G10

Workshop # 1

29 Jan 2025

Welcome & Being Tidy

Dr. Riley

Workshop # 2

5 Feb 2025

Git with it!

Elizabeth Stregger

Workshop # 3

12 Feb 2025

Science Writing in R

Christelinda Laurejis

About this Workshop Series

WHAT CAN YOU EXPECT?

# 1	# 2	# 3
Welcome & Being Tidy	Git with it!	Science Writing in R
Both a mix of lecture and activities.		- 45 min hybrid lecture - 45 min in-person activitu

About this Workshop Series

WHAT CAN YOU EXPECT?

# 1 Welcome & Being Tidy	# 2 Git with it!	# 3 Science Writing in R
Both a mix of lecture and activities.		- 45 min hybrid lecture - 45 min in-person activity

One person will lead each workshop, and the two others will be "floaters". If you have an issue or question, put a RED post-it note on top of your laptop. Floaters will be by to help you out!

Open Science Skills in R – A Workshop Series

Welcome to Git with It!

Original black cat illustrations by Nellie Littlehale Umbstaetter, Available from the Internet Archive

Let's Git Set Up

OUR SOFTWARE TOOLKIT FOR THESE WORKSHOPS

R Software

- Open-source statistical programming language
- Also an environment for statistical computing and graphics that is easily extendable using *packages*

R Studio

- R Studio is a convenient interface for R called an IDE (integrated development environment; *e.g., "I write R code in the R Studio IDE"*)
- It is not a requirement for programming with R, but it is very commonly used by data scientists

Let's Git Set Up

OUR SOFTWARE TOOLKIT FOR THESE WORKSHOPS

☑ Git

- Open-source version control system
- Unique branching features

☑ GitHub

- Cloud-based platform
- *Share work, track and manage changes, collaborate*

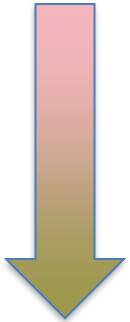
Git's Set Up!

So far, you've successfully:

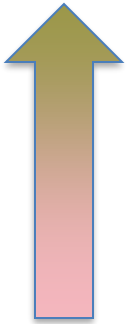
- Created a GitHub account
- Installed Git
- Created GitHub credentials and saved them in Rstudio

And if you ran into trouble, let's meet and work through it.

RStudio & GitHub: Push & Pull



Pull
Updates your local repository (Rstudio on your computer) with changes from the remote repository (in this case, GitHub)



Push
Updates the remote repository (in this case, GitHub) with changes in your local repository (Rstudio on your computer)

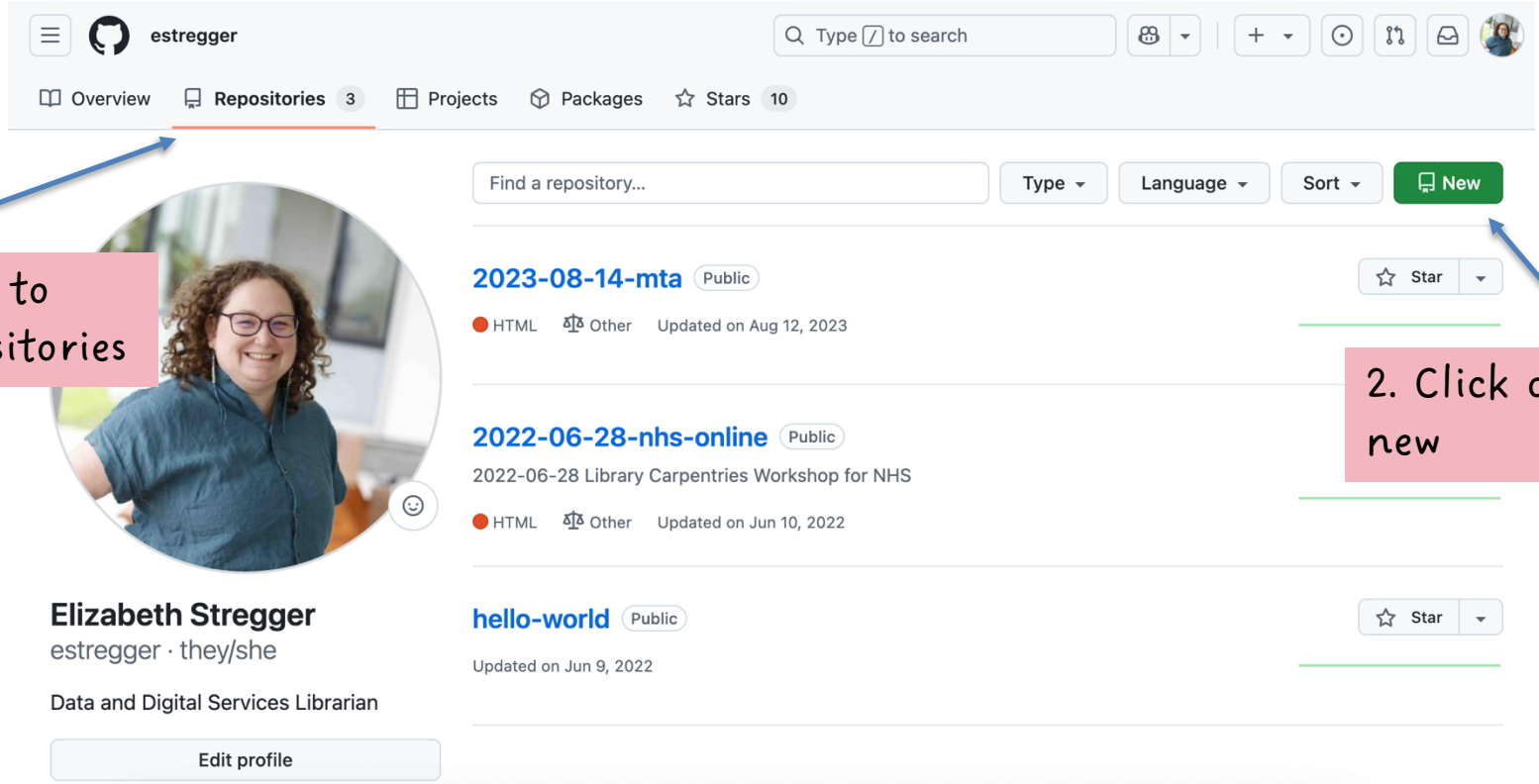
Tip for working with other people
Pull at the beginning of each work session to update your local repository

GitHub and R: Approaches

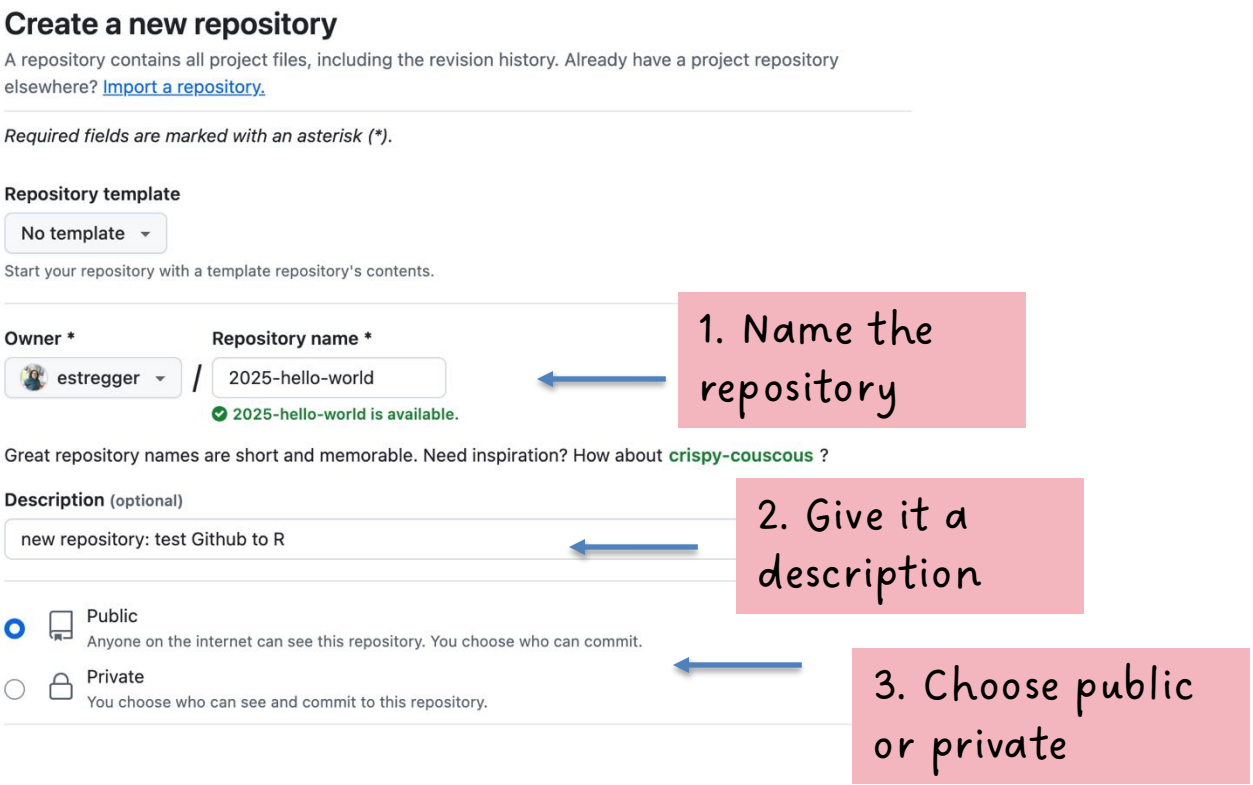
3 general approaches to getting started:

1. New project from GitHub pulled into R
2. Existing project from GitHub pulled into R
3. Existing R project made into Git repository and pushed to R

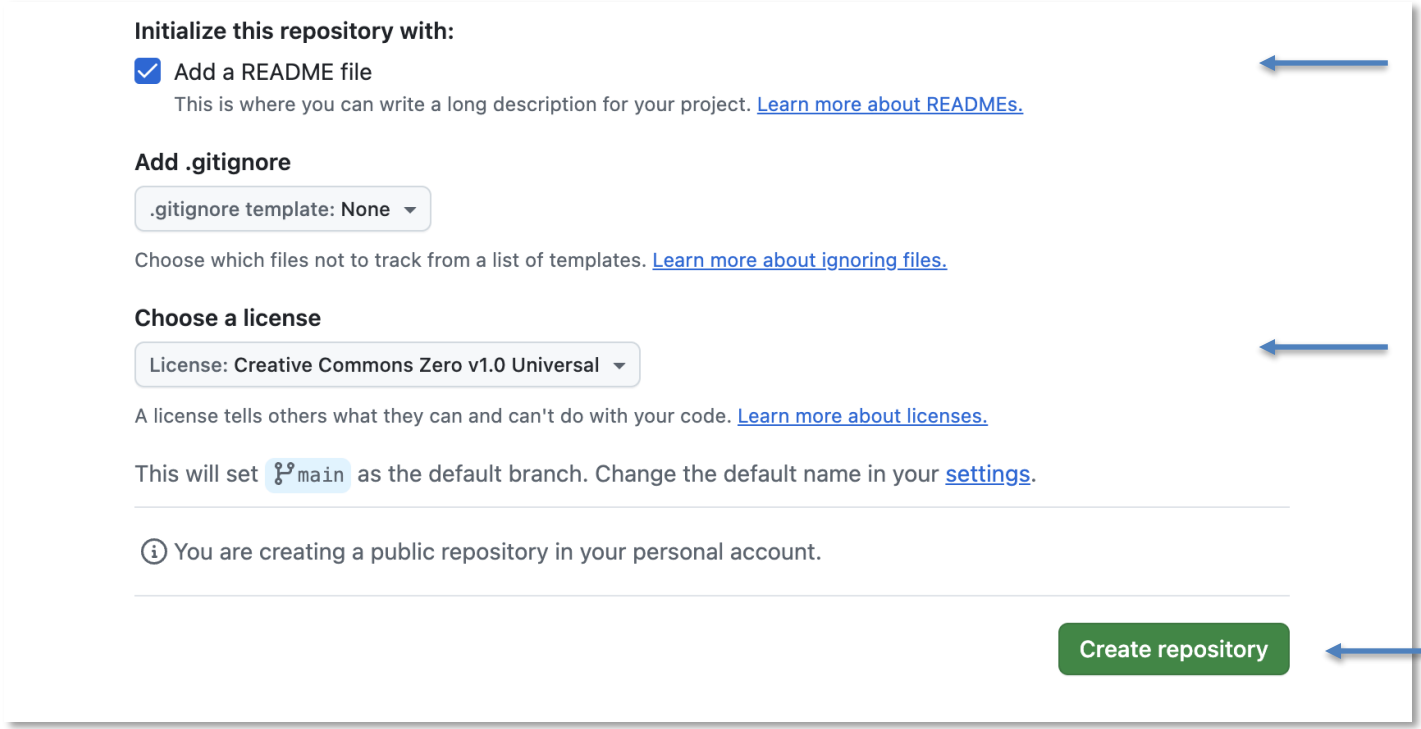
New: GitHub to R



New: GitHub to R



New: Github to R

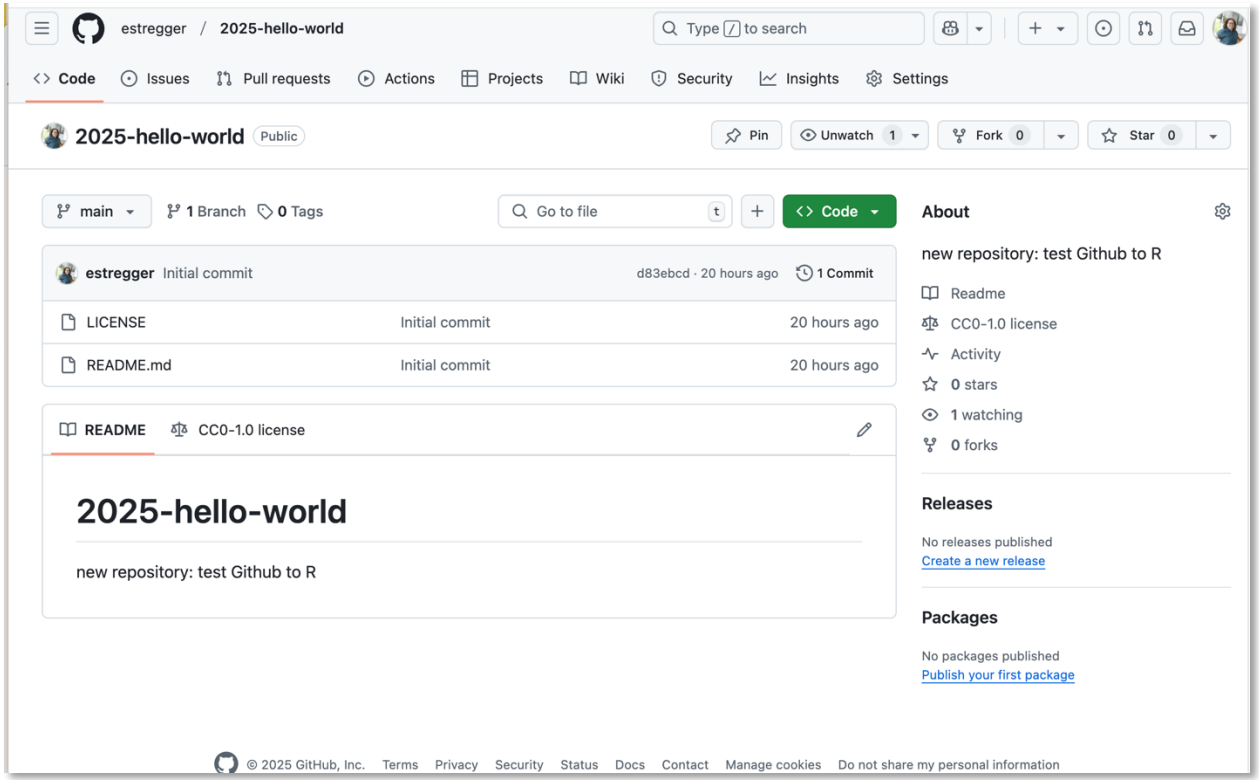


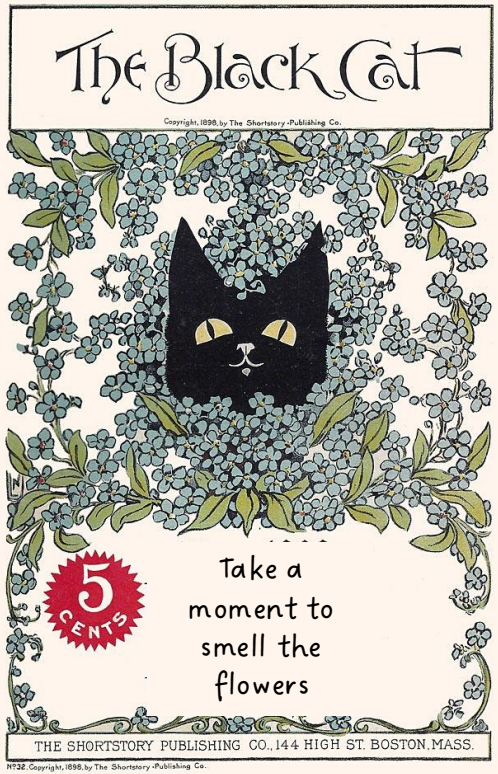
4. Add a README file

5. Choose a license

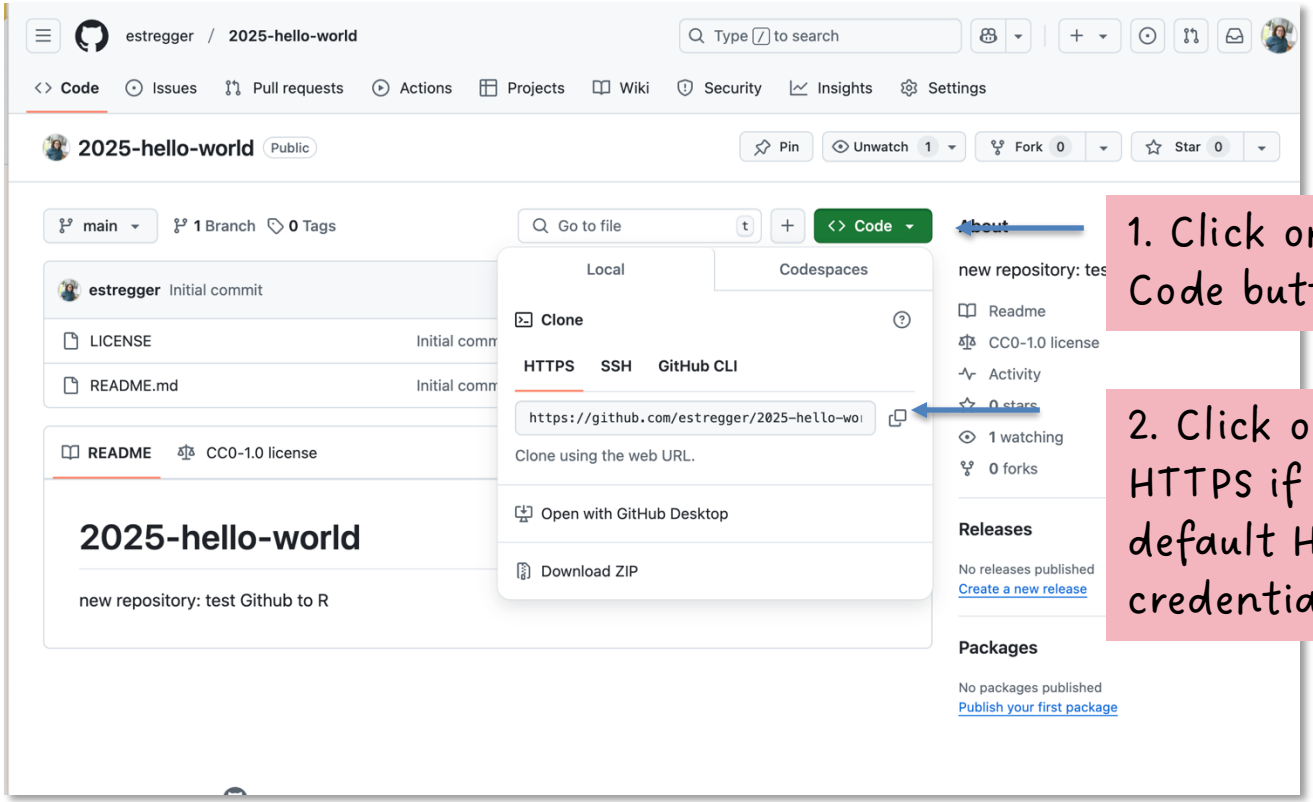
6. Create repository

New: GitHub to R





New: GitHub to R

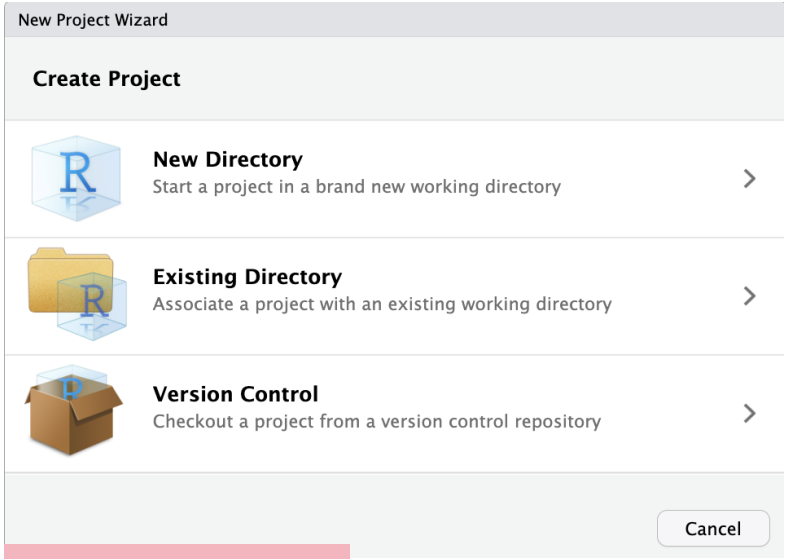


1. Click on Code button

2. Click on copy. Choose HTTPS if you used the default HTTPS credentials.

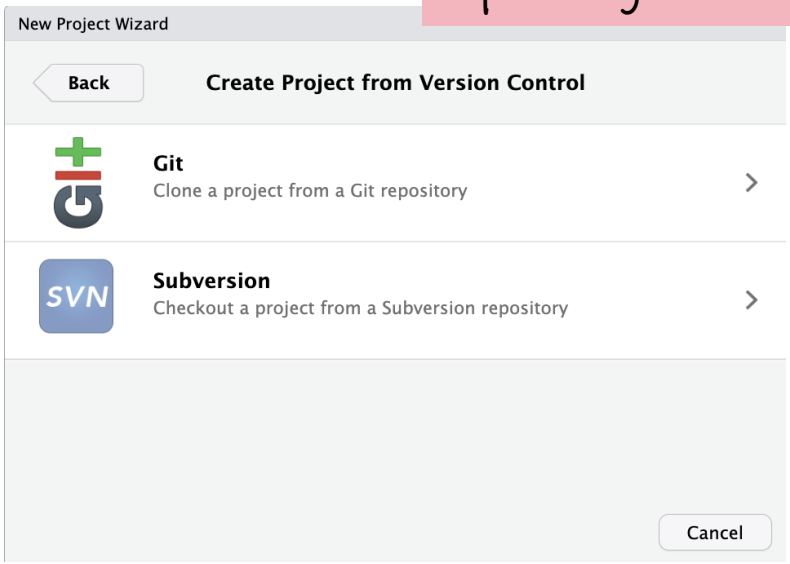
New: GitHub to R

1. Open Rstudio and go to File > New Project

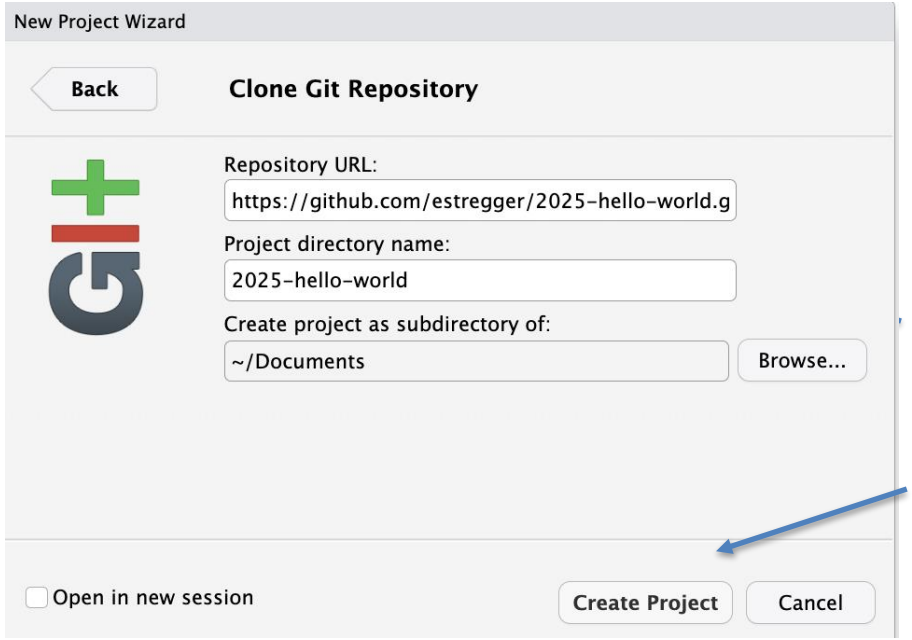


2. Select version control

3. Select Git: Clone a project from a Git repository



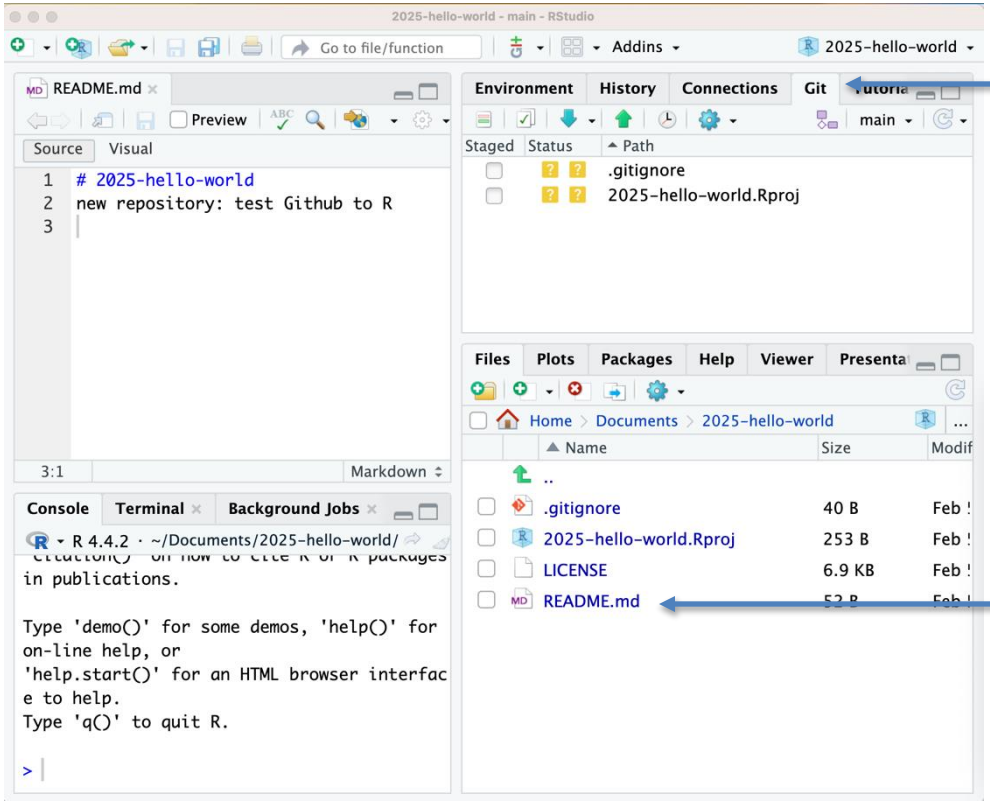
New: GitHub to R



1. Paste URL you got from the code button in GitHub

2. Give it a project directory name
And
3. Create Project

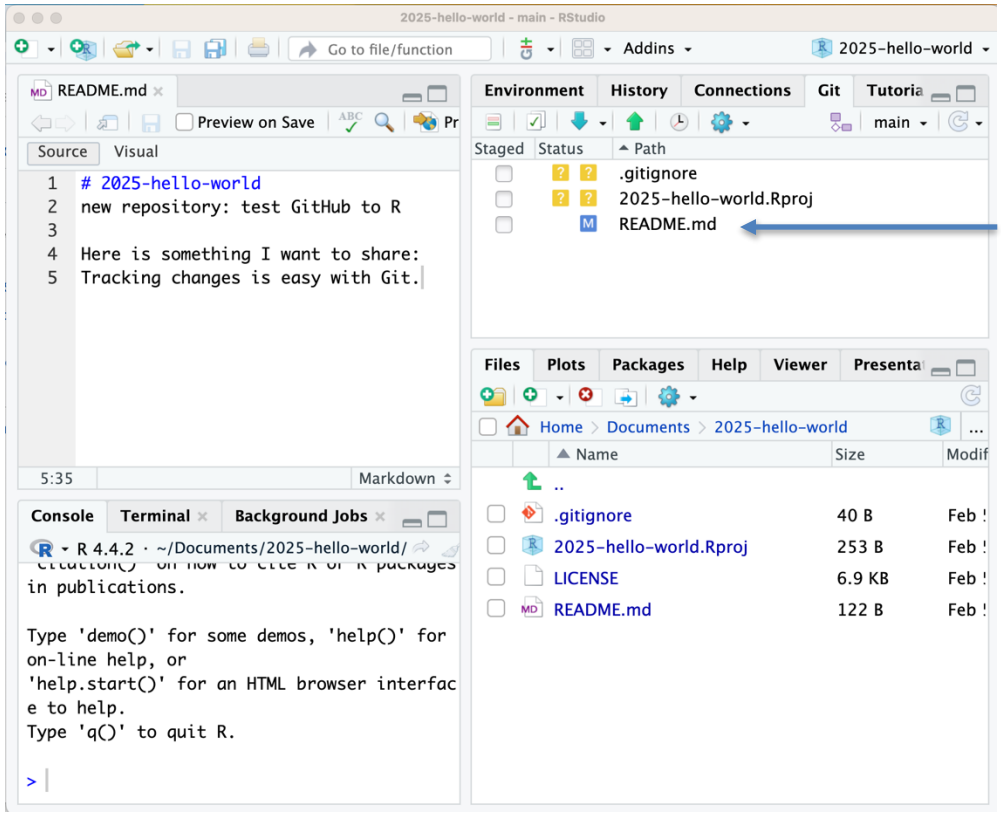
New: GitHub to R



1. Notice new tab: Git

2. Click on README.md file. This is the file you created when you created your repository

New: GitHub to R

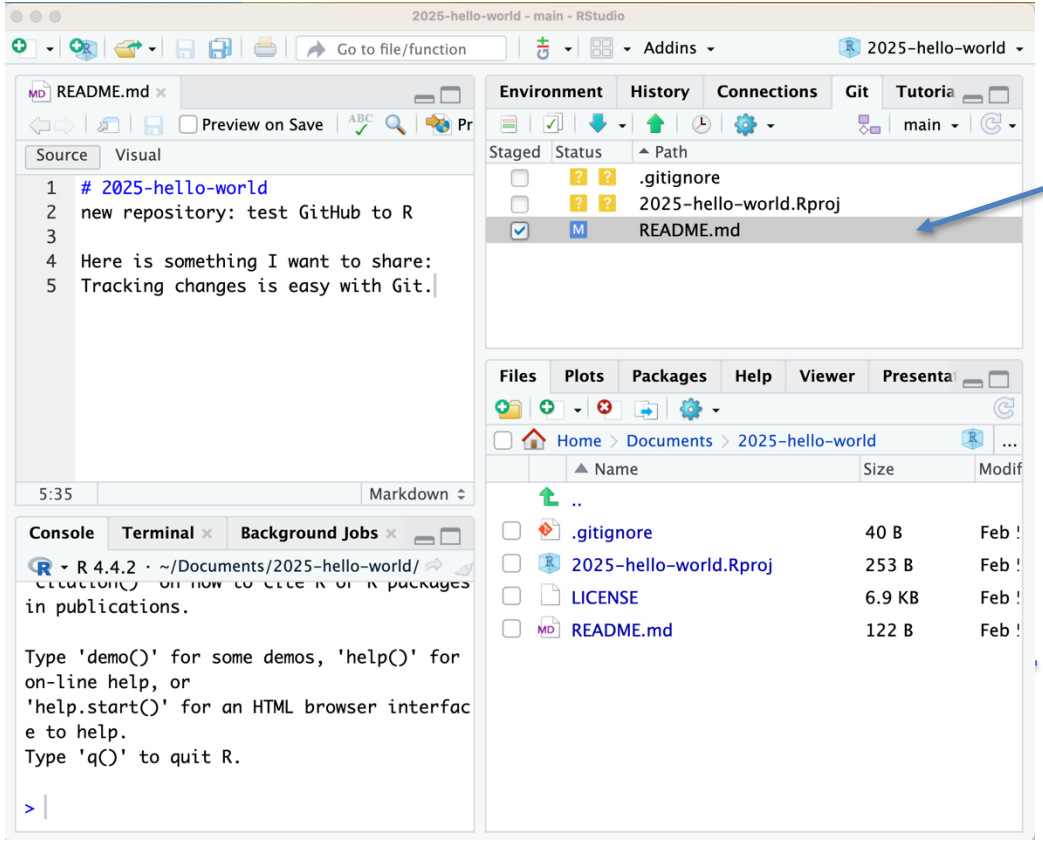


1. Make some changes to your README.md file and save it.

2. Your README.md file will appear in the Git pan with a blue M box next to it.

Git pane symbols:
? Untracked
M Modified
D Deleted
R Renamed

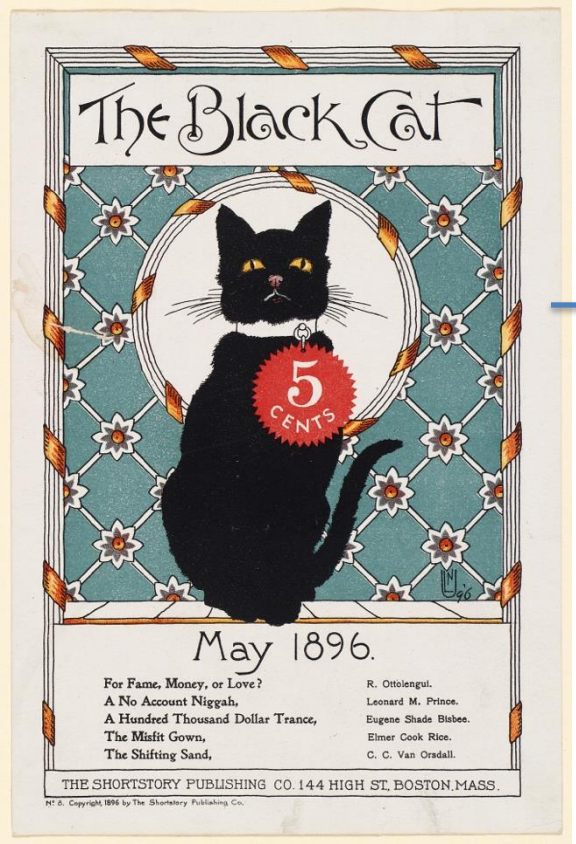
New: GitHub to R

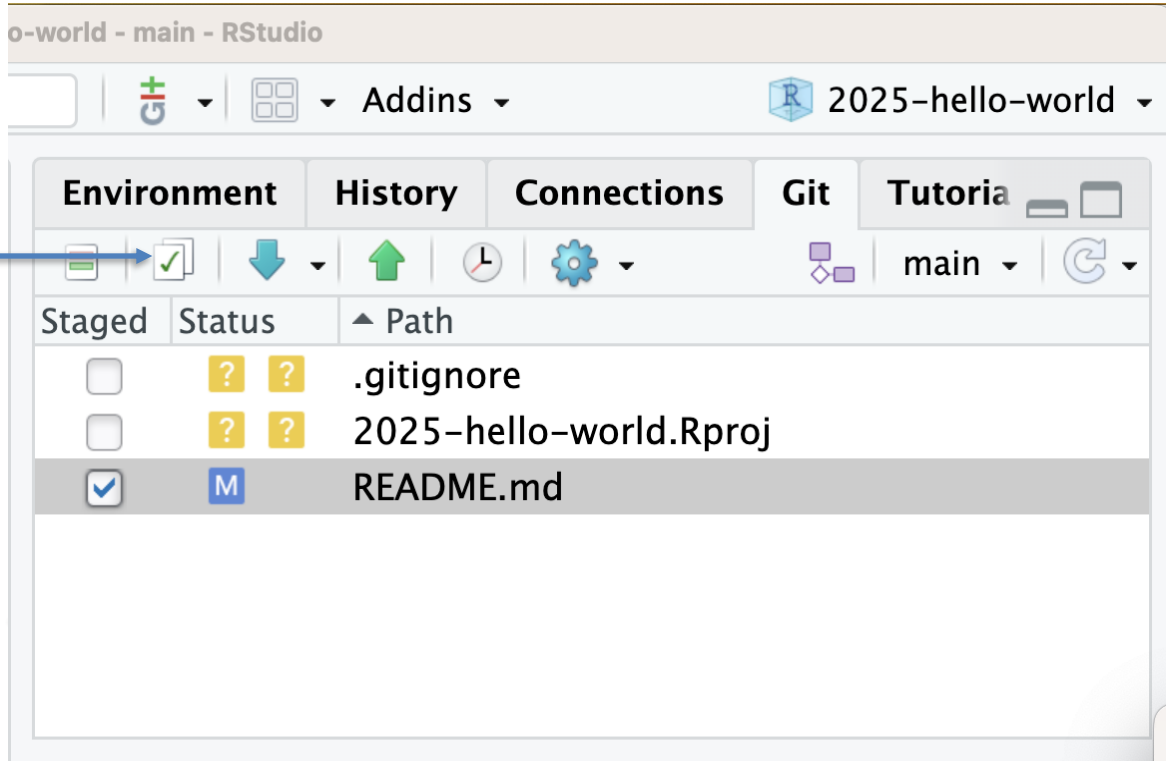


1. Tick the box next to README.md to stage it.

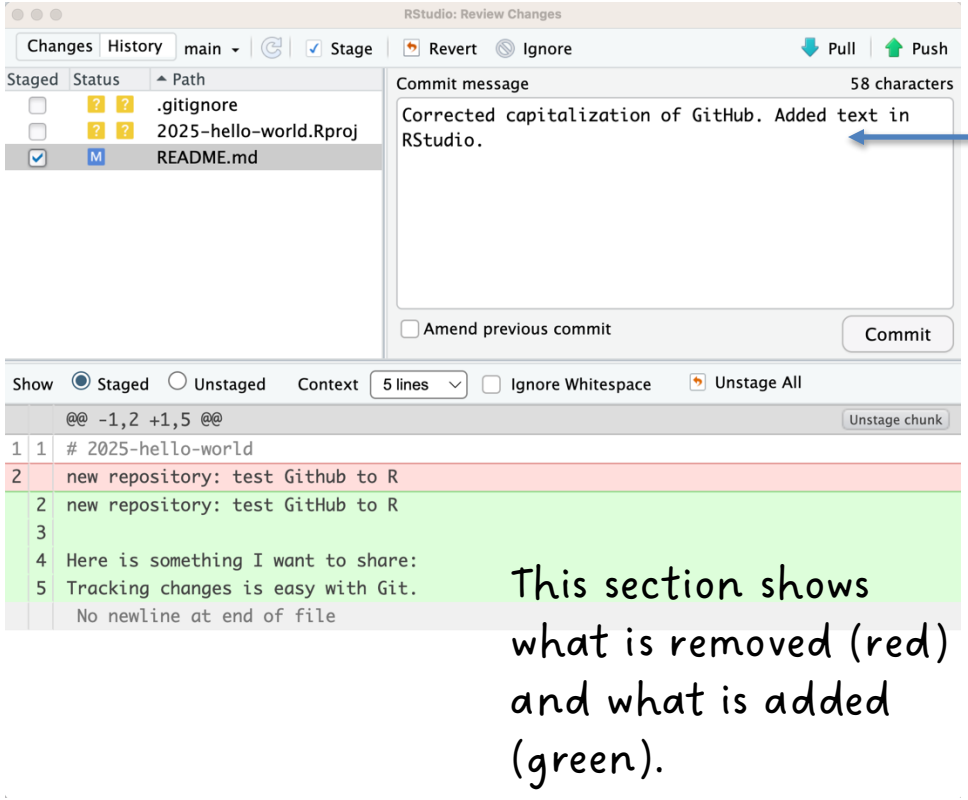
Stage a file when:
You're ready to take a snapshot of this new version of your file.
The staged files will be included in your next commit.

New: GitHub to R





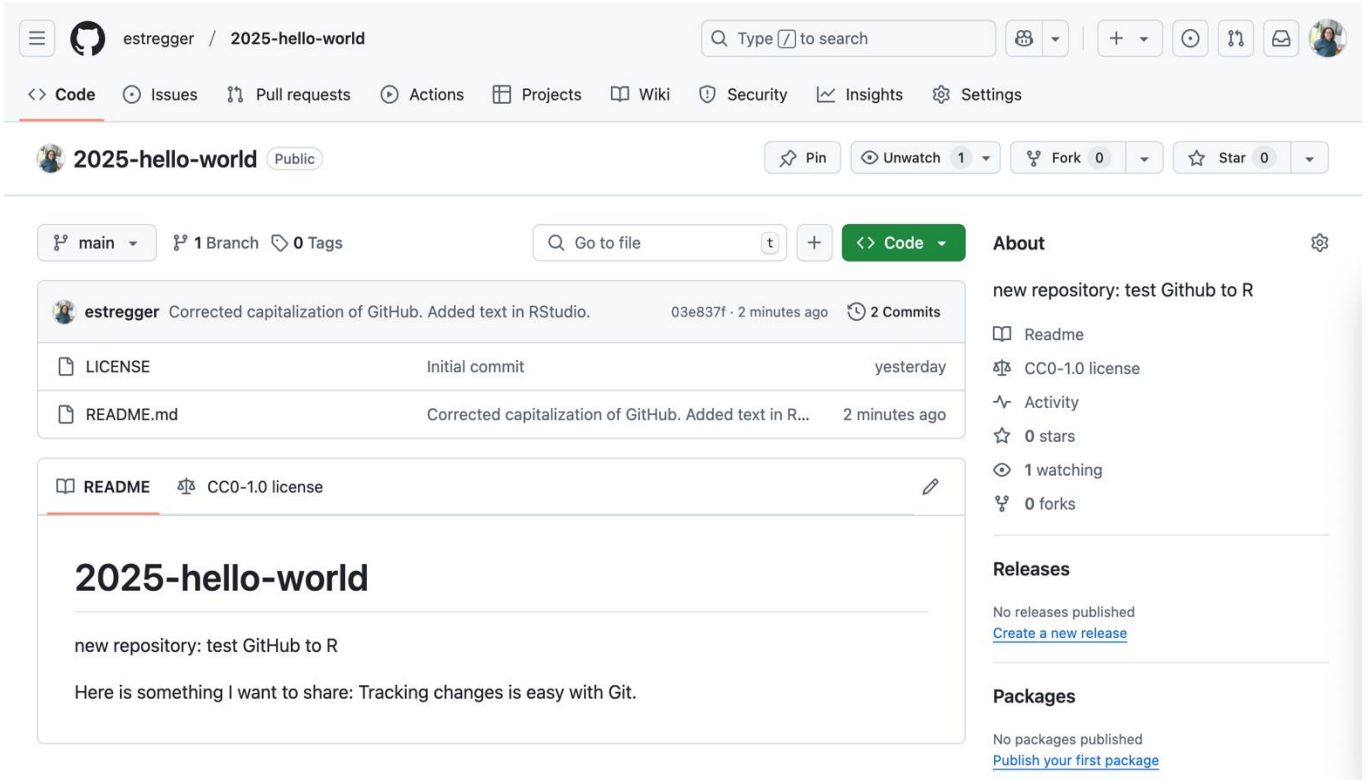
New: GitHub to R



This section shows what is removed (red) and what is added (green).

1. Describe changes in the commit message. Being detailed will save you (and others) time later.
2. Click Commit
3. Push using arrow.

New: GitHub to R



Back to GitHub (refresh if necessary) Which files changed? When?

Activity

Share account names with a few people.
Invite them to collaborate on your repository.
Accept invitations from others.
Clone their repositories to a new Rstudio project on your computer.
Create a new Rmd file, save it.
Pull to see if any additional changes have been made.
Stage your changes, commit them, and push them to the other person's repository.

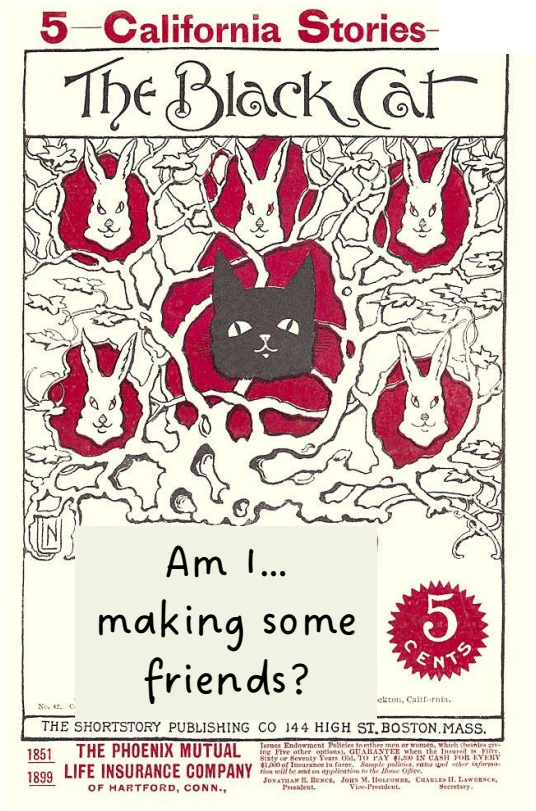
Branching

We were working on the main branch.

If you have multiple people working on solving a problem, creating a new branch can preserve a clean history of all approaches and reduce the frequency of merges.
The same is true if you are testing multiple approaches on your own!

Once you've decided on a solution, you can merge a branch back into the main branch.

Fork and Clone



Fork and Clone is the polite way to work on someone else's project if you're not already a collaborator.

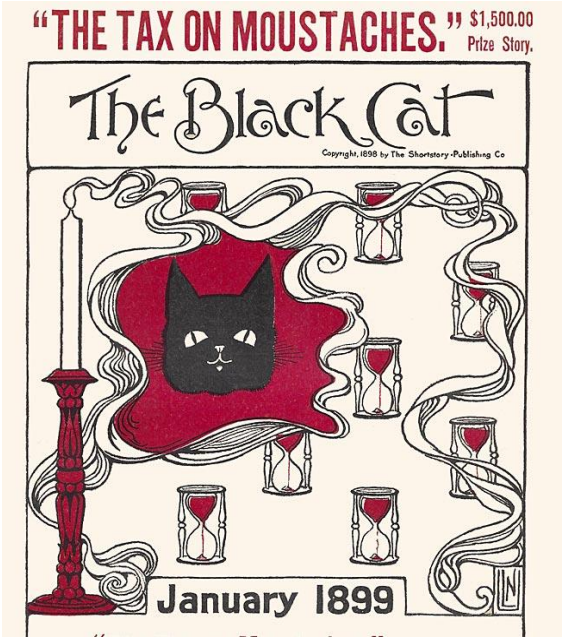
The original repository is called the upstream repository.

Instead of a branch, work is done on a fork.

Instead of pushing the changes to the repository, the person solving the problem submits a pull request.

The collaborators on the original project then decide if they want to pull code from the fork and merge it.

ANY LAST QUESTIONS?



Thank you.