

# FH Cluster Guide

September 09, 2022



# Contents

<b>About this Course</b>	<b>5</b>
<b>Cluster 101</b>	<b>9</b>
<b>1 What is a Cluster?</b>	<b>9</b>
<b>2 Account Setup</b>	<b>11</b>
2.1 Check your HutchNet ID . . . . .	11
2.2 Contacting the SciComp Team . . . . .	11
<b>3 Terminal Setup</b>	<b>13</b>
3.1 What is a terminal? . . . . .	13
3.2 Windows Setup . . . . .	13
3.3 Mac Setup . . . . .	15
3.4 Linux Setup . . . . .	16
<b>4 Logging In</b>	<b>19</b>
4.1 What is SSH? . . . . .	19
4.2 Windows Login . . . . .	19
4.3 Mac Login . . . . .	21
4.4 Linux Login . . . . .	22
<b>5 Tour your space</b>	<b>23</b>
<b>6 Submit your first job</b>	<b>25</b>

<b>7 Interactive cluster</b>	<b>27</b>
<b>8 File upload and download</b>	<b>29</b>
8.1 Globus . . . . .	29
<b>Cluster 201</b>	<b>33</b>
<b>9 Resource Optimization</b>	<b>33</b>
<b>Appendix</b>	<b>37</b>
<b>10 Where to get help</b>	<b>37</b>
Submit a Ticket . . . . .	37
Visit the SciWiki . . . . .	37
<b>11 Provide Feedback</b>	<b>39</b>
<b>FAQ</b>	<b>41</b>
11.1 How can I manually install PuTTY? . . . . .	41
<b>About the Authors</b>	<b>43</b>
<b>References</b>	<b>47</b>

# About this Course

Our goal is to get you running on the Fred Hutch cluster quickly and efficiently with this quick-start guide. As a wise Drivers' Ed instructor once said, **you need to go slow to go fast!**

In this short course, you'll invest a bit of time now to save you time and frustration down the road. Follow along at **any time on your own schedule**. We hope that the following modules will help you take advantage of the powerful resources the Fred Hutch has to offer!



# Cluster 101





# Chapter 1

## What is a Cluster?

A computing cluster is a set of many computers networked together. Because there are many computers working together, the network is able to handle computationally expensive tasks, like genome assemblies or advanced algorithms. Imagine you're building a house. It would take a long time by yourself! It's much better to have many builders working together.

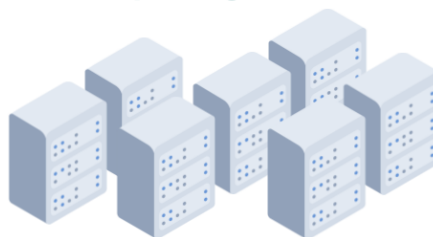
Now that we have a team of workers, the next challenge is task management. A home construction team will need a manager to help delegate tasks. Similarly, the computing cluster uses management software to prioritize tasks, delegate workers (resources), and check on progress. The Fred Hutch cluster uses a common management and scheduling tool called Slurm.

### Your Computer



1 node  
8 CPUs  
8 GB memory

### Fred Hutch Computing Cluster



600 nodes  
4,000 CPUs  
40 TB memory

CC-by hutchdatascience.org

How is the cluster different from a laptop or desktop? First, you might use an operating system like Windows or MacOS. The Fred Hutch server is a Linux

system. Second, because many people use the cluster for many tasks, there isn't a central screen and keyboard. You access the cluster remotely from your computer! We will talk more about how to connect to the cluster in a following chapter.

**Computing cluster**

A set of computers networked together to perform large tasks.

## Chapter 2

# Account Setup

You will need an account to log in to the cluster. This ensures that data stays protected.

### 2.1 Check your HutchNet ID

A HutchNet ID is the standard login you receive when you start working at the Hutch or are an official affiliate. You can use it to login to most resources at the Center (Desktop Computer, Employee Self Service, VPN, Webmail) and our Scientific Computing systems.

For example:

- my email is `jsmith3@fredhutch.org`.
- my HutchNet ID is `jsmith3`.

If one of your collaborators requires access to the Fred Hutch network you can submit a non-employee action form. Non-employee is a generic administrative term for affiliates, students, contractors, etc.

### 2.2 Contacting the SciComp Team

To use Scientific Computing clusters at Hutch, your HutchNet ID must be associated with a PI account.

The Scientific Computing Team (SciComp) tries to set some users up ahead of time. However, not everyone is set up automatically. Please fill out this Account Setup Form and we will ensure you are set up correctly!

Errors similar to “Invalid account or account/partition” typically indicate that the account hasn’t been set up by SciComp. This is a quick fix if you use the form above.

Now, let’s set up our Terminal!

## Chapter 3

# Terminal Setup

The next step is getting familiar with your Terminal. This is your portal to the cluster.

### 3.1 What is a terminal?

The Terminal is a command line interface. In other words, the Terminal is a software application that allows you to issue commands directly to your laptop or desktop computer. The Terminal is very useful because it allows you to run commands that don't have a graphical user interface (GUI). It can also connect you to computer networks, such as the Fred Hutch cluster! The Terminal setup is different depending on your operating system. Jump to the Windows, MacOS, or Linux sections below.

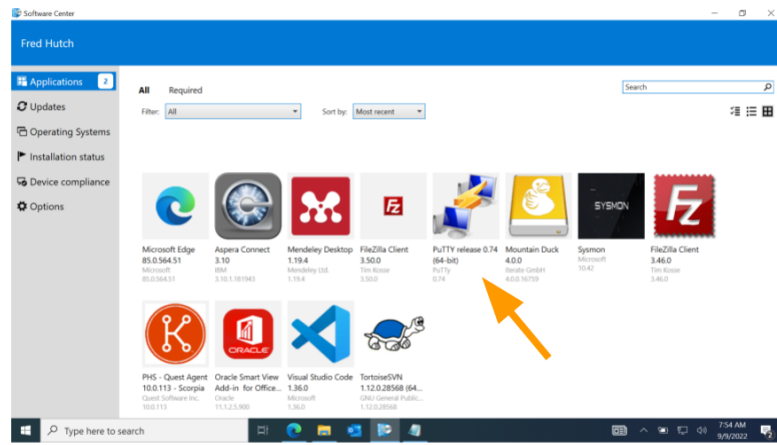
“Terminal” used to be synonymous with “computer”. With the creation of operating systems like Windows and MacOS, computers became much easier to use and exploded in popularity! Your colleagues are almost always referring to the command line application when they say “Terminal”.

### 3.2 Windows Setup

Click to view steps

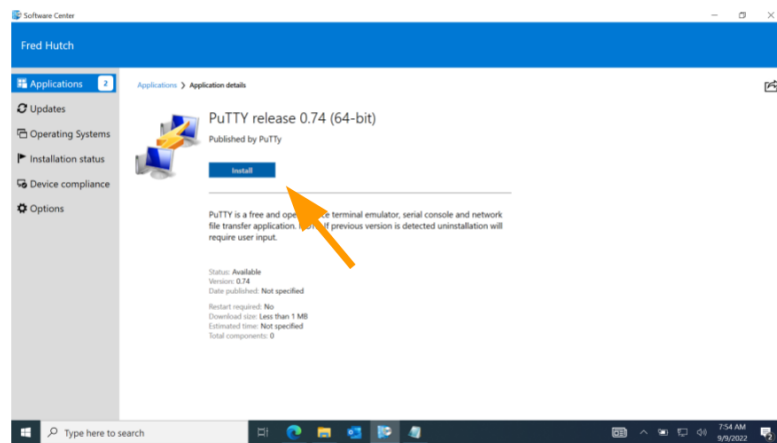
You will need to install a Terminal application called PuTTY to connect to the Fred Hutch Cluster.

1. You should then see PuTTY available in the Software Center. Click “Install” and go through the Setup Wizard.



Screenshot of the Software Center application.

CC-by hutchedatascience.org

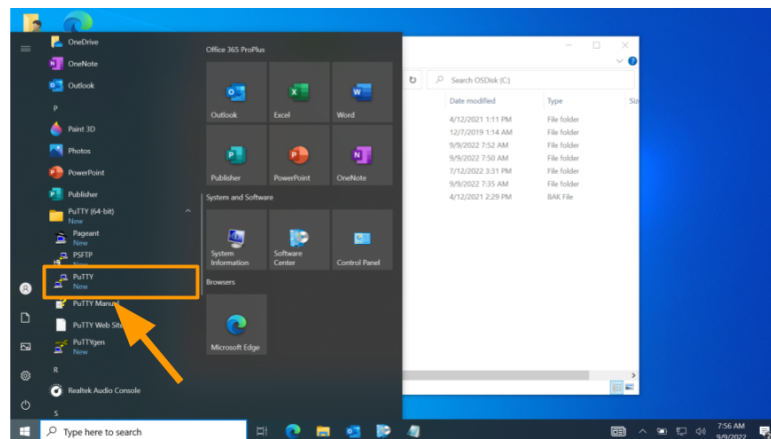


Screenshot of the Software Center PuTTY installation page.

CC-by hutchedatascience.org

You can also install PuTTY manually if you don't see it in the Software Center.

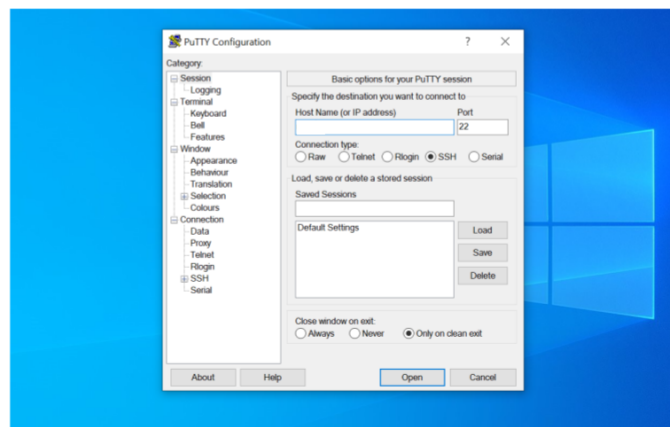
1. PuTTY should now be available in your applications. Click on PuTTY to open.



Screenshot of the Windows application menu.

CC-by hutchdatascience.org

2. You should now see the PuTTY Configuration menu.



Screenshot of the PuTTY Configuration menu.

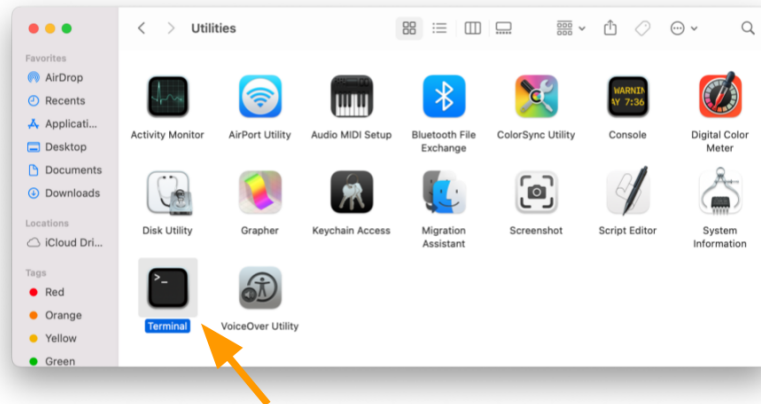
CC-by hutchdatascience.org

## 3.3 Mac Setup

Click to view steps

Mac machines come with a Terminal installed.

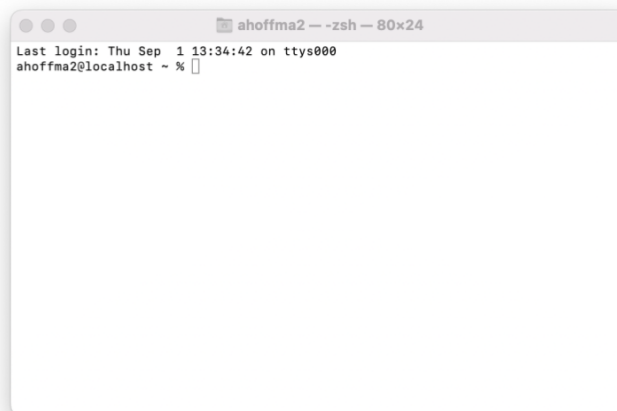
1. Go to Finder > Applications > Utilities > Terminal and double-click.



Screenshot of the Mac Utilities directory.

CC-by hutchdatascience.org

2. Your Terminal should look like this:



Screenshot of the Mac Terminal.

CC-by hutchdatascience.org

## 3.4 Linux Setup

Click to view steps

The commonly used Linux distribution, Ubuntu, already comes with a Terminal installed.

1. Press `ctrl + alt + T`. Your open Terminal window should look like this:

[SCREENSHOT]



1. Update the Terminal and prepare it for connecting to the cluster by running:

```
sudo apt install openssh-server
```

Enter your password and enter Y when prompted.



## Chapter 4

# Logging In

Now that you have your Terminal application ready, you want to connect to the cluster. You will do this using a method called SSH, which stands for “Secure SHell”.

### 4.1 What is SSH?

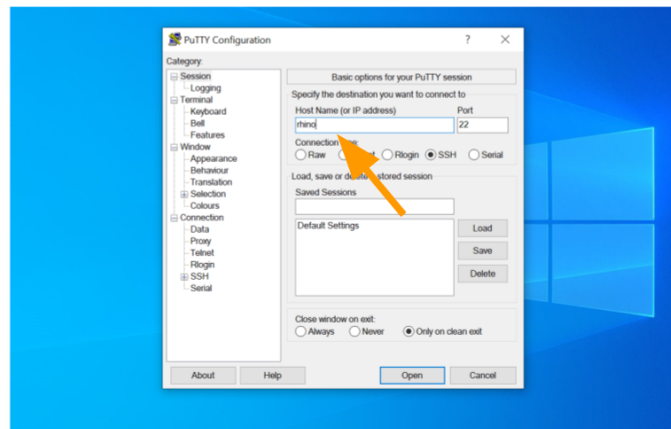
SSH is a secure way to remotely connect to another computer or network of computers. In other words, SSH helps us protect your data and the data on the Fred Hutch cluster through authentication.

Before moving on, you will need to connect to the Fred Hutch VPN. This is the first layer of security. The next set of steps are specific to your operating system.

### 4.2 Windows Login

Click to view steps

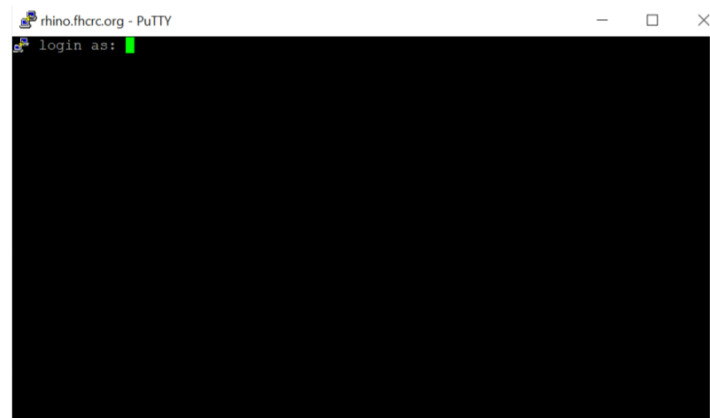
1. Go to the PuTTY Configuration menu. Under “Host Name” type **rhino** and click “Open”.



Screenshot of the PuTTY Configuration menu.

CC-by hutchdatascience.org

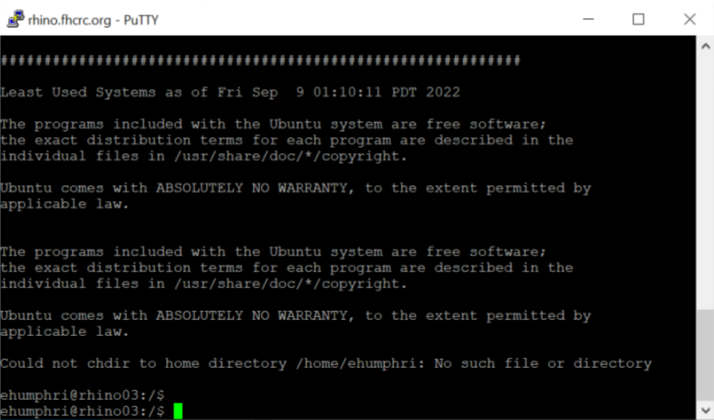
2. You will be prompted to login. Type in your HutchNetID (e.g., `jsmith3`).



Screenshot of the PuTTY login prompt for rhino.

CC-by hutchdatascience.org

3. Enter your password. No\* or symbols will show up, so type it in carefully!
4. You are now logged in! There should be a login message, with your name at the bottom.



```

rhino.fhcr.org - PuTTY
*****
Least Used Systems as of Fri Sep 9 01:10:11 EDT 2022

The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

Could not chdir to home directory /home/ehumphri: No such file or directory
ehumphri@rhino03:/$
ehumphri@rhino03:/$

```

Screenshot of the Fred Hutch cluster login message. CC-by hutchdatascience.org

Congratulations! You are now logged in to the Fred Hutch cluster!

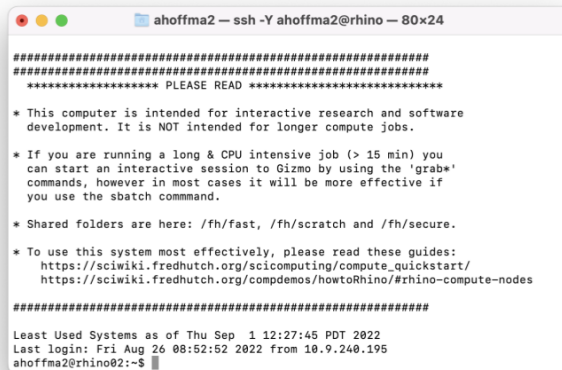
## 4.3 Mac Login

Click to view steps

1. Type the following commands, substituting in your HutchNet ID:

```
ssh -Y HutchID@rhino
```

2. You will see a message that looks like `The authenticity of host 'rhino (XXX.XXX.XX.XX)' can't be established.` Type in `yes` and hit enter.
3. Enter your password. No\* or symbols will show up, so type it in carefully!
4. You are now logged in! There should be a login message, with your name at the bottom.



```
ahoffma2 — ssh -Y ahoffma2@rhino — 80x24

#####
***** PLEASE READ *****

* This computer is intended for interactive research and software
  development. It is NOT intended for longer compute jobs.

* If you are running a long & CPU intensive job (> 15 min) you
  can start an interactive session to Gizmo by using the 'grab*'
  commands, however in most cases it will be more effective if
  you use the sbatch command.

* Shared folders are here: /fh/fast, /fh/scratch and /fh/secure.

* To use this system most effectively, please read these guides:
  https://scwiki.fredhutch.org/scicomputing/compute_quickstart/
  https://scwiki.fredhutch.org/compdemos/howtoRhino/#rhino-compute-nodes

#####

Least Used Systems as of Thu Sep  1 12:27:45 PDT 2022
Last login: Fri Aug 26 08:52:52 2022 from 10.9.240.195
ahoffma2@rhino02:~$
```

Screenshot of the Fred Hutch cluster login message.

CC-by hutchdatascience.org

Congratulations! You are now logged in to the Fred Hutch cluster!

## 4.4 Linux Login

Click to view steps

Congratulations! You are now logged in to the Fred Hutch cluster!

## Chapter 5

# Tour your space





## Chapter 6

# Submit your first job



## Chapter 7

# Interactive cluster



## Chapter 8

# File upload and download

### 8.1 Globus



## Cluster 201





## Chapter 9

# Resource Optimization

It is best practice to...



# Appendix



## Chapter 10

# Where to get help

We want to help! Here are some ways you can get help for your work on the cluster.

### Submit a Ticket

Submitting a good ticket helps the SciComp Team address your needs quickly and efficiently. We suggest you submit the following in a ticket:

- 1.
- 2.
- 3.

### Visit the SciWiki

The SciWiki Scientific Computing page is full of useful tips and guides.



## Chapter 11

# Provide Feedback

Please submit an issue at our GitHub repo. You can also click the edit button on the top of the page in question.





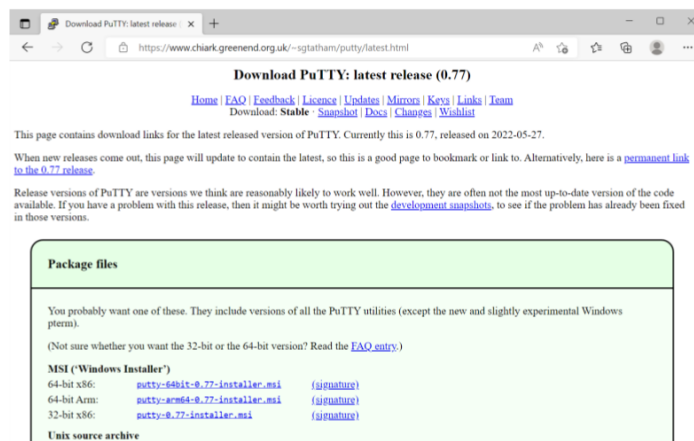
# FAQ

Here are some issues you might encounter.

## 11.1 How can I manually install PuTTY?

Click to view steps

1. Click here to install the latest version of PuTTY. You will choose the 64-bit x86 installation with few exceptions.



Screenshot of the PuTTY download page.

CC-by hutchdatascience.org

2. Click through to install via the Setup Wizard.



# About the Authors

These credits are based on our course contributors table guidelines.

Credits	Names
<b>Pedagogy</b>	
Lead Content Instructor(s)	FirstName LastName
Lecturer(s) (include chapter name/link in parentheses if only for specific chapters) - make new line if more than one chapter involved	Delivered the course in some way - video or audio
Content Author(s) (include chapter name/link in parentheses if only for specific chapters) - make new line if more than one chapter involved	If any other authors besides lead instructor
Content Contributor(s) (include section name/link in parentheses) - make new line if more than one section involved	Wrote less than a chapter
Content Editor(s)/Reviewer(s)	Checked your content
Content Director(s)	Helped guide the content direction
Content Consultants (include chapter name/link in parentheses or word “General”) - make new line if more than one chapter involved	Gave high level advice on content
Acknowledgments	Gave small assistance to content but not to the level of consulting
<b>Production</b>	
Content Publisher(s)	Helped with publishing platform
Content Publishing Reviewer(s)	Reviewed overall content and aesthetics on publishing platform

Credits	Names
<b>Technical</b>	
Course Publishing Engineer(s)	Helped with the code for the technical aspects related to the specific course generation
Template Publishing Engineers	Candace Savonen, Carrie Wright
Publishing Maintenance Engineer	Candace Savonen
Technical Publishing Stylists	Carrie Wright, Candace Savonen
Package Developers (ottrpal)	Candace Savonen, John Muschelli, Carrie Wright
<b>Art and Design</b>	
Illustrator(s)	Created graphics for the course
Figure Artist(s)	Created figures/plots for course
Videographer(s)	Filmed videos
Videography Editor(s)	Edited film
Audiographer(s)	Recorded audio
Audiography Editor(s)	Edited audio recordings
<b>Funding</b>	
Funder(s)	Institution/individual who funded course including grant number
Funding Staff	Staff members who help with funding

```
## - Session info -----
## setting value
## version R version 4.0.2 (2020-06-22)
## os      Ubuntu 20.04.3 LTS
## system  x86_64, linux-gnu
## ui      X11
## language (EN)
## collate en_US.UTF-8
## ctype   en_US.UTF-8
## tz      Etc/UTC
## date    2022-09-09
##
## - Packages -----
## package      * version      date      lib source
## assertthat    0.2.1        2019-03-21 [1] RSPM (R 4.0.3)
## bookdown      0.24         2022-02-15 [1] Github (rstudio/bookdown@88bc4ea)
```

```
## callr          3.4.4      2020-09-07 [1] RSPM (R 4.0.2)
## cli            2.0.2      2020-02-28 [1] RSPM (R 4.0.0)
## crayon         1.3.4      2017-09-16 [1] RSPM (R 4.0.0)
## desc           1.2.0      2018-05-01 [1] RSPM (R 4.0.3)
## devtools       2.3.2      2020-09-18 [1] RSPM (R 4.0.3)
## digest         0.6.25     2020-02-23 [1] RSPM (R 4.0.0)
## ellipsis       0.3.1      2020-05-15 [1] RSPM (R 4.0.3)
## evaluate       0.14       2019-05-28 [1] RSPM (R 4.0.3)
## fansi          0.4.1      2020-01-08 [1] RSPM (R 4.0.0)
## fs             1.5.0      2020-07-31 [1] RSPM (R 4.0.3)
## glue           1.6.1      2022-01-22 [1] CRAN (R 4.0.2)
## htmltools      0.5.0      2020-06-16 [1] RSPM (R 4.0.1)
## knitr          1.33       2022-02-15 [1] Github (yihui/knitr@a1052d1)
## lifecycle      1.0.0      2021-02-15 [1] CRAN (R 4.0.2)
## magrittr       2.0.2      2022-01-26 [1] CRAN (R 4.0.2)
## memoise        1.1.0      2017-04-21 [1] RSPM (R 4.0.0)
## pkgbuild       1.1.0      2020-07-13 [1] RSPM (R 4.0.2)
## pkgload        1.1.0      2020-05-29 [1] RSPM (R 4.0.3)
## prettyunits    1.1.1      2020-01-24 [1] RSPM (R 4.0.3)
## processx       3.4.4      2020-09-03 [1] RSPM (R 4.0.2)
## ps             1.3.4      2020-08-11 [1] RSPM (R 4.0.2)
## purrr          0.3.4      2020-04-17 [1] RSPM (R 4.0.3)
## R6             2.4.1      2019-11-12 [1] RSPM (R 4.0.0)
## remotes        2.2.0      2020-07-21 [1] RSPM (R 4.0.3)
## rlang          0.4.10     2022-02-15 [1] Github (r-lib/rlang@f0c9be5)
## rmarkdown      2.10       2022-02-15 [1] Github (rstudio/rmarkdown@02d3c25)
## rprojroot      2.0.2      2020-11-15 [1] CRAN (R 4.0.2)
## sessioninfo    1.1.1      2018-11-05 [1] RSPM (R 4.0.3)
## stringi        1.5.3      2020-09-09 [1] RSPM (R 4.0.3)
## stringr        1.4.0      2019-02-10 [1] RSPM (R 4.0.3)
## testthat       3.0.1      2022-02-15 [1] Github (R-lib/testthat@e99155a)
## usethis        2.1.5.9000 2022-02-15 [1] Github (r-lib/usethis@57b109a)
## withr          2.3.0      2020-09-22 [1] RSPM (R 4.0.2)
## xfun           0.32       2022-08-10 [1] CRAN (R 4.0.2)
## yaml           2.2.1      2020-02-01 [1] RSPM (R 4.0.3)
##
## [1] /usr/local/lib/R/site-library
## [2] /usr/local/lib/R/library
```



## References