

Using git with RStudio

Lind & Cariveau

February 5, 2018

What is git?

Distributed version control system



Search entire site...

Git is a **free and open source distributed version control system** designed to handle everything from small to very large projects with speed and efficiency.

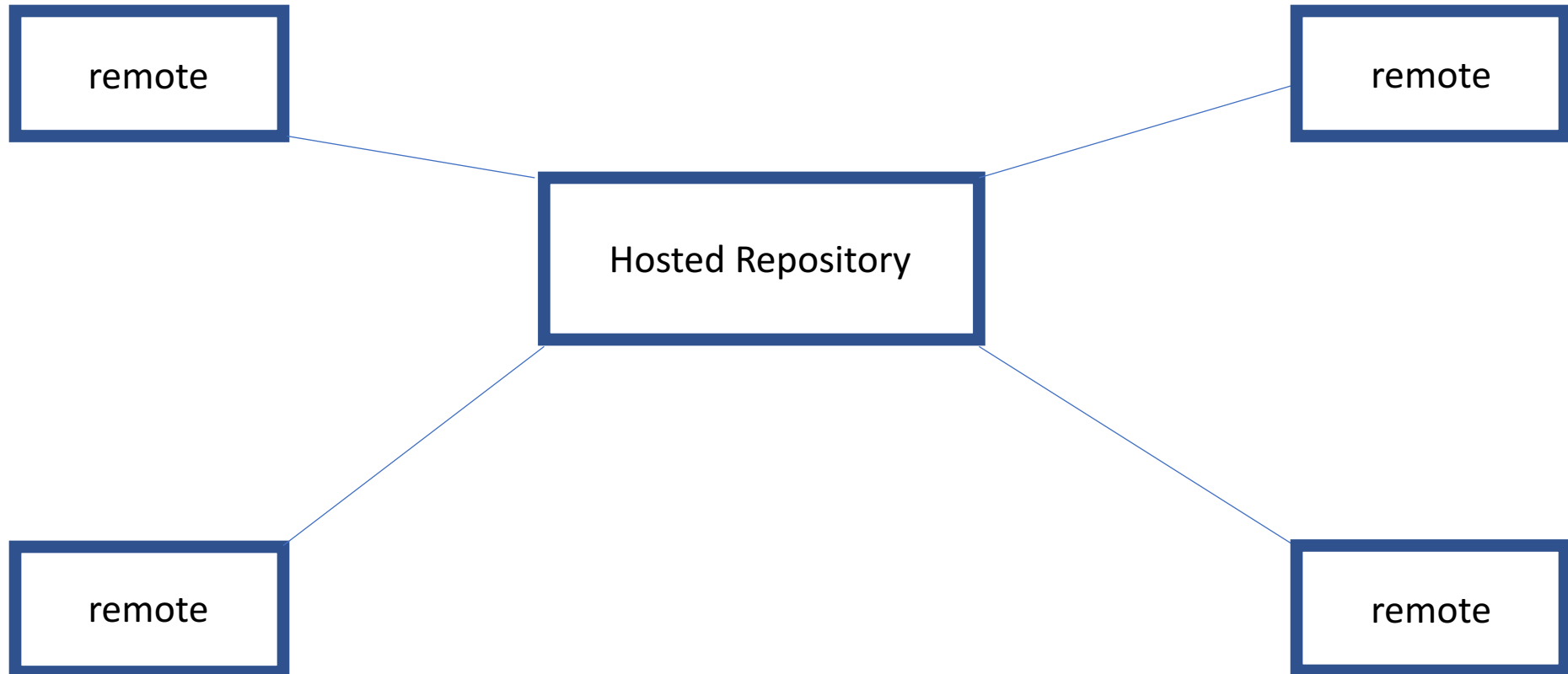
Git is **easy to learn** and has a **tiny footprint with lightning fast performance**. It outclasses SCM tools like Subversion, CVS, Perforce, and ClearCase with features like **cheap local branching**, convenient **staging areas**, and **multiple workflows**.



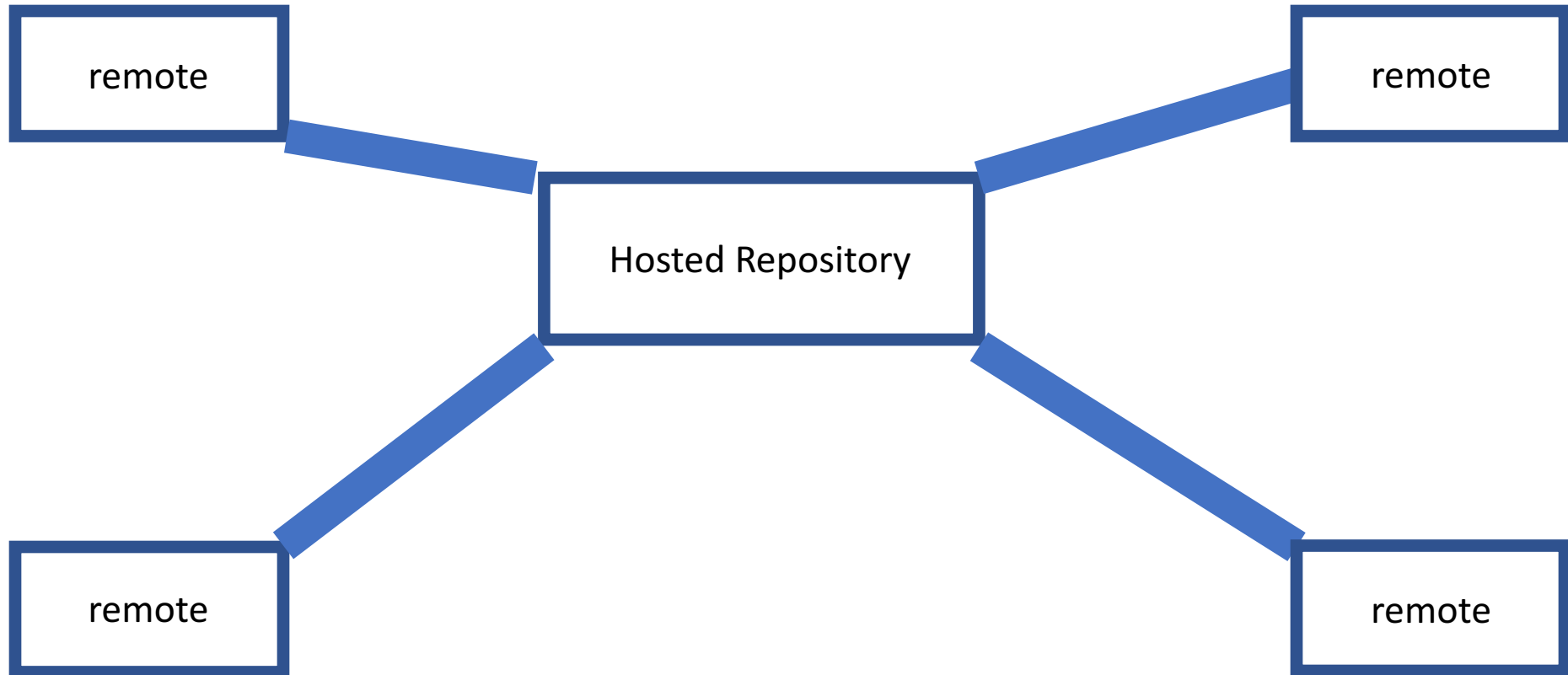
Learn Git in your browser for free with **Try Git**.



Distributed version control system



Distributed version control system



Repositories (Repos)

Projects

Tracked by git

Each repo is tracked independently

Repositories (Repos)

Public Repository: everyone can download and suggest edits (pull request); *collaborators* can make most changes

Private Repository: only *collaborators* can see and make changes

git vs. github vs. github.umn vs. bitbucket

git: distributed version control **system**

Repository hosting sites:

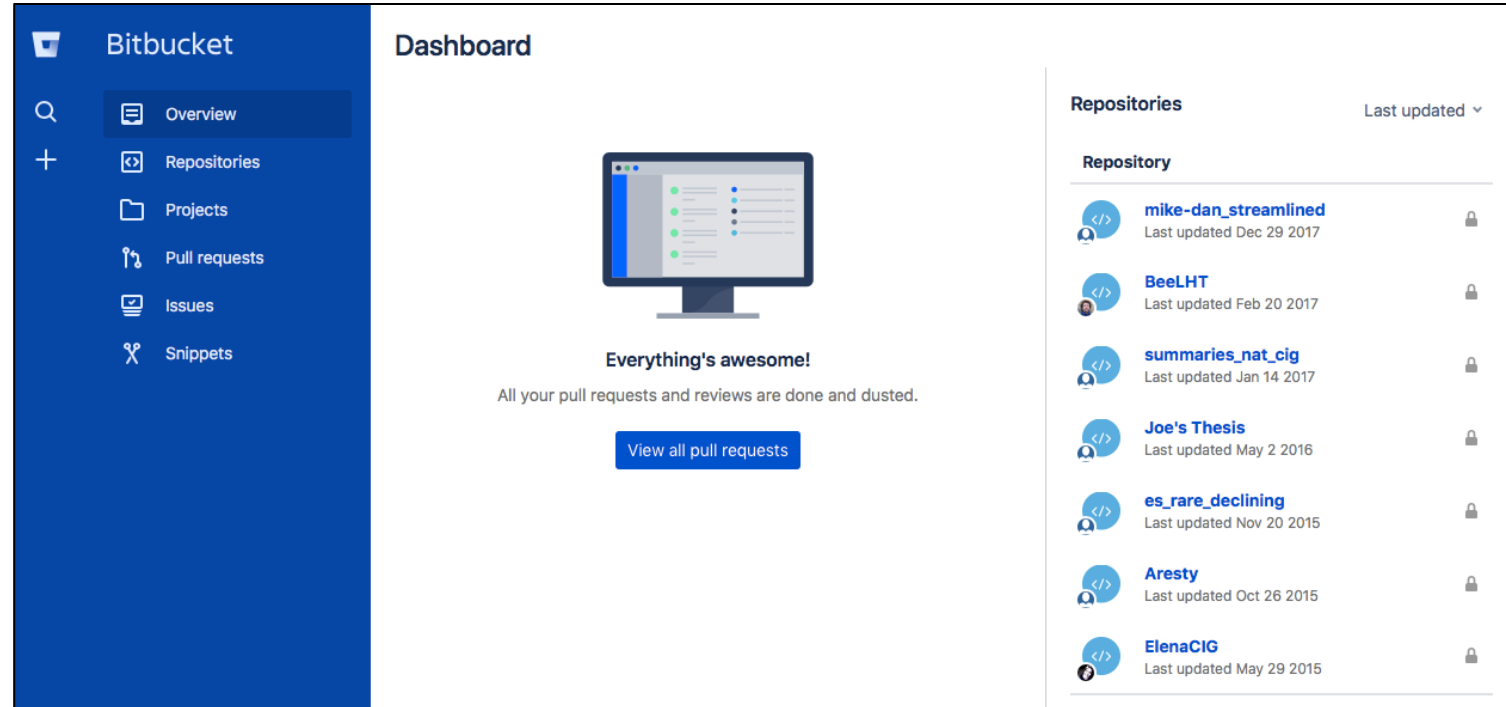
github, github.umn.edu, Bitbucket

Bitbucket

Free versions

Pay based on # users

Unlimited private
and public
repositories



The screenshot shows the Bitbucket dashboard interface. On the left is a blue sidebar with the Bitbucket logo and navigation options: Overview (selected), Repositories, Projects, Pull requests, Issues, and Snippets. The main content area is titled 'Dashboard' and features a central graphic of a computer monitor displaying a code diff. Below the graphic, the text reads 'Everything's awesome!' and 'All your pull requests and reviews are done and dusted.', with a 'View all pull requests' button. On the right side, there is a 'Repositories' section with a 'Last updated' dropdown menu. It lists several repositories with their names and last update dates: mike-dan_streamlined (Dec 29 2017), BeeLHT (Feb 20 2017), summaries_nat_cig (Jan 14 2017), Joe's Thesis (May 2 2016), es_rare_declining (Nov 20 2015), Aresty (Oct 26 2015), and ElenaCIG (May 29 2015). Each repository entry includes a lock icon on the right.

github.com

Most used

Free for public
repositories

Pay based # private
repositories

The screenshot shows the GitHub profile page for Daniel dancariveau. The profile includes a pink pixelated avatar, the name 'Daniel dancariveau', and a bio 'University of Minnesota, St. Paul, Minnesota'. It lists contact information: email 'dancariveau@gmail.com' and website 'http://dancariveau.com'. The profile statistics show 6 repositories, 0 stars, 2 followers, and 2 following. The 'Popular repositories' section features 'Main_Trial' (HTML) and 'data-mgmt-4-biologists' (Jupyter Notebook). A contribution graph shows 92 contributions in the last year, with a legend indicating the number of contributions per day.

Search GitHub

Pull requests Issues Marketplace Explore

Overview Repositories 6 Stars 0 Followers 2 Following 2

Popular repositories Customize your pinned repositories

Main_Trial
This is just a trial
HTML

data-mgmt-4-biologists
Forked from ericlind/data-mgmt-4-biologists
Jupyter Notebook

Daniel
dancariveau
Add a bio

University of Minnesota
St. Paul, Minnesota
dancariveau@gmail.com
http://dancariveau.com

92 contributions in the last year Contribution settings

Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan

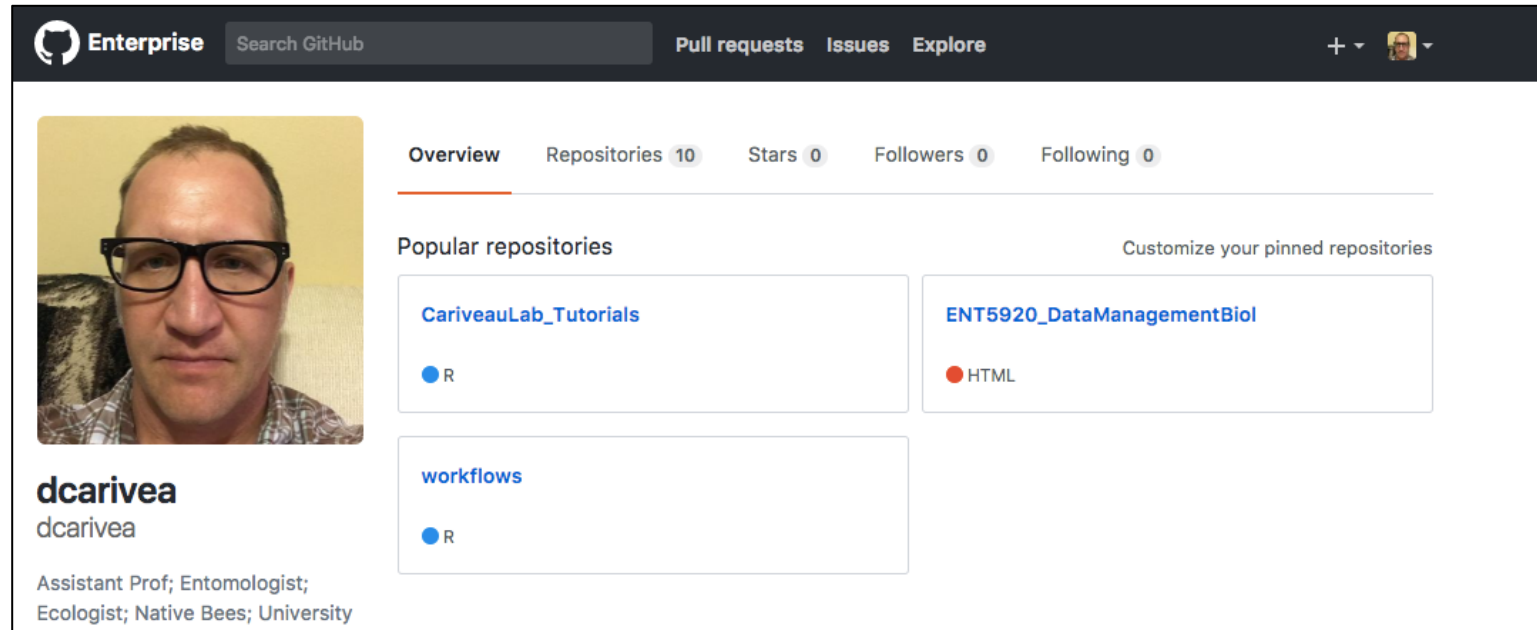
Mon
Wed
Fri

Learn how we count contributions. Less More

github.umn.edu

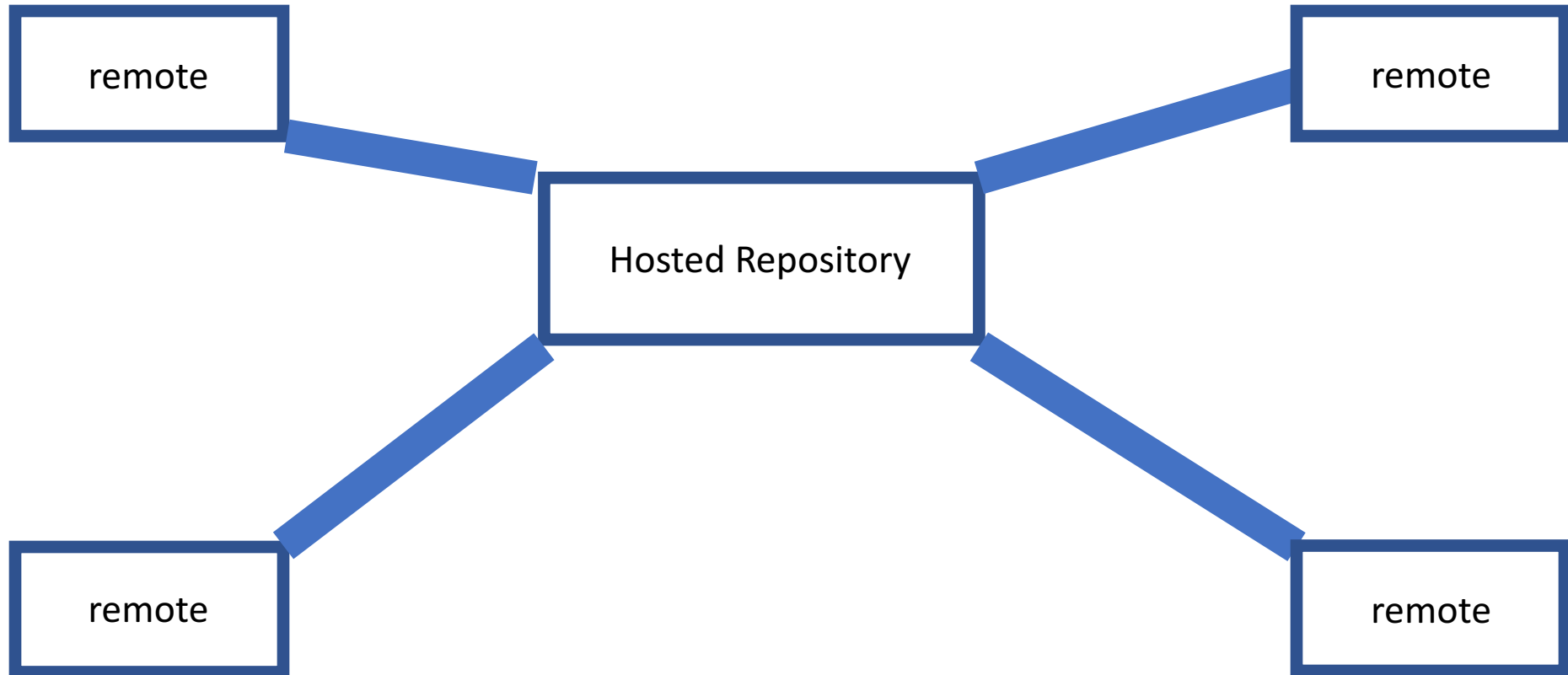
Free for “public” and private repositories

Public = those with umn github accounts



The screenshot shows a GitHub profile page for a user named **dcarivea**. The page is part of a GitHub Enterprise interface. At the top, there is a navigation bar with the GitHub logo, the word "Enterprise", a search bar labeled "Search GitHub", and links for "Pull requests", "Issues", and "Explore". On the right side of the navigation bar, there is a user profile icon and a plus sign. Below the navigation bar, the profile page is displayed. On the left, there is a profile picture of a man with glasses. Below the picture, the username **dcarivea** is shown, followed by the GitHub handle `dcarivea` and a bio: "Assistant Prof; Entomologist; Ecologist; Native Bees; University". To the right of the profile picture, there are statistics: "Overview", "Repositories 10", "Stars 0", "Followers 0", and "Following 0". Below these statistics, there is a section titled "Popular repositories" with a link to "Customize your pinned repositories". Three repositories are listed: "CariveauLab_Tutorials" (R), "ENT5920_DataManagementBiol" (HTML), and "workflows" (R).

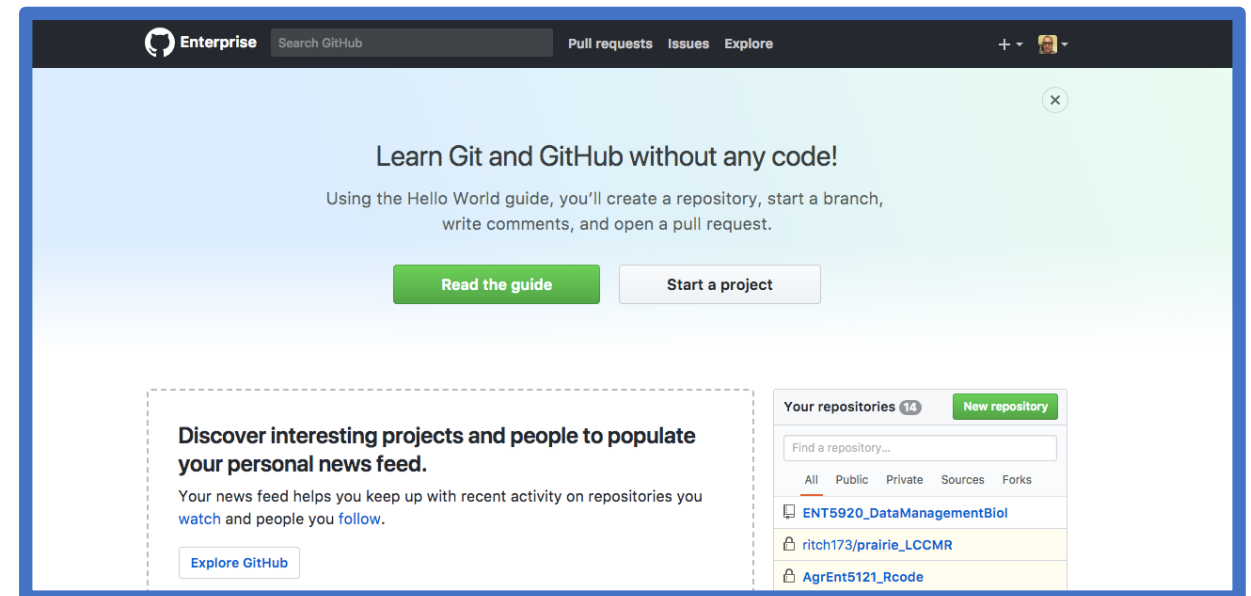
Distributed version control system



Version Control:

Step 1 Create Remote Repo

Hosted
Repository



The screenshot displays the GitHub Enterprise user interface. At the top, there is a navigation bar with the GitHub logo, the word "Enterprise", a search bar for GitHub, and links for "Pull requests", "Issues", and "Explore". A user profile icon is visible in the top right corner. The main content area features a large light blue and green gradient background with the text "Learn Git and GitHub without any code!" and a sub-header "Using the Hello World guide, you'll create a repository, start a branch, write comments, and open a pull request." Below this text are two buttons: "Read the guide" (green) and "Start a project" (white with a grey border). In the bottom right corner, there is a section titled "Your repositories 14" with a "New repository" button. Below this is a search bar "Find a repository..." and a list of repositories: "ENT5920_DataManagementBiol", "ritch173/prairie_LCCMR", and "AgrEnt5121_Rcode". On the bottom left, there is a dashed box containing the text "Discover interesting projects and people to populate your personal news feed." and a button "Explore GitHub".

Version Control:

Step 2 Clone Repo to Workspace

Local
Workspace



clone

Hosted
Repository

```
getRare.R x
1 library(RMySQL)
2 library(tidyverse)
3
4 source("psw.R") #simply save a .R file as psw="PASSWORD"
5 source("user.R")
6
7 lower <- function(x) {tolower(names(x))}
8 #specs<-read.csv("data/other/spectimens.csv")
9
10
11- #####
12- #### amnh #####
13- #####
14
15 conn <- dbConnect(driver="MySQL", user=user, password=psw,
16 dbname="mydatabase", port = "myport", host= "my_IP")
17
18 dbLstTables(conn)
19
20 amnh = dbReadTable(conn, "dryad_amnh")
21 amnh[1:nrow(amnh), "amnh"]
16.63 [limited]
R Script
>

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
```

Enterprise Search GitHub Pull requests Issues Explore

Learn Git and GitHub without any code!

Using the Hello World guide, you'll create a repository, start a branch, write comments, and open a pull request.

[Read the guide](#) [Start a project](#)

Your repositories 14 [New repository](#)

Find a repository...

All Public Private Sources Forks

- ENT5920_DataManagementBiol
- ritch173/prairie_LCCMR
- AgrEnt5121_Rcode

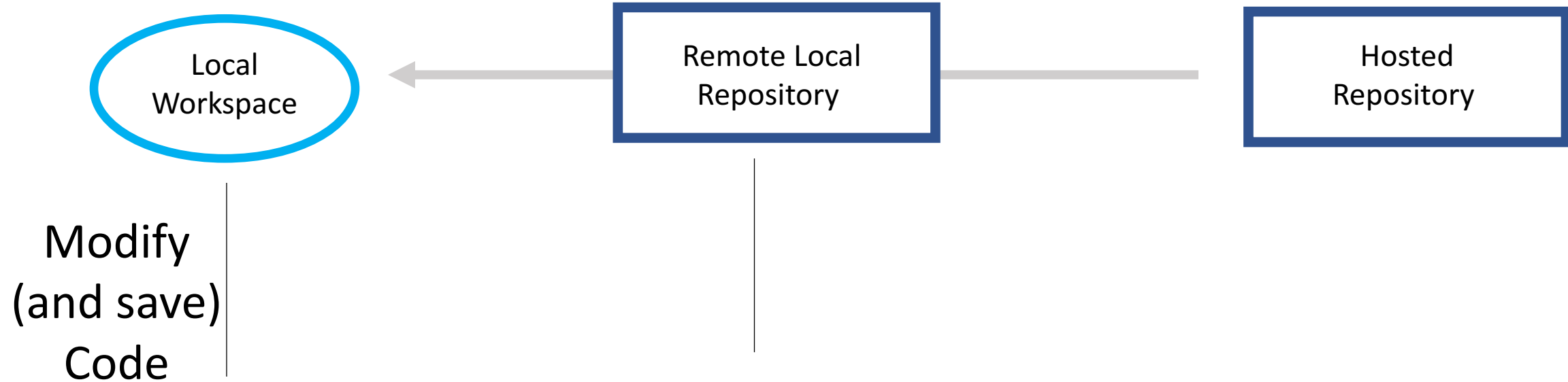
Discover interesting projects and people to populate your personal news feed.

Your news feed helps you keep up with recent activity on repositories you watch and people you follow.

[Explore GitHub](#)

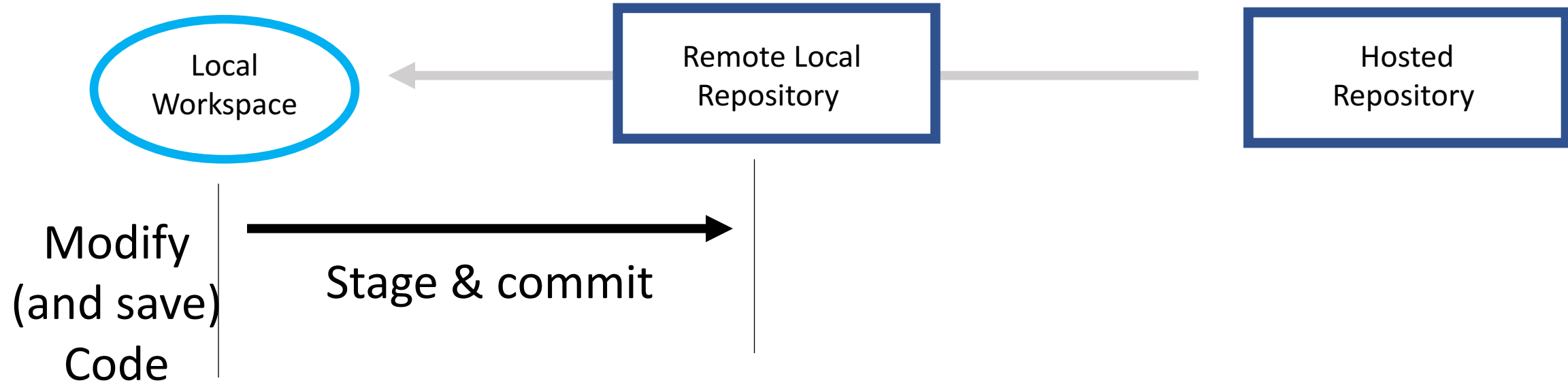
Version Control:

Step 3 Modify, Stage and Commit



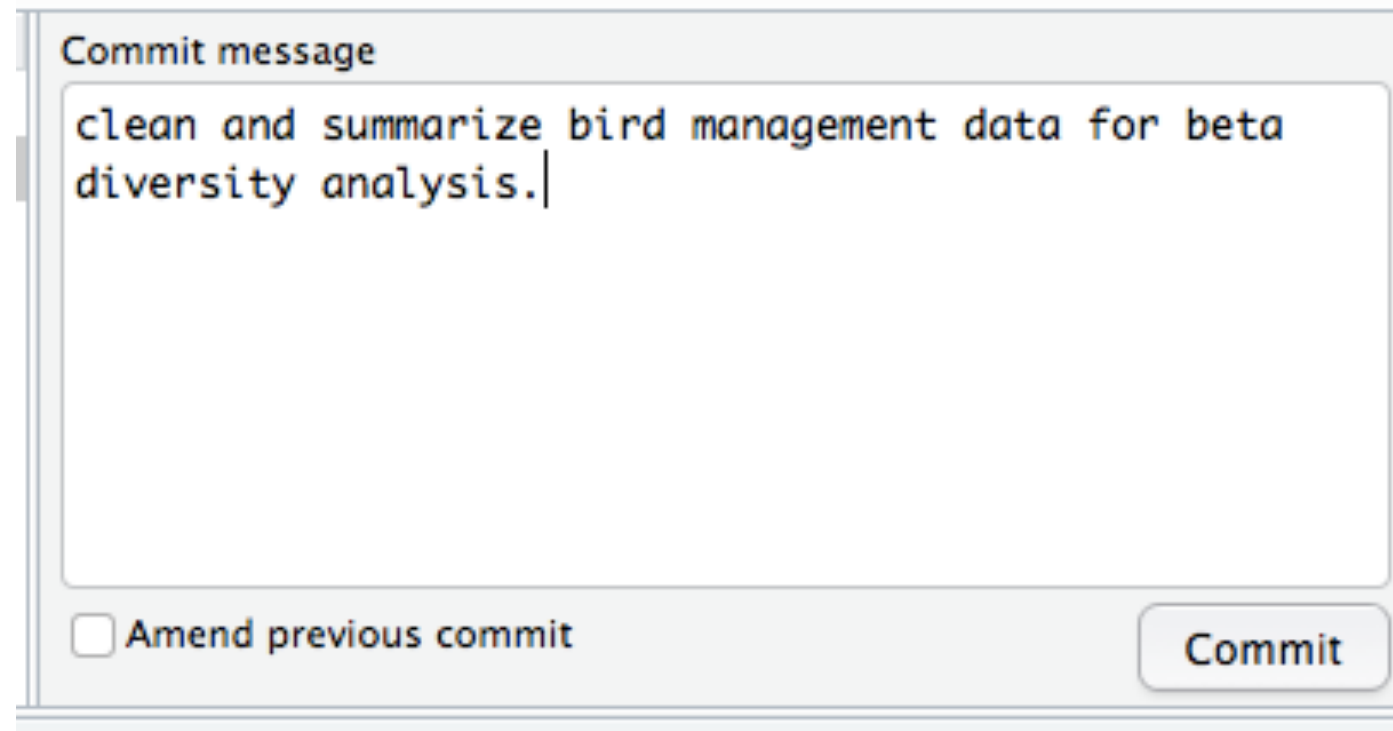
Version Control:

Step 3 Modify, Stage and Commit



Version Control: Good Commits

If accepted this commit will...{commit message here}



A screenshot of a commit message dialog box. The title bar reads "Commit message". The main text area contains the message: "clean and summarize bird management data for beta diversity analysis." followed by a cursor. At the bottom left, there is a checkbox labeled "Amend previous commit" which is currently unchecked. At the bottom right, there is a button labeled "Commit".

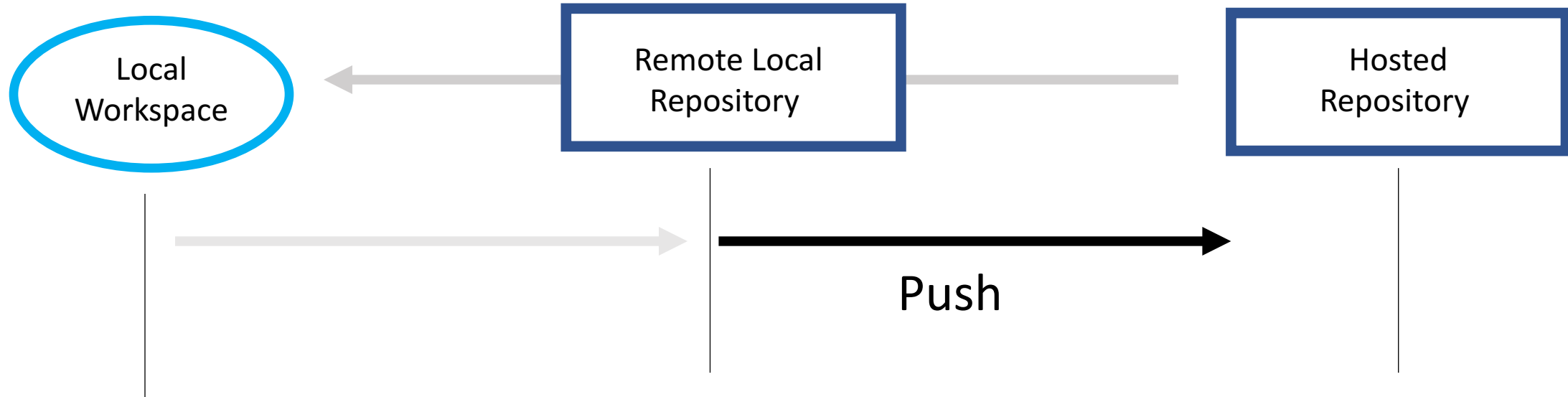
Version Control: Good Commits

	COMMENT	DATE
○	CREATED MAIN LOOP & TIMING CONTROL	14 HOURS AGO
○	ENABLED CONFIG FILE PARSING	9 HOURS AGO
○	MISC BUGFIXES	5 HOURS AGO
○	CODE ADDITIONS/EDITS	4 HOURS AGO
○	MORE CODE	4 HOURS AGO
○	HERE HAVE CODE	4 HOURS AGO
○	AAAAAAAAA	3 HOURS AGO
○	ADKFJSLKDFJSDKLFJ	3 HOURS AGO
○	MY HANDS ARE TYPING WORDS	2 HOURS AGO
○	HAAAAAAAAAANDS	2 HOURS AGO

AS A PROJECT DRAGS ON, MY GIT COMMIT MESSAGES GET LESS AND LESS INFORMATIVE.

Version Control:

Step 4 Push



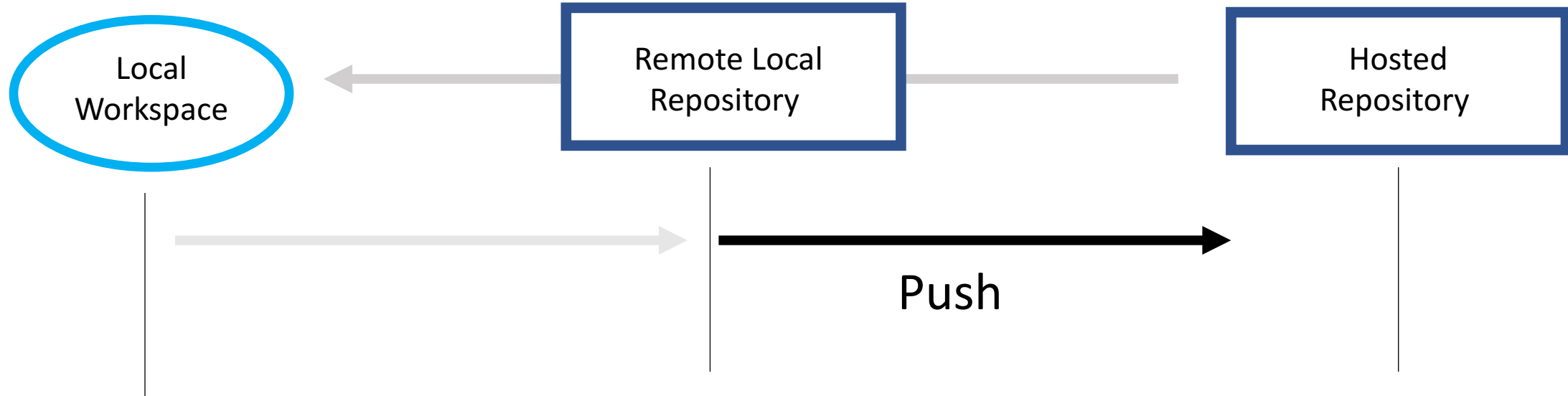
1. Collaborators can push without getting an “OK”

2. Public repositories: Others can create a pull request
(We will not cover today)

Version Control:

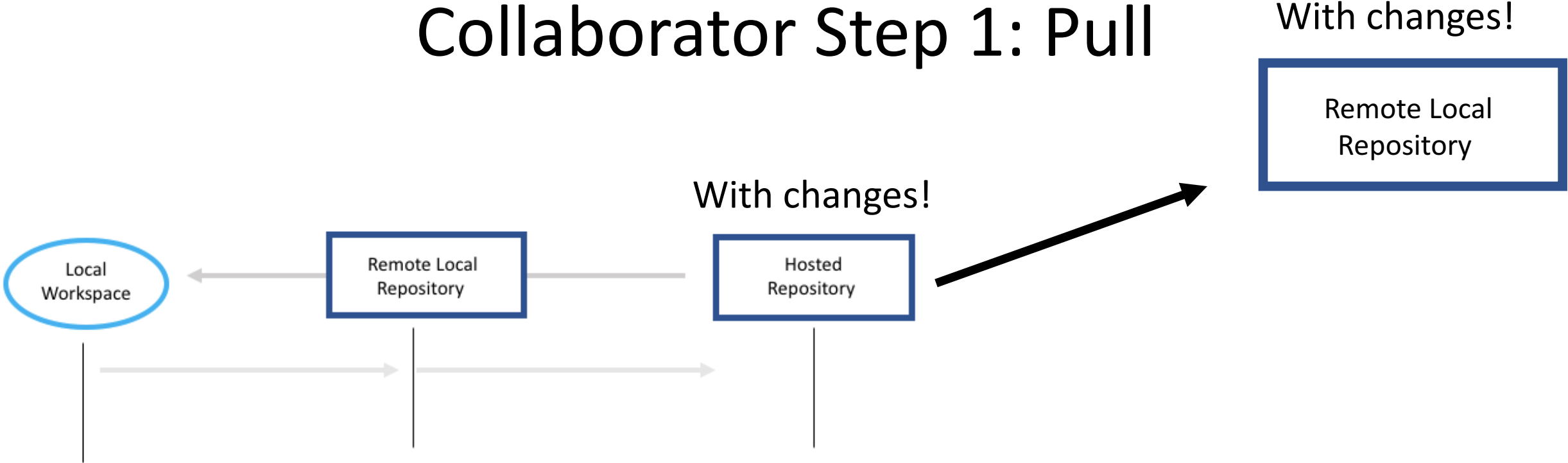
Step 4 Push

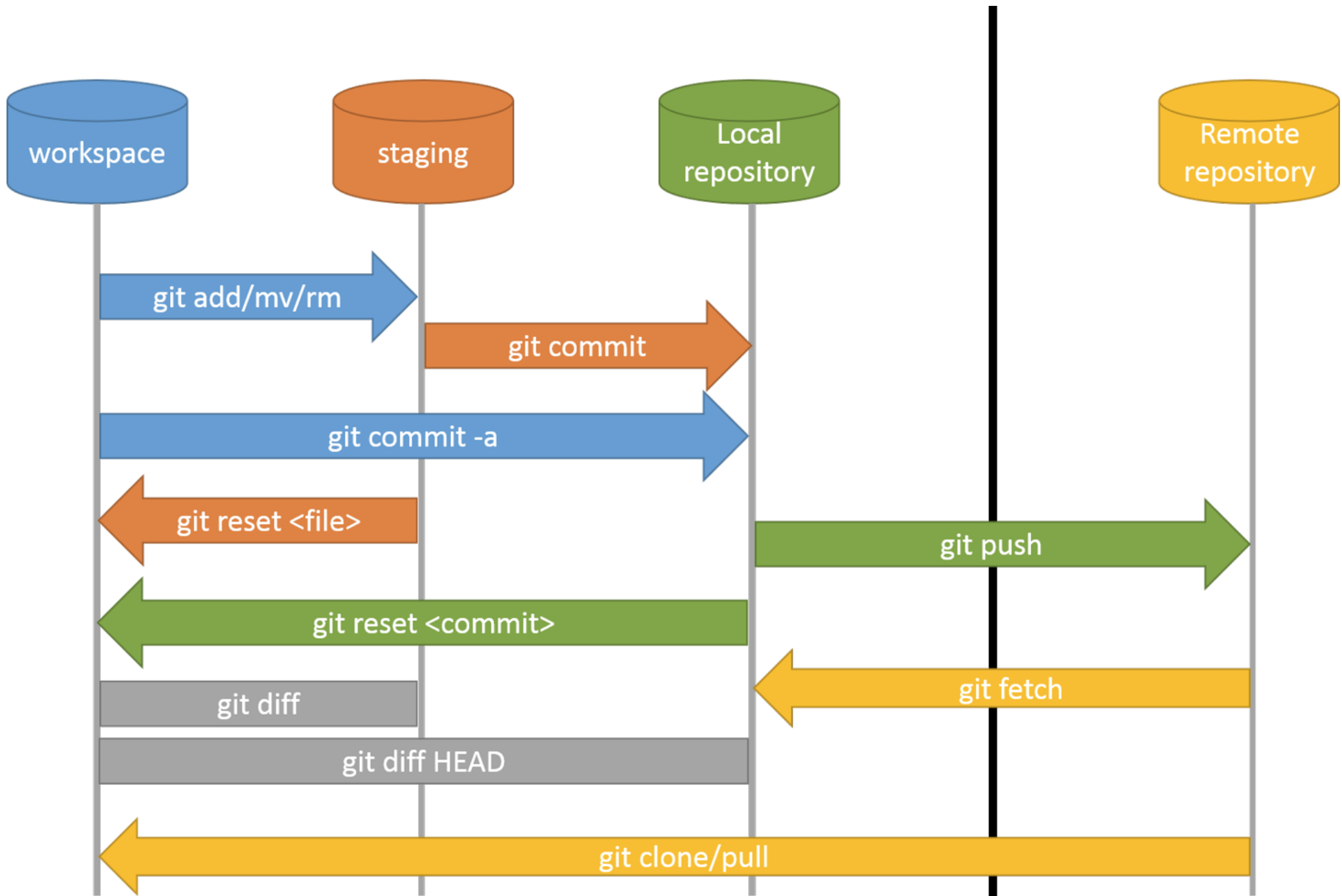
With changes!



Version Control:

Collaborator Step 1: Pull

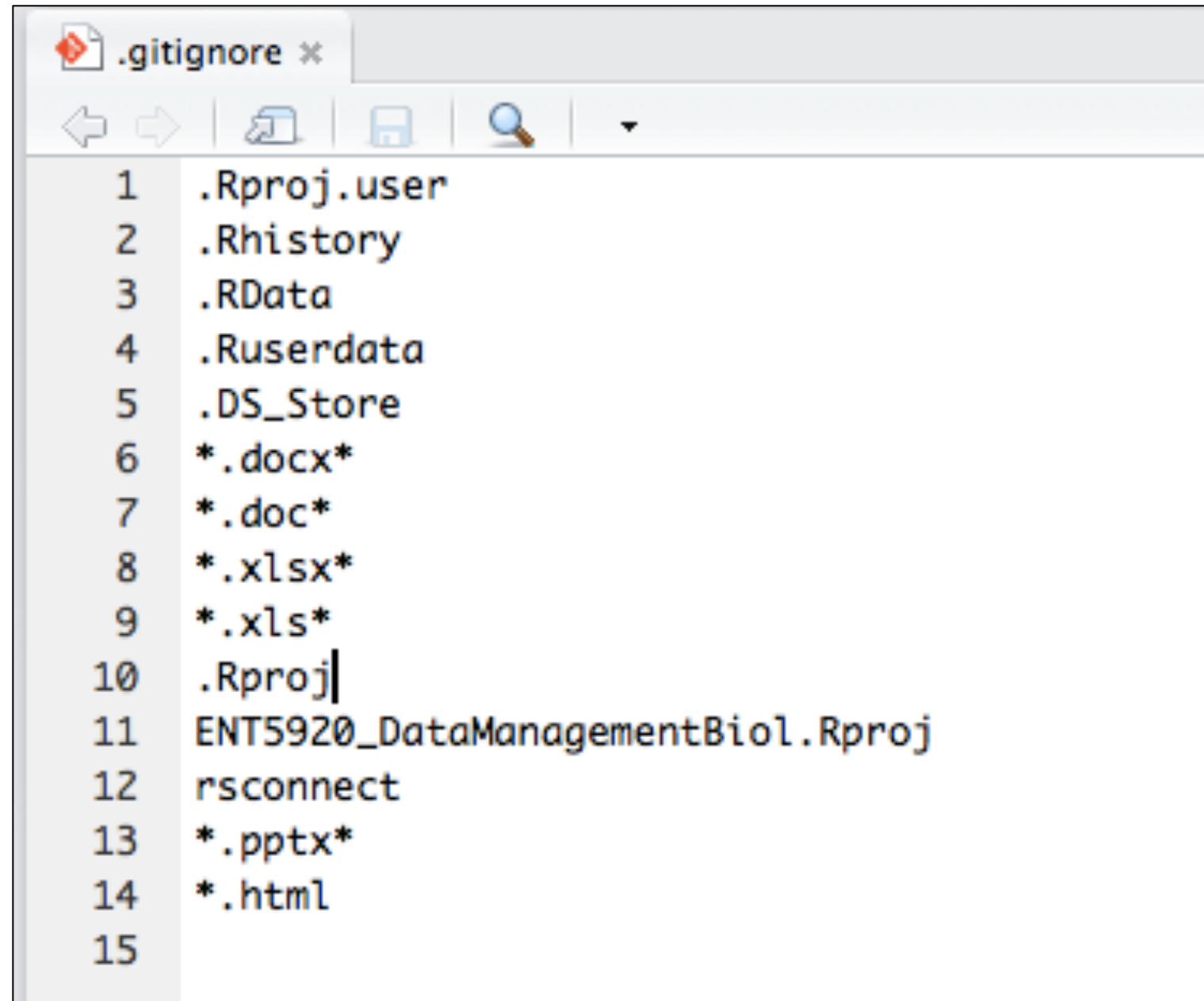




A few other things

- git is not typically for data storage (but can work fine here)
- Don't use version control for .xlsx, .pptx, and other complex files

.gitignore



A screenshot of a code editor window showing a .gitignore file. The window title is ".gitignore *". The editor has a toolbar with icons for back, forward, copy, save, search, and a dropdown menu. The content of the file is as follows:

```
1 .Rproj.user
2 .Rhistory
3 .RData
4 .Ruserdata
5 .DS_Store
6 *.docx*
7 *.doc*
8 *.xlsx*
9 *.xls*
10 .Rproj|
11 ENT5920_DataManagementBiol.Rproj
12 rsconnect
13 *.pptx*
14 *.html
15
```