



# PROJECTS' WORKFLOW FOR REPRODUCIBILITY AND REPLICABILITY USING R

## PART 2 – COMPUTATIONAL REPRODUCIBILITY

Second Rostock Open Science Workshop, MPIDR, Rostock  
17-18 March 2025

*Egor Kotov & Jonas Schöley*

# BUILDING BLOCKS OF COMPUTATIONAL REPRODUCIBILITY



Lock R packages and  
specific versions



# BUILDING BLOCKS OF COMPUTATIONAL REPRODUCIBILITY



Lock R packages and  
specific versions



Reproducible  
analysis workflow



# BUILDING BLOCKS OF COMPUTATIONAL REPRODUCIBILITY



Lock R packages and specific versions



Reproducible analysis workflow



Complete computational environment



# BUILDING BLOCKS OF COMPUTATIONAL REPRODUCIBILITY



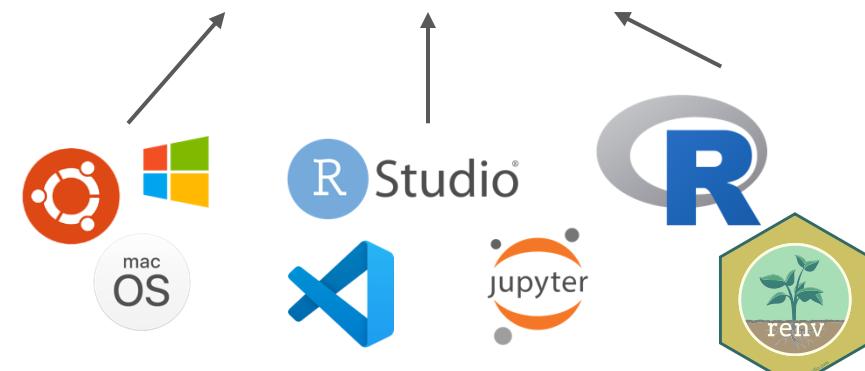
Lock R packages and specific versions



Reproducible analysis workflow



Complete computational environment





# renv: PACKAGE VERSION CONTROL

You are here

Lock R packages and specific versions



Reproducible analysis workflow



Complete computational environment



Logos by respective copyright holders



# renv: PACKAGE VERSION CONTROL



**/bin/R/library** (*base R packages*)



## renv: PACKAGE VERSION CONTROL

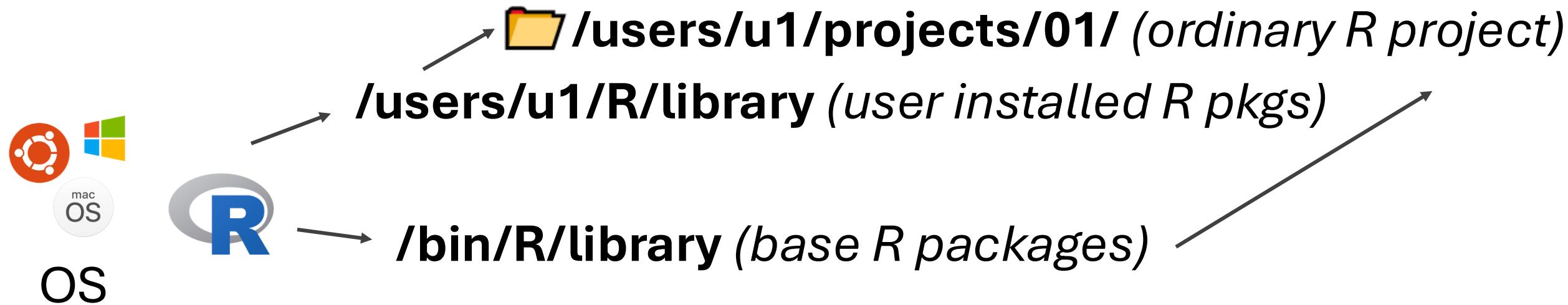


**/users/u1/R/library** (*user installed R pkgs*)

**/bin/R/library** (*base R packages*)

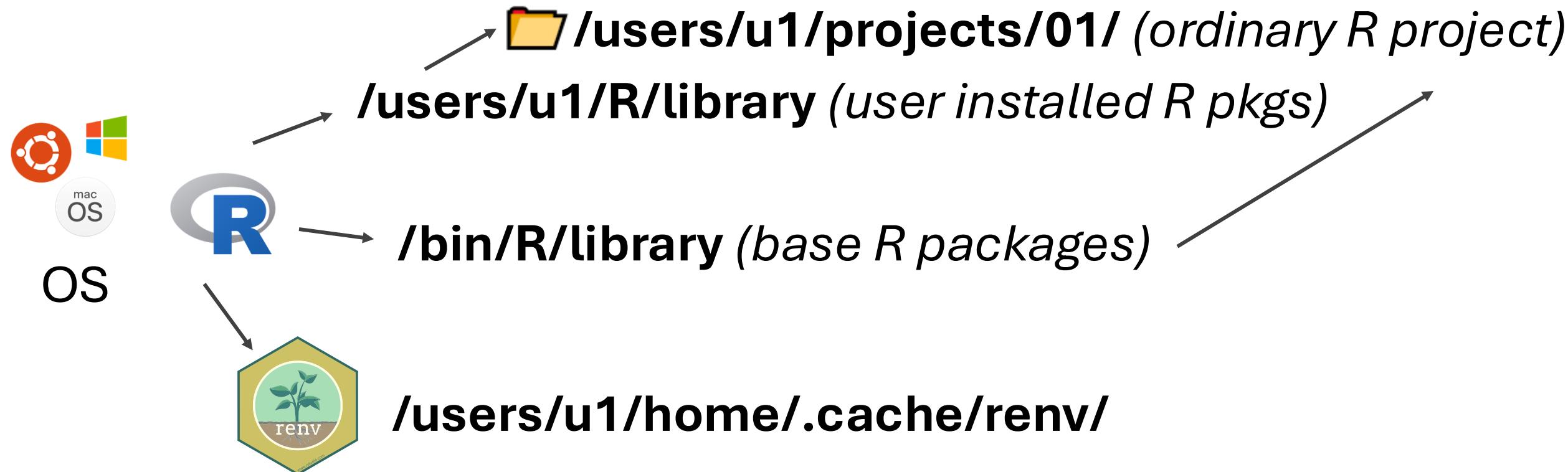


## renv: PACKAGE VERSION CONTROL



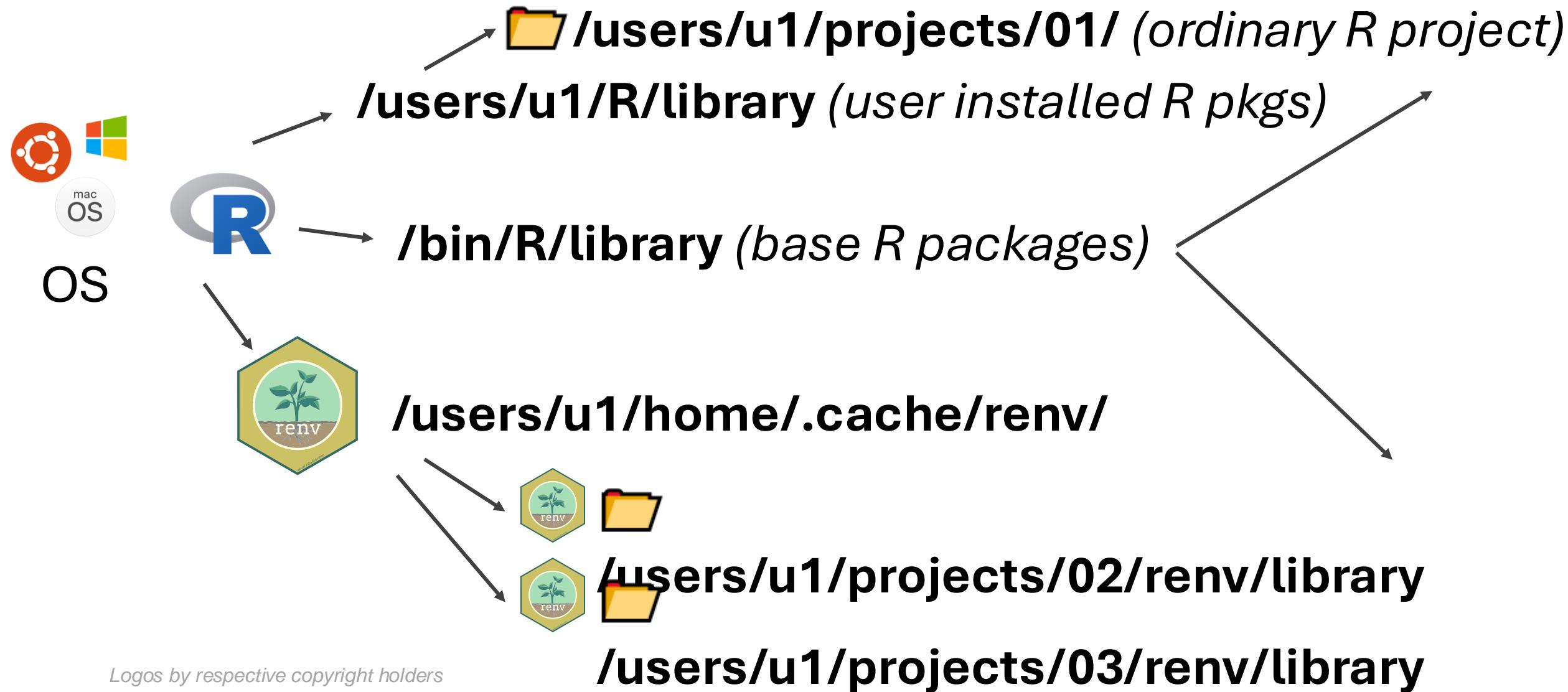


## renv: PACKAGE VERSION CONTROL

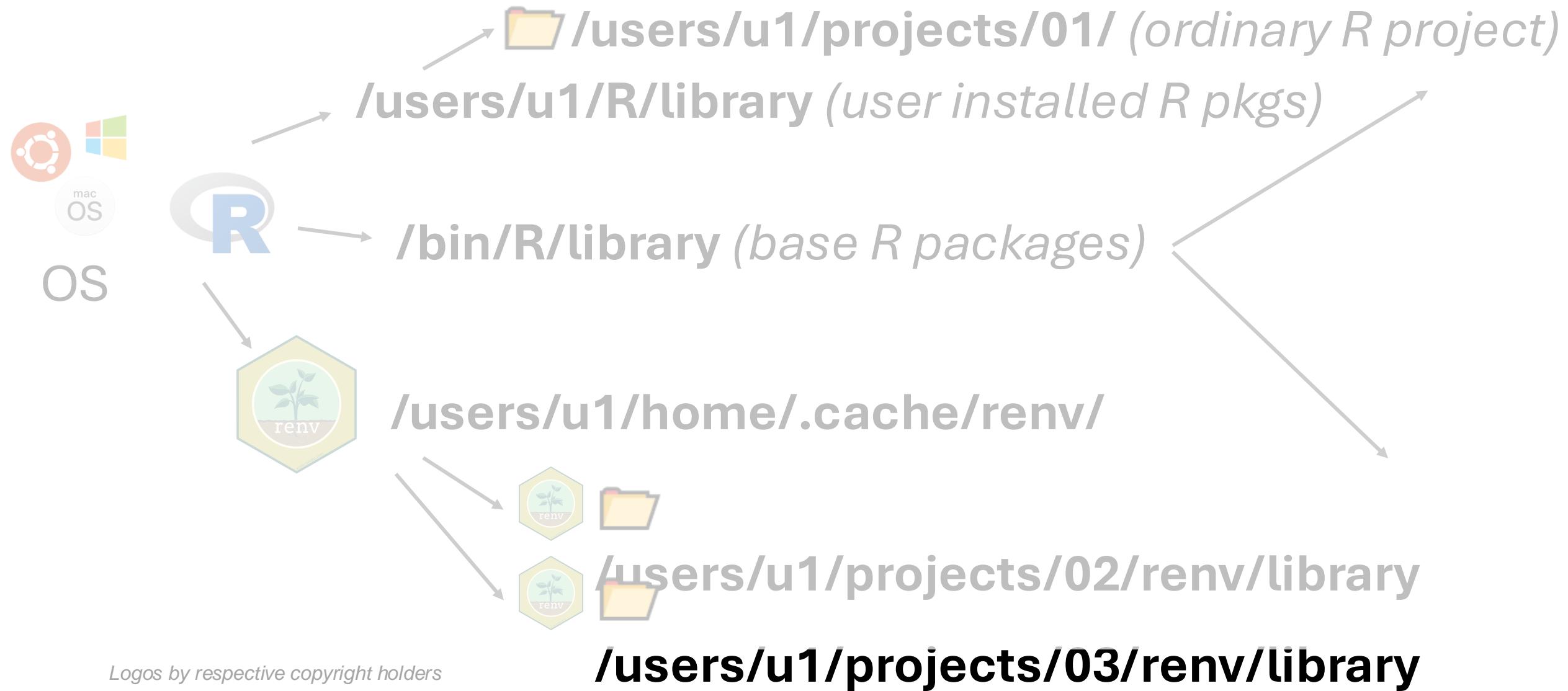




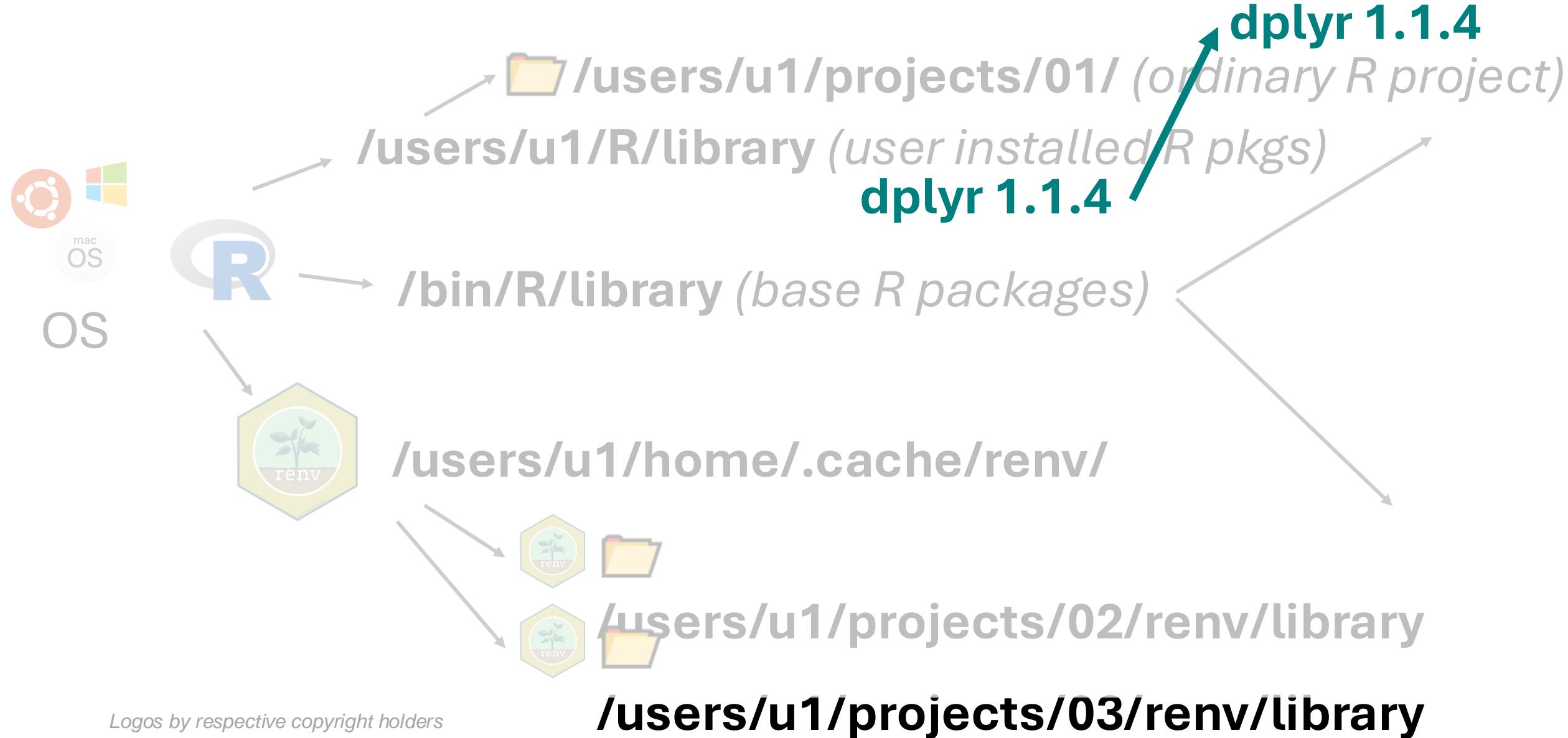
# renv: PACKAGE VERSION CONTROL



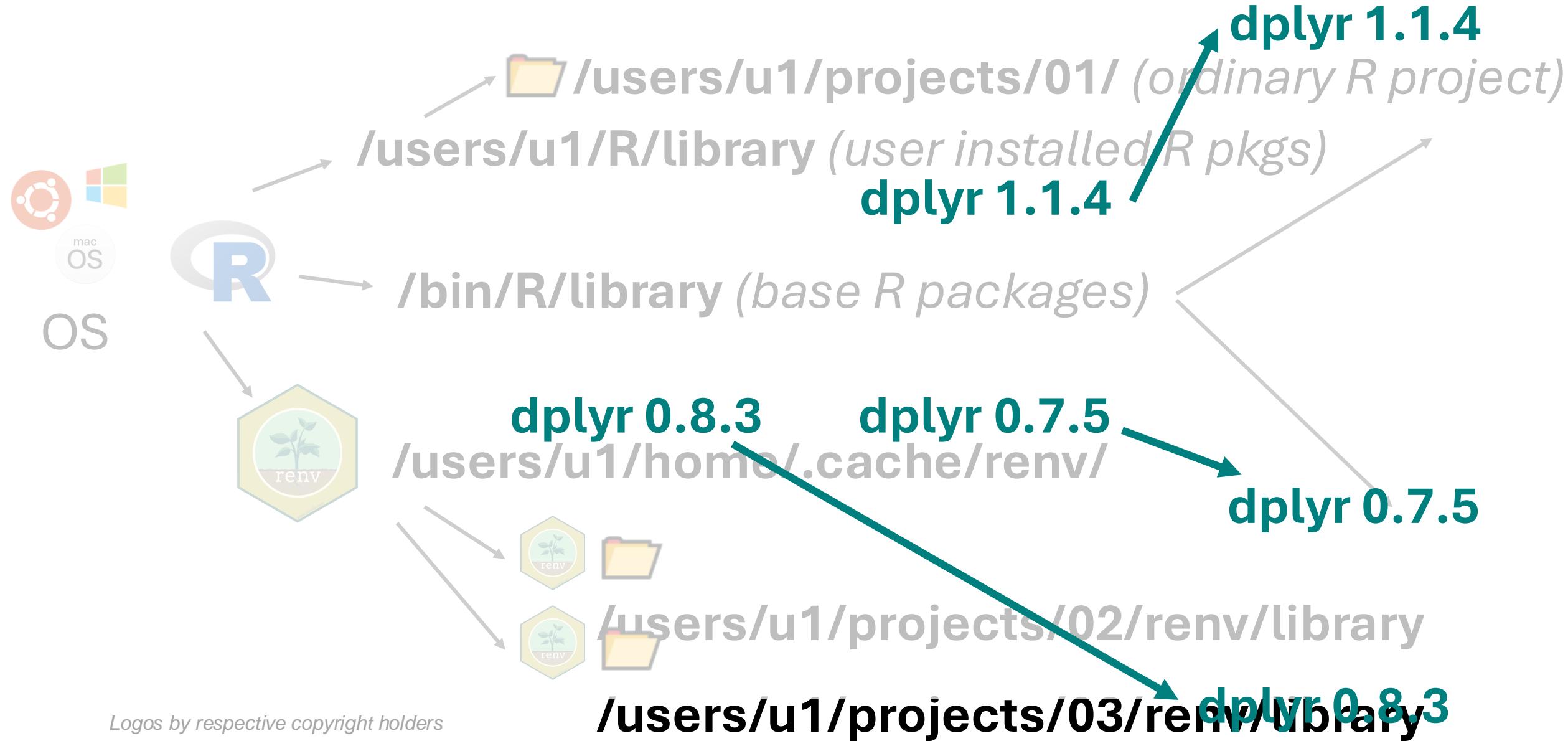
# renv: PACKAGE VERSION CONTROL



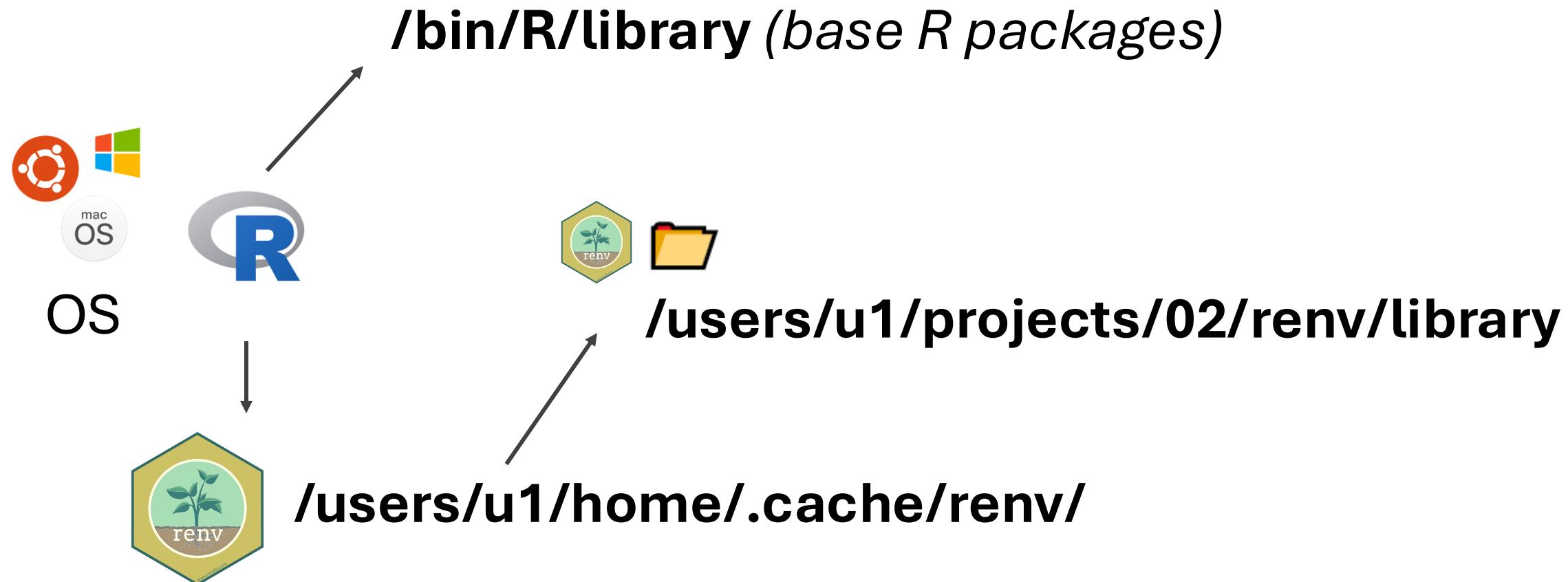
# renv: PACKAGE VERSION CONTROL



# renv: PACKAGE VERSION CONTROL

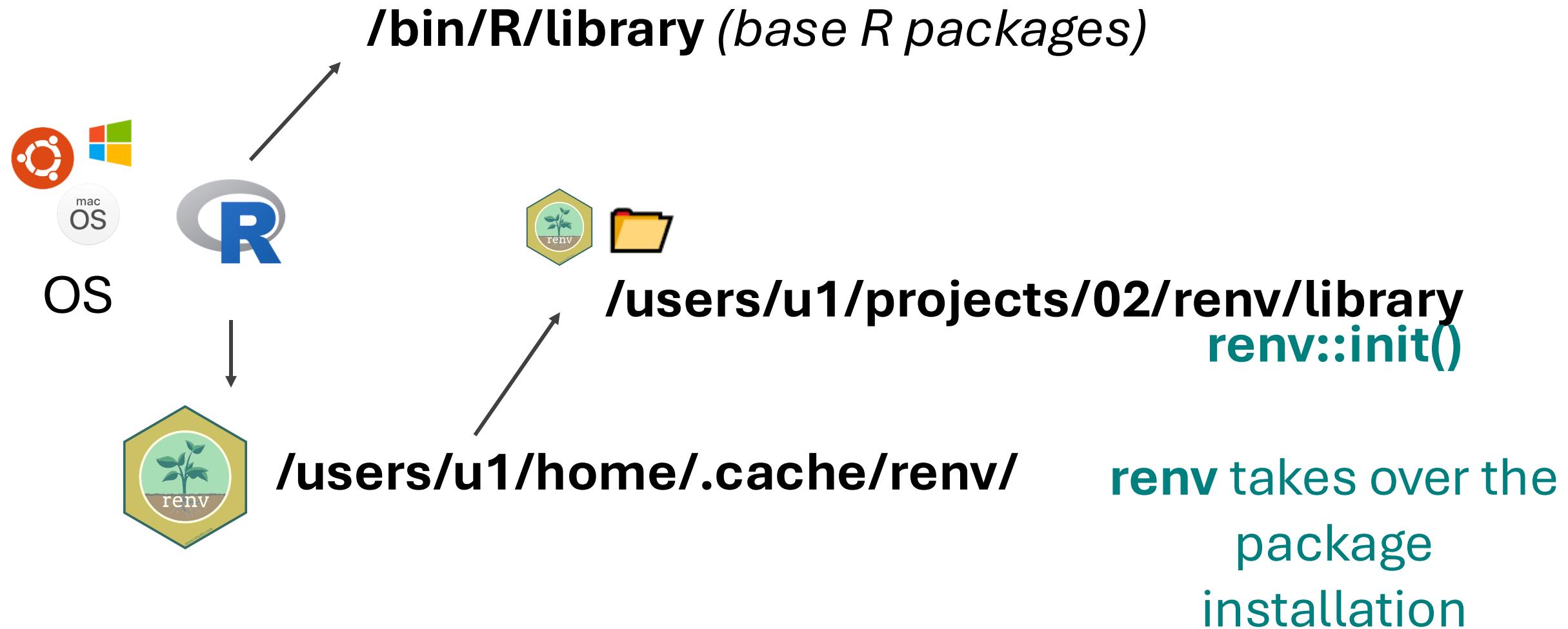


# renv: PACKAGE VERSION CONTROL

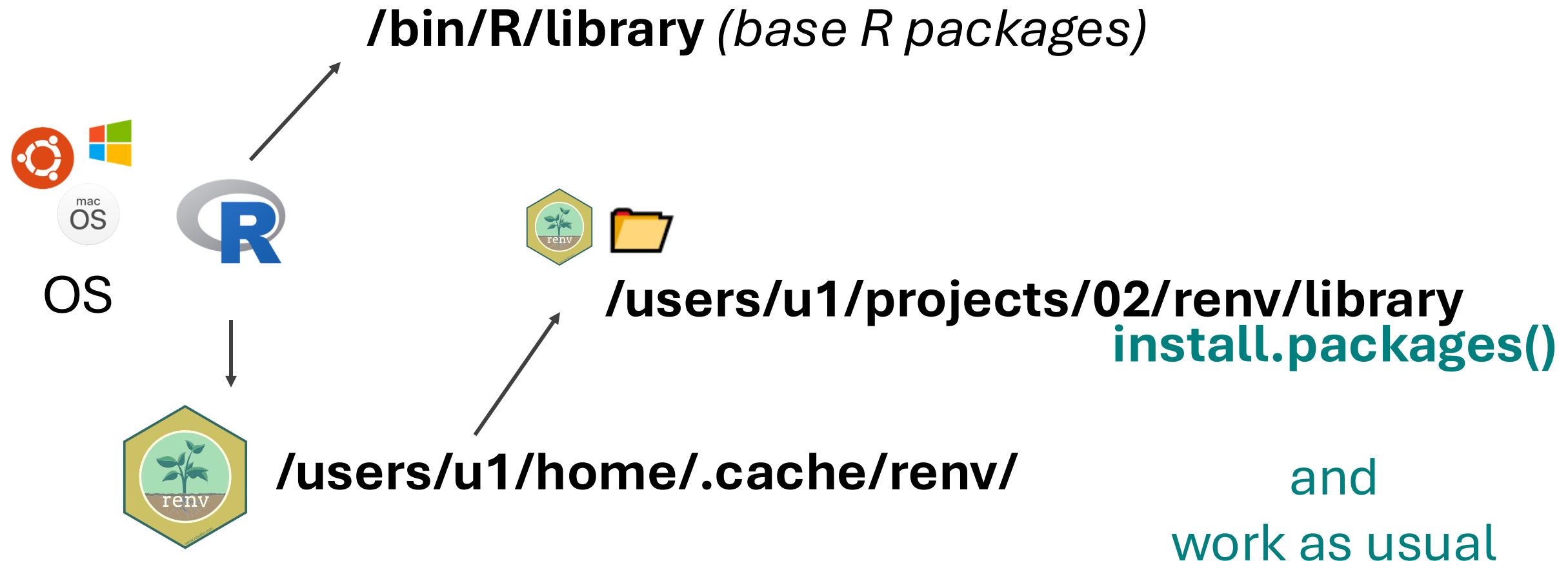




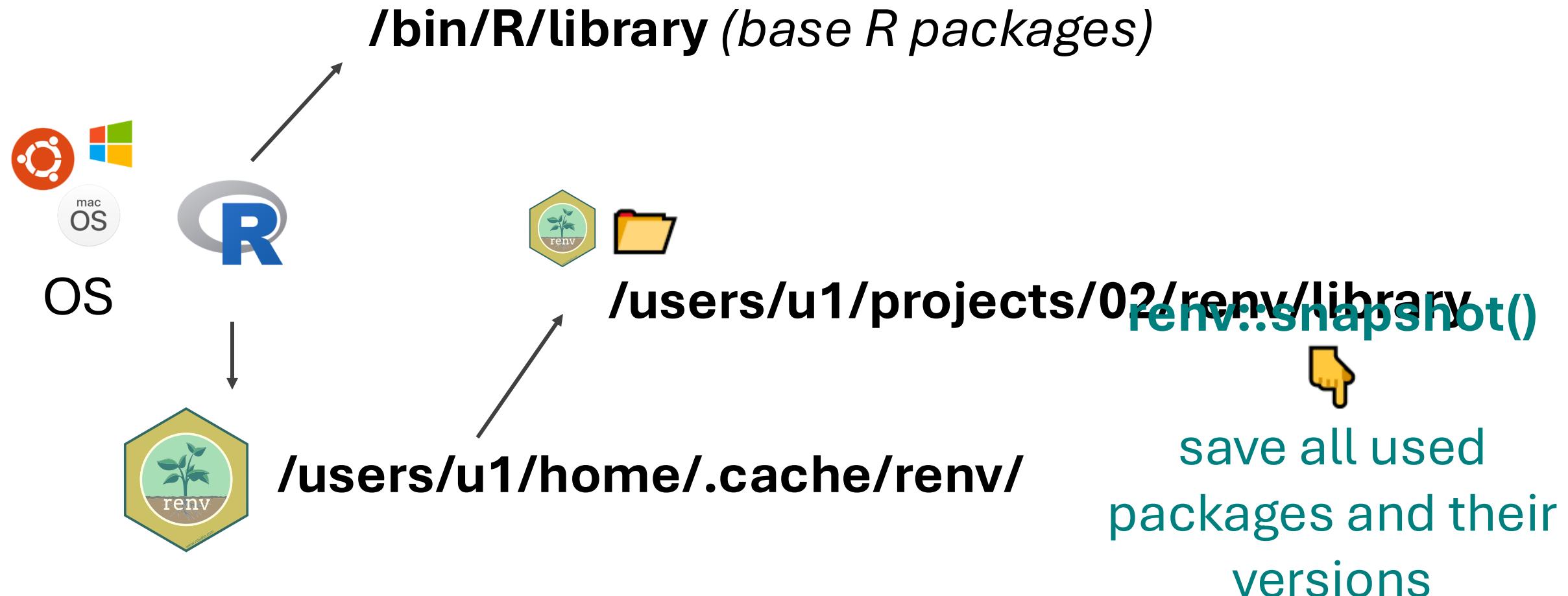
# renv: PACKAGE VERSION CONTROL



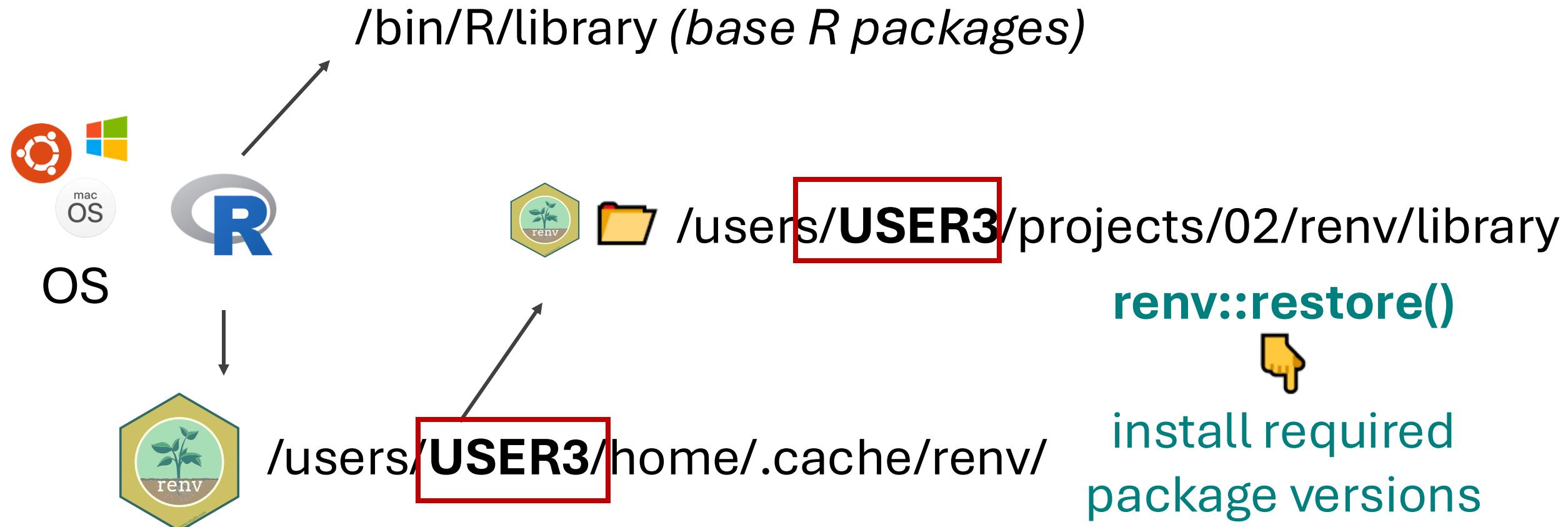
# renv: PACKAGE VERSION CONTROL



# renv: PACKAGE VERSION CONTROL

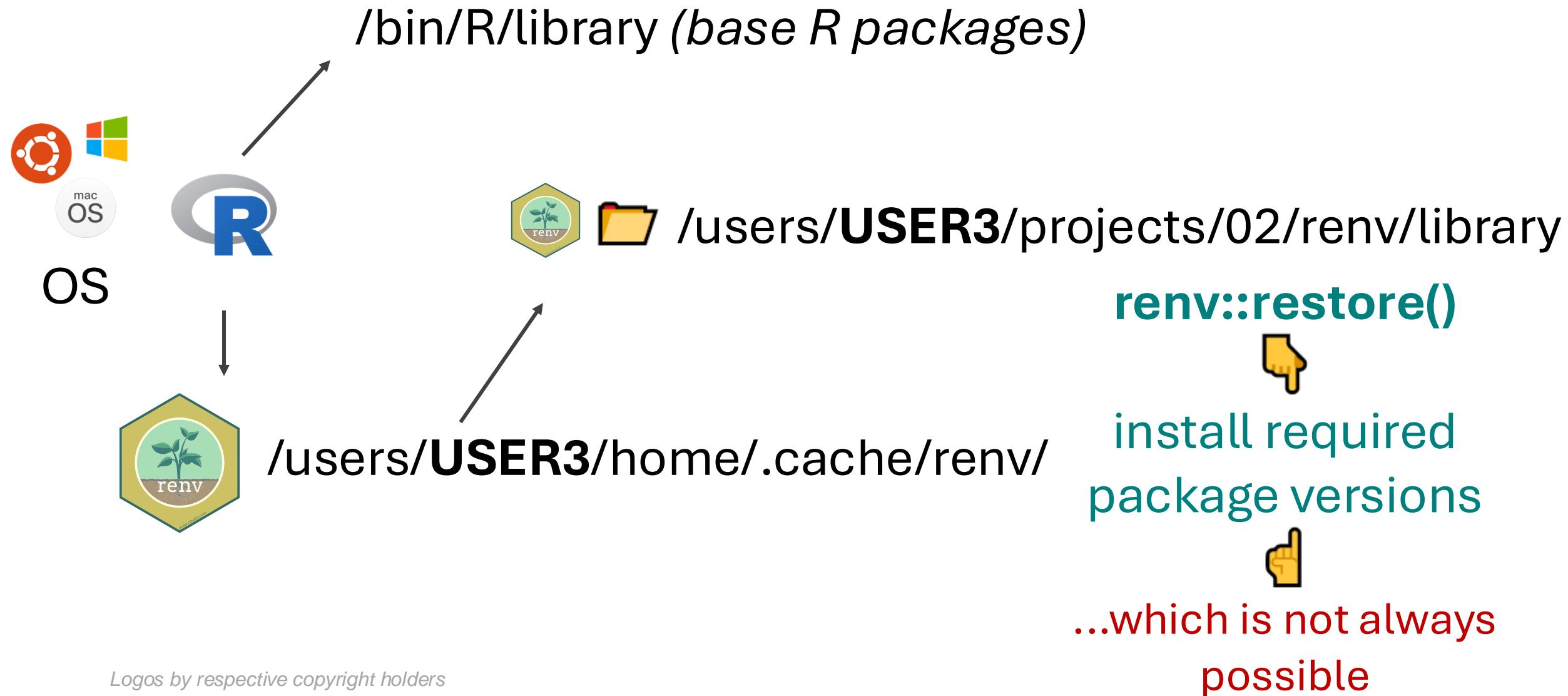


# renv: PACKAGE VERSION CONTROL





# renv: PACKAGE VERSION CONTROL





# targets : ANALYSIS WORKFLOW

Lock R packages and specific versions



You are here

Reproducible analysis workflow



Complete computational environment



SINGULARITYCE



APPTAINER



mac OS



R Studio



Logos by respective copyright holders



# targets: ANALYSIS WORKFLOW

```
tar_target(mtcars_data,  
          mtcars) ,  
  
tar_target(mpg_mean,  
          mean(mtcars_data$mpg)) ,  
  
tar_target(plot_mpg,  
          ggplot2::qplot(mpg,data=mtcars_data))
```



# targets: ANALYSIS WORKFLOW

```
tar_target(mtcars_data,  
          mtcars),  function saves result to
```

```
tar_target(mpg_mean,  
          mean(mtcars_data$mpg)),  function saves result to
```

```
tar_target(plot_mpg,  
          ggplot2::qplot(mpg,data=mtcars_data))  
  
 function saves result to
```



# targets: ANALYSIS WORKFLOW

```
library(targets)
tar_dir({
tar_script({
  library(targets)
  list(
    tar_target(mtcars_data,
               mtcars) ,

    tar_target(mpg_mean,
               mean(mtcars_data$mpg)) ,

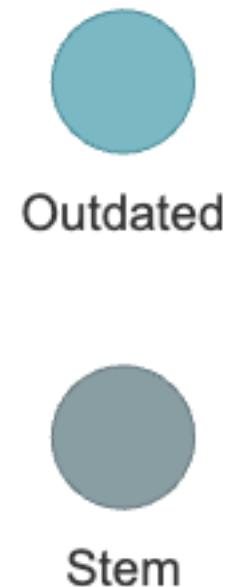
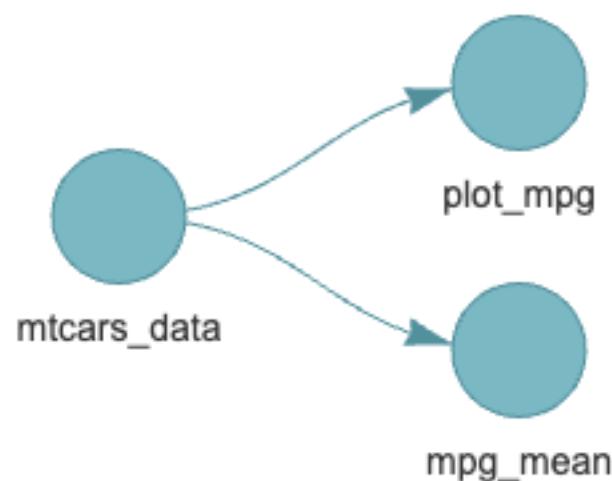
    tar_target(plot_mpg,
               ggplot2::qplot(mpg,data=mtcars_data)) 

  ) }, ask = FALSE)
tar_visnetwork()
})
```



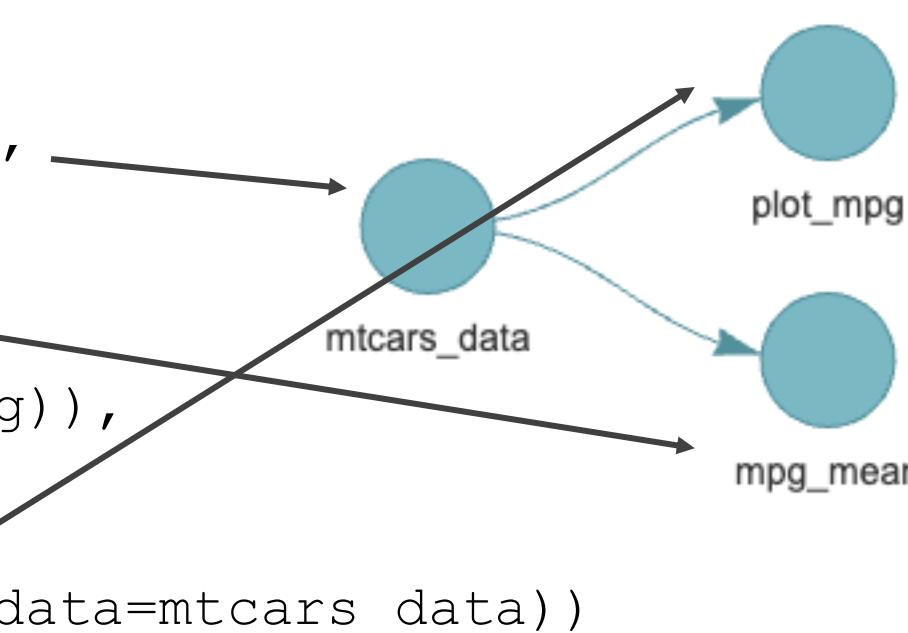
# targets: ANALYSIS WORKFLOW

```
library(targets)
tar_dir({
tar_script({
  library(targets)
  list(
    tar_target(mtcars_data,
               mtcars),
    tar_target(mpg_mean,
               mean(mtcars_data$mpg)),
    tar_target(plot_mpg,
               ggplot2::qplot(mpg,data=mtcars_data))
  )}, ask = FALSE)
tar_visnetwork()
})
```



# targets: ANALYSIS WORKFLOW

```
library(targets)
tar_dir({
tar_script({
  library(targets)
  list(
    tar_target(mtcars_data,
               mtcars),
    tar_target(mpg_mean,
               mean(mtcars_data$mpg)),
    tar_target(plot_mpg,
               ggplot2::qplot(mpg, data=mtcars_data))
  ) }, ask = FALSE)
tar_visnetwork()
})
```



Outdated



Stem

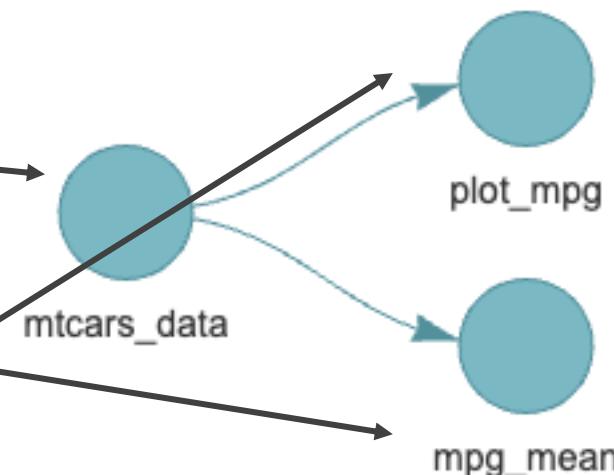


# targets: ANALYSIS WORKFLOW

```
tar_target(mtcars_data,  
          mtcars),
```

```
tar_target(mpg_mean,  
          mean(mtcars_data$mpg)),
```

```
tar_target(plot_mpg,  
          ggplot2::qplot(mpg, data=mtcars_data))
```



Outdated

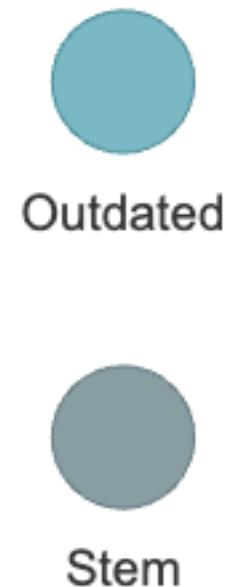
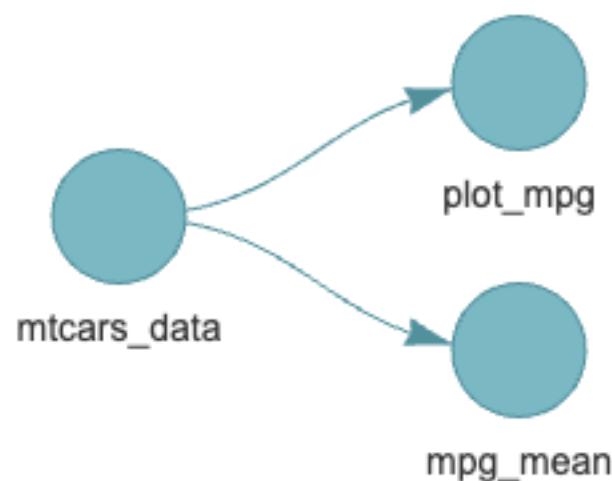


Stem



# targets: ANALYSIS WORKFLOW

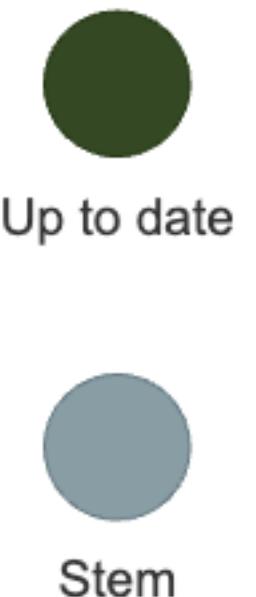
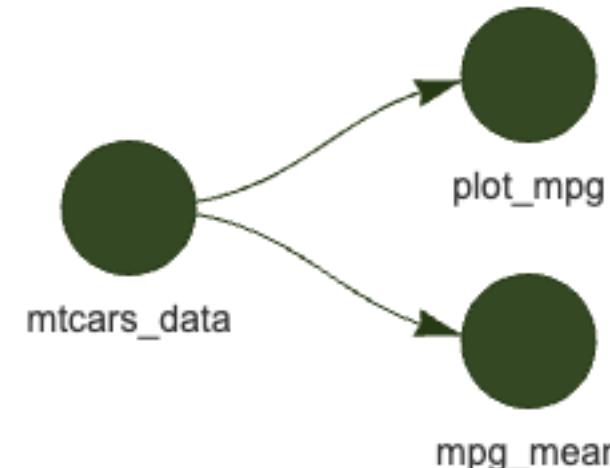
```
library(targets)
tar_dir({
tar_script({
  library(targets)
  list(
    tar_target(mtcars_data,
               mtcars),
    tar_target(mpg_mean,
               mean(mtcars_data$mpg)),
    tar_target(plot_mpg,
               ggplot2::qplot(mpg,data=mtcars_data))
  )}, ask = FALSE)
tar_visnetwork()
})
```



# targets: ANALYSIS WORKFLOW

```
library(targets)
tar_dir({
  tar_script({
    library(targets)
    list(
      tar_target(mtcars_data,
                 mtcars),
      tar_target(mpg_mean,
                 mean(mtcars_data$mpg)),
      tar_target(plot_mpg,
                 ggplot2::qplot(mpg, data=mtcars_data))
    ) }, ask = FALSE)
  tar_make(); tar_visnetwork()
})
```

**targets::tar\_make()** ➡️ run all steps





# targets: ANALYSIS WORKFLOW

```
library(targets)
tar_dir({
tar_script({
  library(targets)
  list(
    tar_target(mtcars_data,
               mtcars),
    tar_target(mpg_mean,
               mean(mtcars_data$mpg)),
    tar_target(plot_mpg,
               ggplot2::qplot(mpg, data=mtcars_data))
  ) }, ask = FALSE)
tar_make(); tar_read(plot_mpg)
})
```

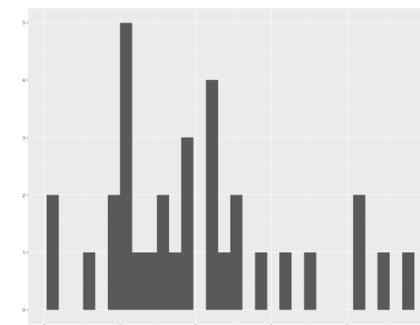
**targets::tar\_read()** ➡ view step results



Up to date



Stem





# targets: PROJECT STRUCTURE

```
library(targets)

list(
  tar_target(output1,
    function_1() ) ,

  tar_target(output2,
    function_2(input1 = output1) ) ,

  tar_target(output3,
    function_3(input1 = output1,
      input2 = output2) ) ,
```



# targets: PROJECT STRUCTURE

```
library(targets)

list(
  tar_target(output1,
    function_1() ) ,

  tar_target(output2,
    function_2(input1 = output1) ) ,

  tar_target(output3,
    function_3(input1 = output1,
      input2 = output2) ) ,
)
```

→ /rpoject01  
→ - **\_targets.R**  
→ - R/  
→ - 01\_scripts.R  
→ - 02\_scripts.R  
→ - ...  
→ - \_targets/  
→ - objects/



# targets: PROJECT STRUCTURE

```
library(targets)                                /rpoject01
list(                                         - _targets.R
  tar_target(output1,                         - R/
    function_1() ),                         - 01_scripts.R
  tar_target(output2,                         - 02_scripts.R
    function_2(input1 = output1) ),           - ...
  tar_target(output3,                         - _targets/
    function_3(input1 = output1,             - objects/
      input 2 = output2) ),
```

)



# targets: PROJECT STRUCTURE

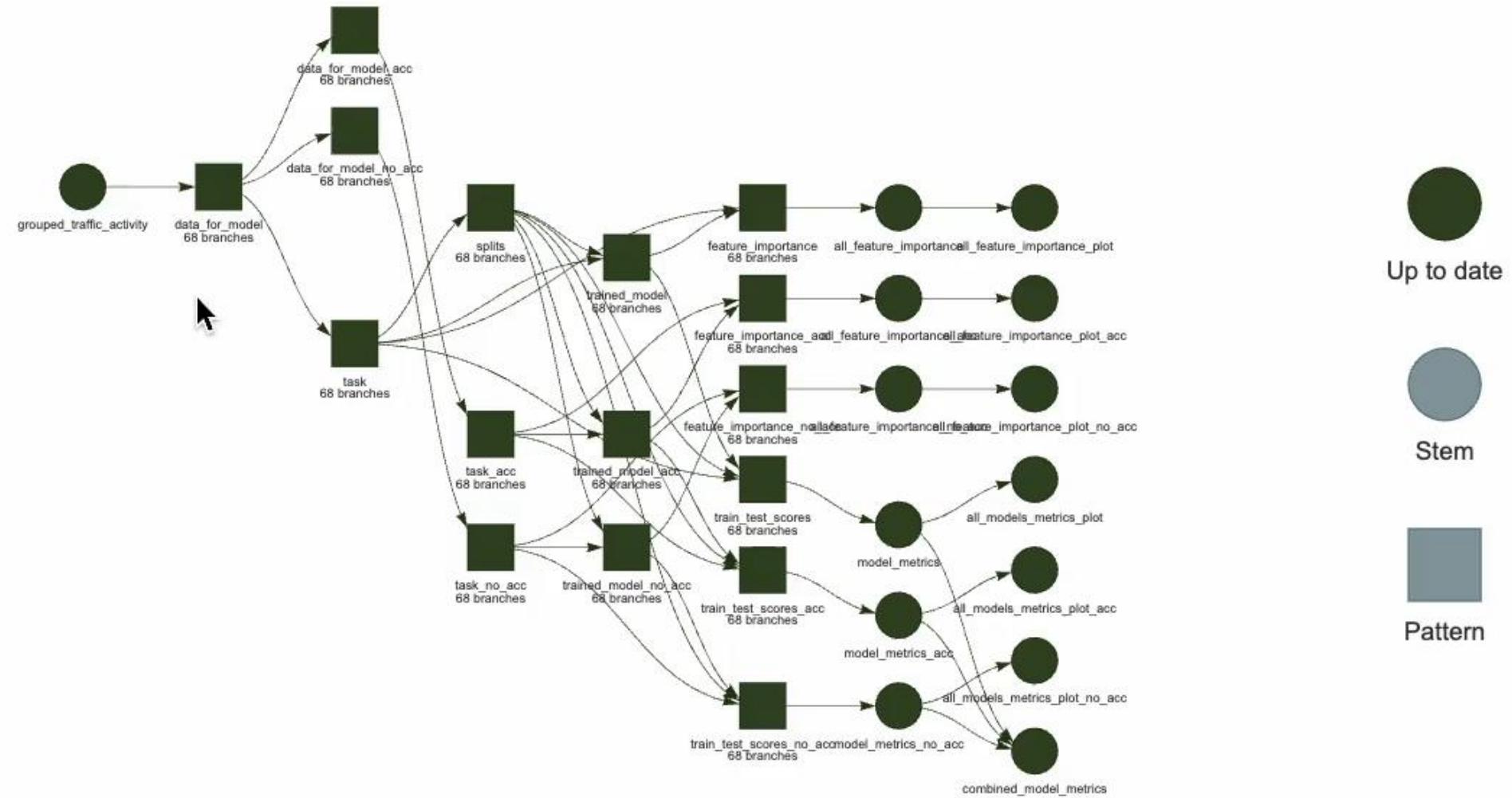
```
library(targets)
list(
  tar_target(output1,
    function_1() ),
  tar_target(output2,
    function_2(input1 = output1) ),
  tar_target(output3,
    function_3(input1 = output1,
      input_2 = output2) ),
)
```

The diagram illustrates the mapping between the target definitions in the R code and the corresponding project structure. Three arrows originate from the 'tar\_target' calls for 'output1', 'output2', and 'output3'. Each arrow points to one of the two directory names: '\_targets/' or '\_objects/'. Specifically, 'output1' maps to '\_targets/' and 'output2' and 'output3' map to '\_objects/'. This visualizes how each target in the list corresponds to a specific directory or object type within the project's file organization.

```
/rpoject01
- _targets.R
- R/
  - 01_scripts.R
  - 02_scripts.R
  - ...
- _targets/
- _objects/
```



# targets: COMPLEX WORKFLOW EXAMPLE





# containers : ENVIRONMENT PRESERVATION

Lock R packages and specific versions



Reproducible analysis workflow



You are here

Complete computational environment



SINGULARITYCE



APPTAINER



mac OS



R Studio



jupyter

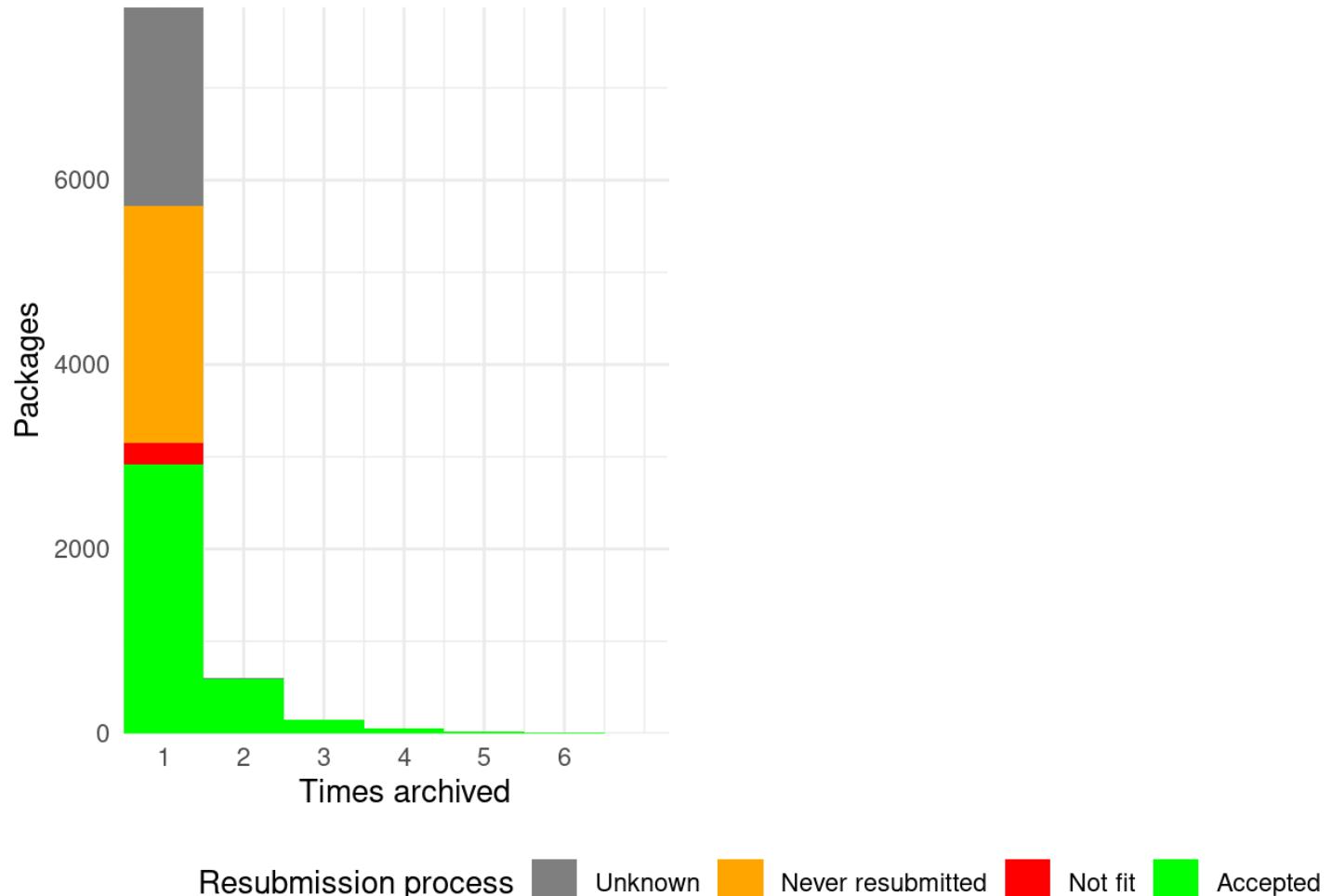


Logos by respective copyright holders



# 40% OF ALL PACKAGES EVER IN CRAN GOT AT ONE POINT ARCHIVED

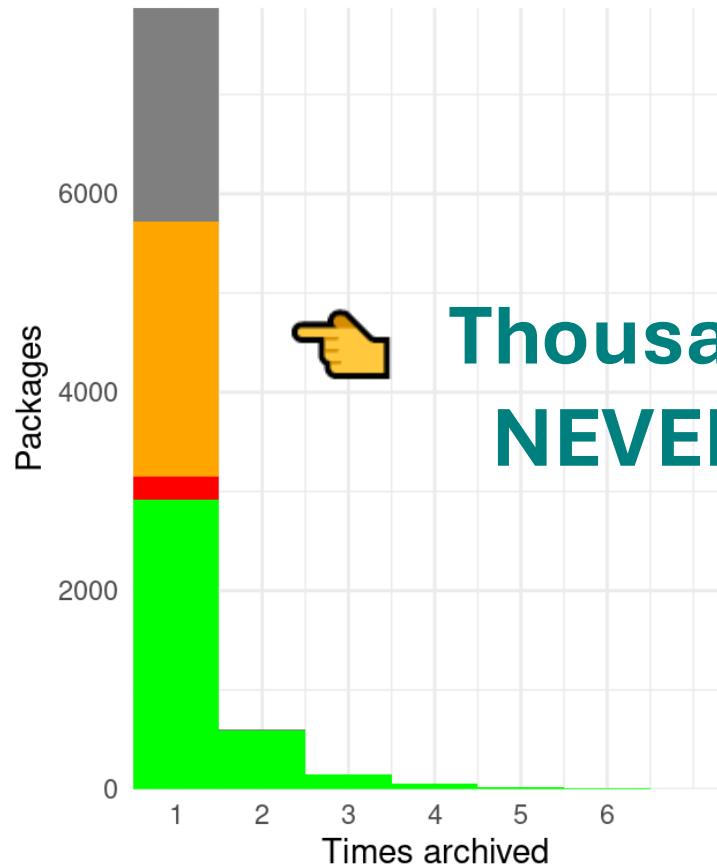
Times a package has been archived





# 40% OF ALL PACKAGES EVER IN CRAN GOT AT ONE POINT ARCHIVED

Times a package has been archived

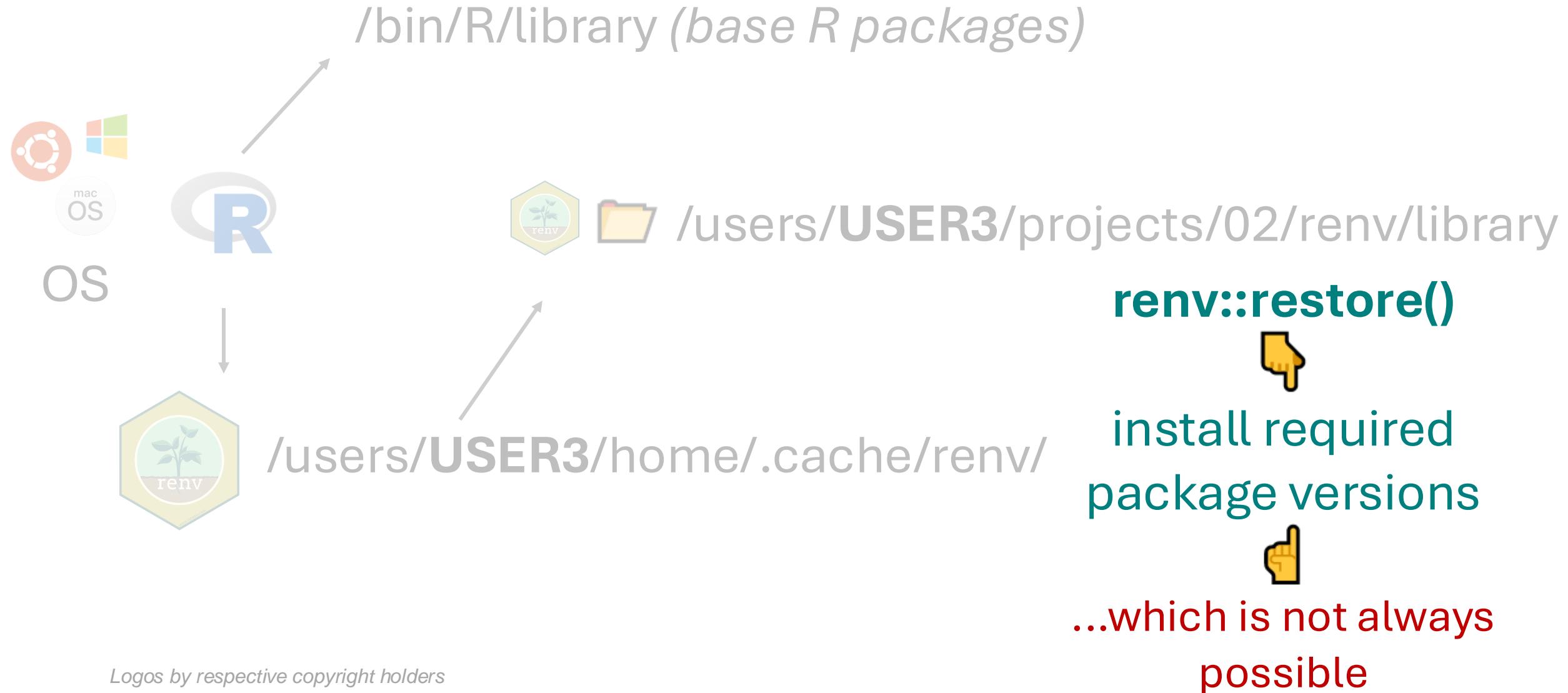


**Thousands on of packages are  
NEVER BACK ON CRAN!**

Resubmission process    Unknown    Never resubmitted    Not fit    Accepted



# renv: PACKAGE VERSION CONTROL



Console Terminal × Background Jobs ×

R 4.4.0 · ~/

```
> Sys.time() ; remotes::install_version("comradesoo") ; sys.time()
[1] "2024-06-17 23:28:52 CEST"
Downloading package from url: https://ftp.gwdg.de/pub/misc/cran//src/contrib/Archive/c
omradesoo/comradesoo_0.1.1.tar.gz
These packages have more recent versions available.
It is recommended to update all of them.
which would you like to update?
```

## remotes::install\_version()

Console Terminal × Background Jobs ×



R 4.4.0 · ~/

```
> Sys.time() ; remotes::install_version("comradesoo") ; sys.time()
[1] "2024-06-17 23:28:52 CEST"
Downloading package from url: https://ftp.gwdg.de/pub/misc/cran//src/contrib/Archive/comradesoo/comradesoo_0.1.1.tar.gz
These packages have more recent versions available.
It is recommended to update all of them.
which would you like to update?
```

1: All  
2: CRAN packages only  
3: None  
4: rlang (1.1.3 -> 1.1.4 ) [CRAN]  
5: fastmap (1.1.1 -> 1.2.0 ) [CRAN]  
6: cachem (1.0.8 -> 1.1.0 ) [CRAN]  
7: xfun (0.44 -> 0.45 ) [CRAN]  
8: highr (0.10 -> 0.11 ) [CRAN]  
9: evaluate (0.23 -> 0.24.0) [CRAN]  
10: knitr (1.46 -> 1.47 ) [CRAN]  
11: textshaping (0.3.7 -> 0.4.0 ) [CRAN]  
12: DBI (1.2.2 -> 1.2.3 ) [CRAN]

Enter one or more numbers, or an empty line to skip updates: 3

Installing 11 packages: crosstalk, reshape2, foreach, plotly, ade4, TopDom, heatmap3, igraph, doParallel, mixtools, seqinr



R 4.4.0 · ~/

```
> Sys.time() ; remotes::install_version("comradesoo") ; Sys.time()
[1] "2024-06-17 23:28:52 CEST"
Downloading package from url: https://ftp.gwdg.de/pub/misc/cran//src/contrib/Archive/comradesoo/comradesoo_0.1.1.tar.gz
These packages have more recent versions available.
It is recommended to update all of them.
which would you like to update?
```

```
1: All
2: CRAN packages only
3: None
4: rlang      (1.1.3 -> 1.1.4 ) [CRAN]
5: fastmap    (1.1.1 -> 1.2.0 ) [CRAN]
6: cachem     (1.0.8 -> 1.1.0 ) [CRAN]
7: xfun       (0.44  -> 0.45   ) [CRAN]
8: highr      (0.10  -> 0.11   ) [CRAN]
9: evaluate   (0.23  -> 0.24.0) [CRAN]
10: knitr     (1.46  -> 1.47   ) [CRAN]
11: textshaping (0.3.7 -> 0.4.0 ) [CRAN]
12: DBI        (1.2.2 -> 1.2.3 ) [CRAN]
```

```
Enter one or more numbers, or an empty line to skip updates: 3
```

```
Installing 11 packages: crosstalk, reshape2, foreach, plotly, ade4, TopDom, heatmap3,
igraph, doParallel, mixtools, seqinr
```

```
Installing packages into 'C:/Users/kotov/AppData/Local/R/win-library/4.4'
(as 'lib' is unspecified)
```

```
trying URL 'https://ftp.gwdg.de/pub/misc/cran/src/contrib/crosstalk_1.2.1.tar.gz'
Content type 'application/octet-stream' length 297970 bytes (290 KB)
downloaded 290 KB
```

```
trvina URL 'https://ftp.awda.de/pub/misc/cran/src/contrib/reshape2_1.4.4.tar.gz'
```

Console Terminal × Background Jobs ×

R 4.4.0 · ~/

> Sys.time() ; remotes::install\_version("comradesoo") ; sys.time()  
[1] "2024-06-17 23:28:52 CEST"

Downloading package from url: [https://ftp.gwdg.de/pub/misc/cran//src/contrib/Archive/comradesoo/comradesoo\\_0.1.1.tar.gz](https://ftp.gwdg.de/pub/misc/cran//src/contrib/Archive/comradesoo/comradesoo_0.1.1.tar.gz)

**2.5 hours later of building from source  
the package is NOT installed**

Warning message:

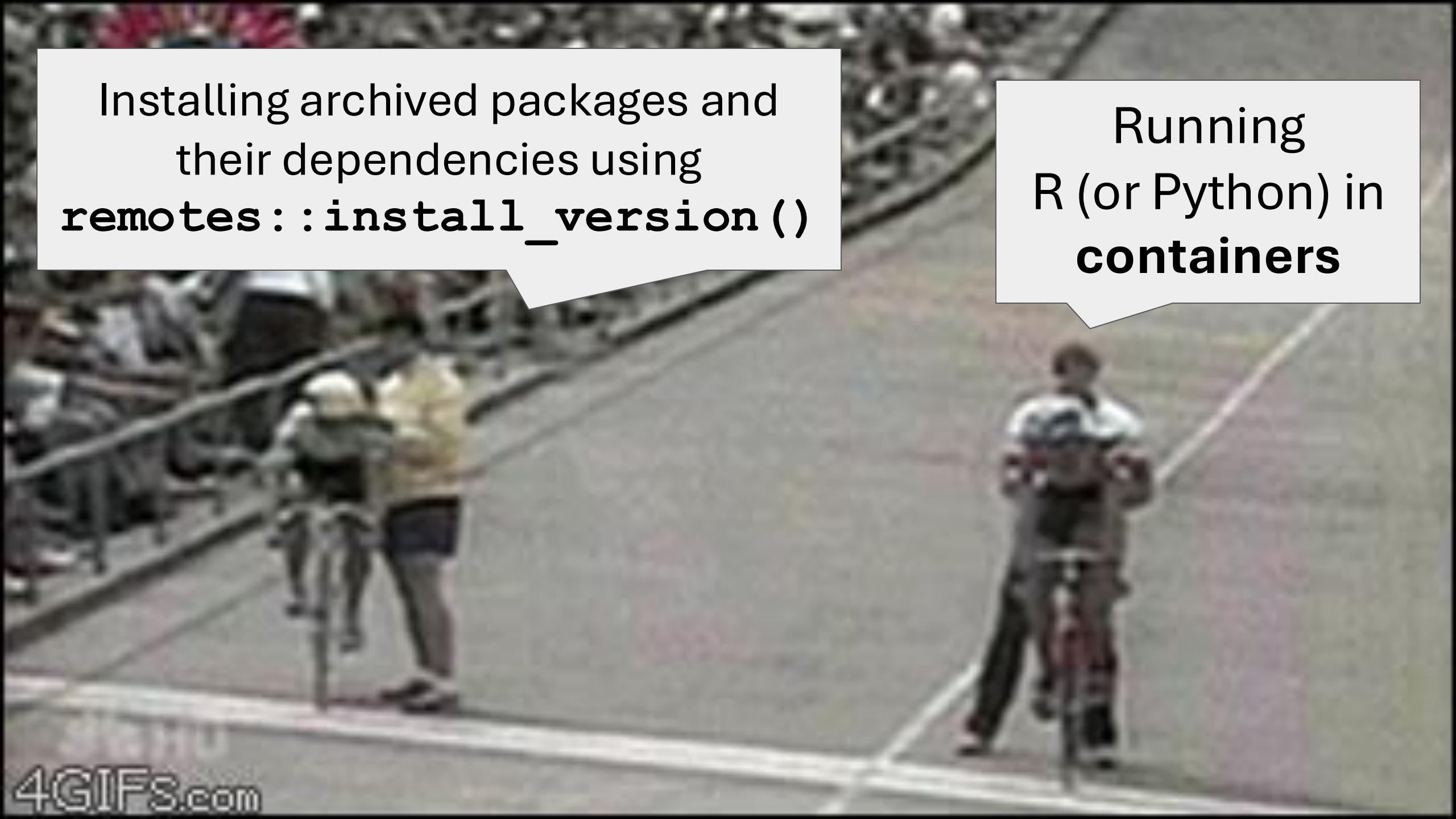
In i.p(...)

installation of package 'c:/users/kotov/AppData/Local/Temp/12/Rtmpu6bw0n/remotescd5475da3580/comradesoo' had non-zero exit status

[1] "2024-06-18 01:53:10 CEST"

Installing archived packages and  
their dependencies using  
`remotes::install_version()`





Installing archived packages and  
their dependencies using  
`remotes::install_version()`

Running  
R (or Python) in  
containers

# CONTAINERS FOR MORE CONSISTENT WORKFLOW

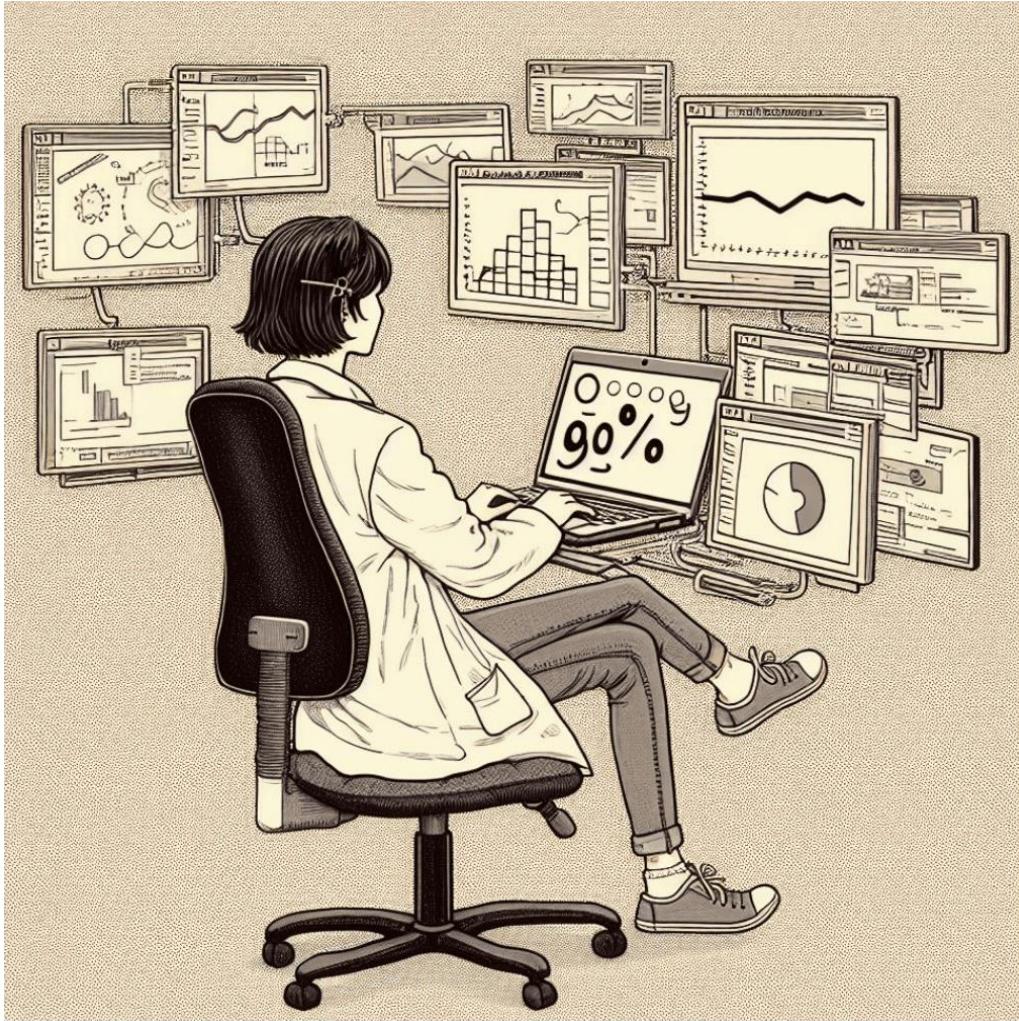
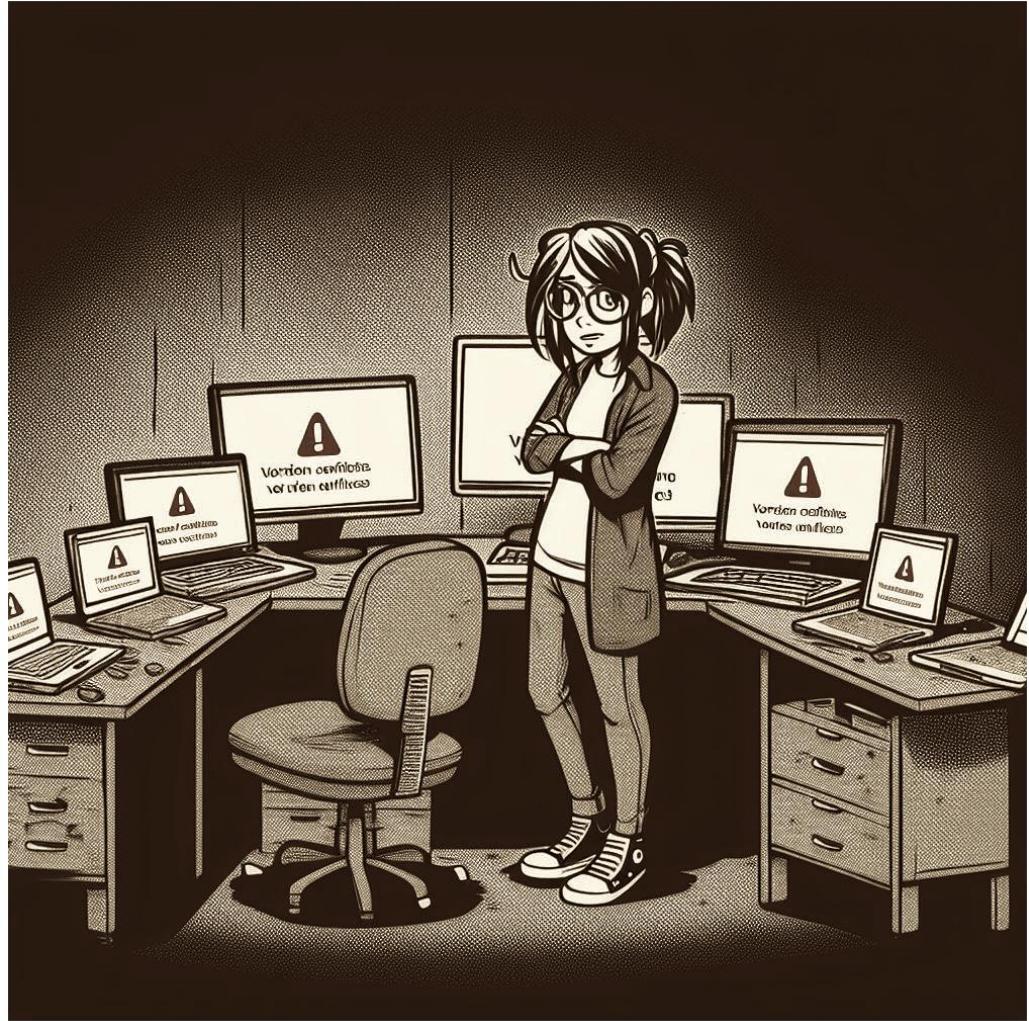


Image generated by Microsoft Image Creator powered by DALL-E 3

Kotov, E., and Denecke, E. (2024). Expanding the Lifespan of Software for Demographic Analysis with Containers: An Application of Spatial Sampling. *The Denominator, Population Dynamics Lab.* <https://doi.org/10.6069/WY8K-D973>



# WHAT ARE CONTAINERS?

Laptop for  
Project 1

Laptop for  
Project 2

Laptop for  
Project ...





## WHAT ARE CONTAINERS?

Hey, I want to  
run your code!



*Image by DALL-E*



## WHAT ARE CONTAINERS?

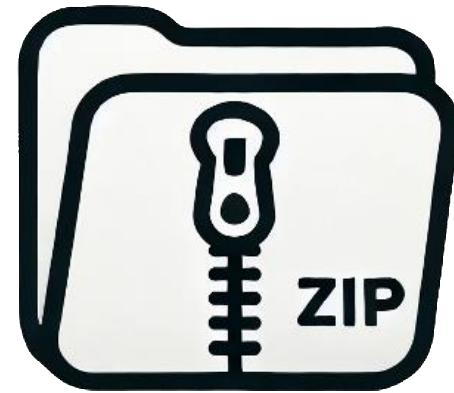
Hey, I want to  
run your code!



Sure, I'll send you  
my laptop #45 via  
DHL. Just give me  
a week to find  
which one it is.



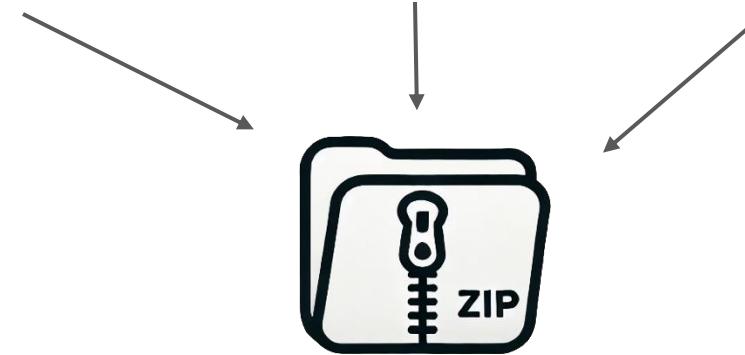
## WHAT ARE CONTAINERS?



**Project\_1\_OS+Rstudio+R+Packages.zip**



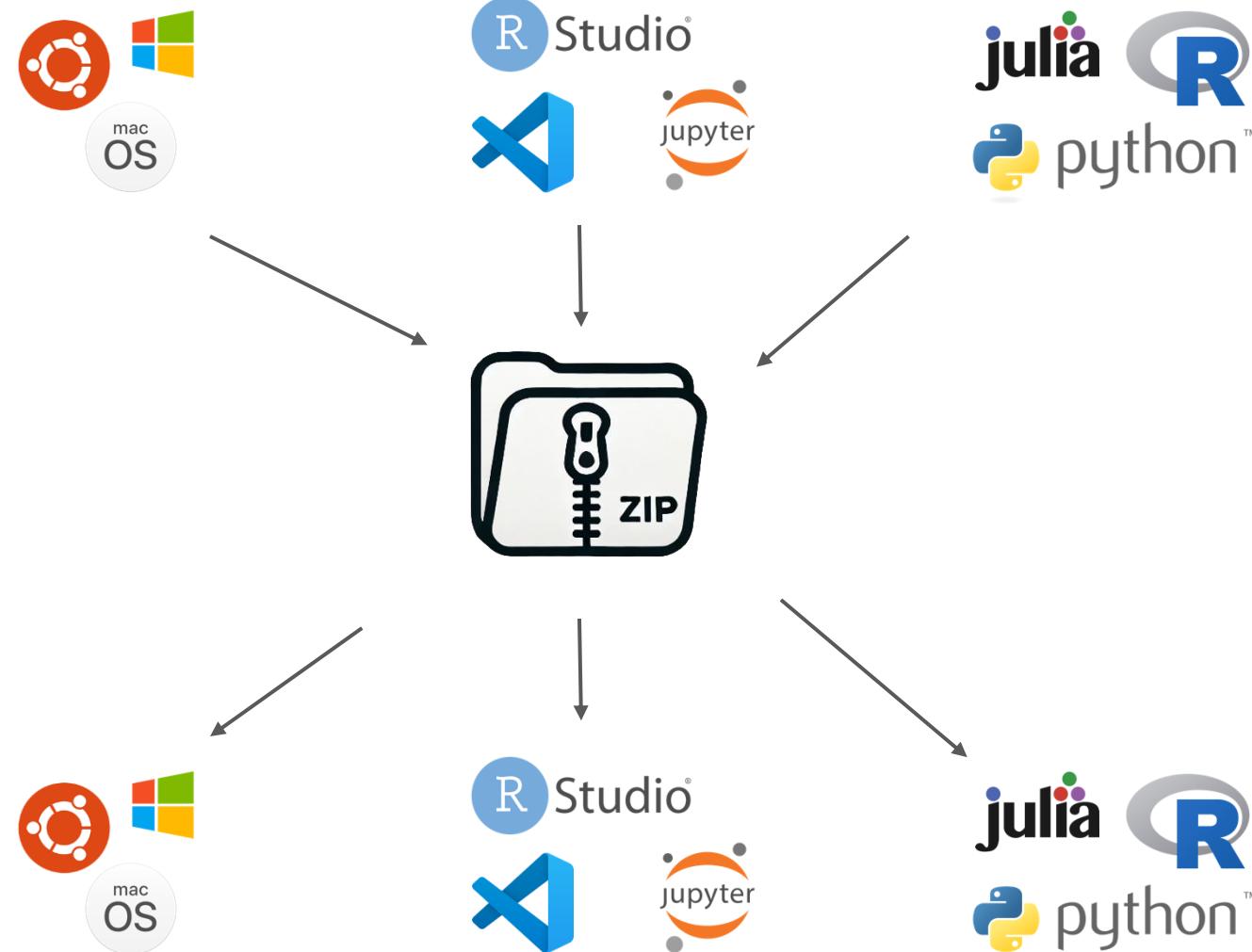
# WHAT ARE CONTAINERS?



*Logos by respective copyright holders*



# WHAT ARE CONTAINERS?



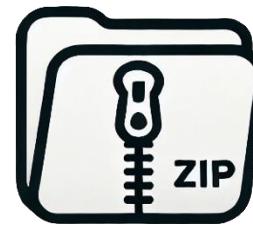
Logos by respective copyright holders



# WHAT ARE CONTAINERS?



That would take a lot of time...





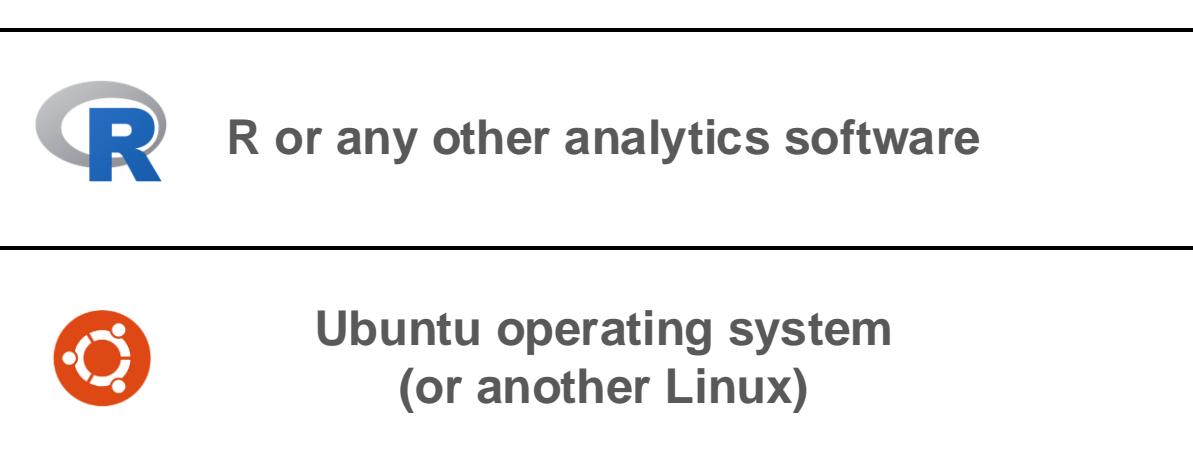
# WHAT ARE CONTAINERS?



Ubuntu operating system  
(or another Linux)



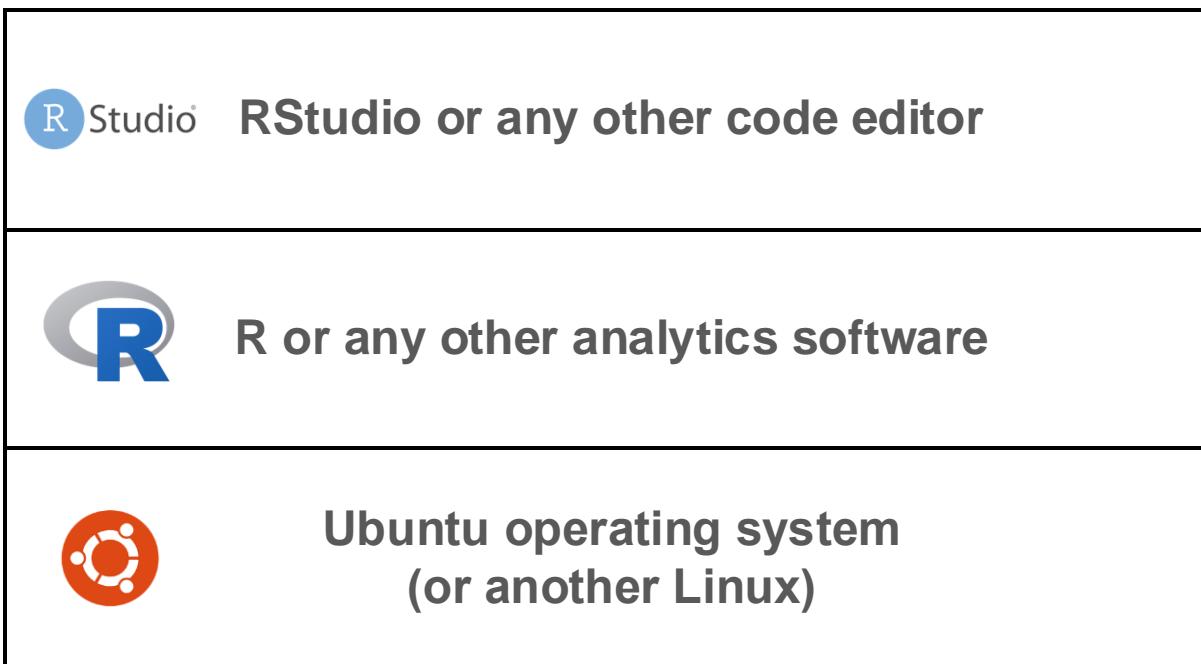
# WHAT ARE CONTAINERS?



*Logos by respective copyright holders*



# WHAT ARE CONTAINERS?





# WHAT ARE CONTAINERS?

**Any other software, scripts, files, anything!**



**RStudio or any other code editor**



**R or any other analytics software**



**Ubuntu operating system  
(or another Linux)**



## SOFTWARE TO BUILD AND RUN CONTAINERS



- Most popular, widely known
- Windows/macOS/Linux
- Many compatible alternative software



## SOFTWARE TO BUILD AND RUN CONTAINERS



- Most popular, widely known
- Windows/macOS/Linux
- Many compatible alternative software

- Often used in academic HPCs
- Mostly Linux
- Can import Docker containers





## SOFTWARE TO BUILD AND RUN CONTAINERS



- Most popular, widely known
- Windows/macOS/Linux
- Many compatible alternative software

- Often used in academic HPCs
- Mostly Linux
- Can import Docker containers



- Run small containers (1-2 GB memory)
- In the cloud, for free
- Based on Docker

# RUNNING ANALYSIS IN A CONTAINER @ MYBINDER.ORG



A screenshot of a GitHub repository page for 'grid-sample-containerized'. The page title is 'GitHub - e-kotov/grid-sample-containerized'. The main content features a large section titled 'Expanding Lifespan of Software for Demographic Analysis with Containers: A Case for a Sampling Technique'. Below this, there's a note: 'Run this repository in a web browser using Binder. Push the button >>' followed by a blue 'Launch Binder' button. A text block explains: 'This repository contains files for the article about using containers for reproducible research with applications in demography.' A 'Files description' table lists three files: Dockerfile, install.R, and main.Rmd. The Dockerfile is described as defining the container image, install.R as containing R package installation scripts, and main.Rmd as the primary analysis script. To the right of the main content area, there are sections for 'No releases published', 'No packages published', and a 'Languages' chart showing 80.5% Dockerfile and 19.5% R.

File	Description
<a href="#">Dockerfile</a>	This text file defines which container image Binder needs to launch.
<a href="#">install.R</a>	This R script contains lines to install packages required for the analysis.
<a href="#">main.Rmd</a>	This is the file with the example code we will run inside the container once it is launched.

Kotov, E., and Denecke, E. (2024). Expanding the Lifespan of Software for Demographic Analysis with Containers: An Application of Spatial Sampling. *The Denominator, Population Dynamics Lab.* <https://doi.org/10.6069/WY8K-D973> || repo <https://github.com/Population-Dynamics-Lab/grid-sample-containerized>

# RUNNING ANALYSIS IN A CONTAINER @ MYBINDER.ORG



GitHub - e-kotov/demographic-research.44-19-containerized x +

https://github.com/e-kotov/demographic-research.44-19-containerized?tab=readme-ov-file

Product Solutions Resources Open Source Enterprise Pricing Sign in Sign up

e-kotov / demographic-research.44-19-containerized Public Notifications Fork 0 Star 0

Code Issues Pull requests Actions Projects Security Insights

main 1 Branch 0 Tags Go to file Code

Author	Commit Message	Date	Commits
e-kotov	update regarding containers preservation	ebaae7b · 3 days ago	3 Commits
	Dockerfile4build	commit all files, initial try	4 days ago
	code	commit all files, initial try	4 days ago
	data	commit all files, initial try	4 days ago
	fig	commit all files, initial try	4 days ago
	media	update regarding containers preservation	3 days ago
	out	Create dummy.txt	4 days ago
	renv	commit all files, initial try	4 days ago
	.Rhistory	commit all files, initial try	4 days ago

About

doi.org/10.5281/zenodo.15024144

Readme Activity 0 stars 1 watching 0 forks Report repository

Packages 1

demographic-research.44-19-containerized

Languages

Schöley, J., 2021. The centered ternary balance scheme: A technique to visualize surfaces of unbalanced three-part compositions. *Demographic Research* 44, 443–458.  
<https://doi.org/10.4054/DemRes.2021.44.19>

Repo link: <https://github.com/e-kotov/demographic-research.44-19-containerized>

# PRESERVING THE COMPUTATIONAL ENVIRONMENT



Screenshot of a GitHub repository page for "e-kotov/demographic-research.44-19-containerized".

The repository is public and contains 1 branch and 0 tags. The main branch has 3 commits by user "e-kotov".

Commit	Message	Date
Dockerfile4build	commit all files, initial try	4 days ago
code	commit all files, initial try	4 days ago
data	commit all files, initial try	4 days ago
fig	commit all files, initial try	4 days ago
media	update regarding containers preservation	3 days ago
out	Create dummy.txt	4 days ago
renv	commit all files, initial try	4 days ago
.Rhistory	commit all files, initial try	4 days ago

The "About" section includes a DOI link: <https://doi.org/10.5281/zenodo.15024144>.

Other repository details shown in the sidebar:

- Readme
- Activity
- 0 stars
- 1 watching
- 0 forks
- Report repository

Packages: 1

- demographic-research.44-19-containerized

Languages

Schöley, J., 2021. The centered ternary balance scheme: A technique to visualize surfaces of unbalanced three-part compositions. *Demographic Research* 44, 443–458.  
<https://doi.org/10.4054/DemRes.2021.44.19>  
Repo link: <https://github.com/e-kotov/demographic-research.44-19-containerized>

# PRESERVING THE COMPUTATIONAL ENVIRONMENT



The screenshot shows a web browser window with the following details:

- Title Bar:** GitHub - e-kotov/demographic-research · Docker and Apptainer container images · https://zenodo.org/records/15024145
- Header:** zenodo
- Message:** **Planned Intervention:** On Tuesday March 18th 06:30 UTC Zenodo will be unavailable for 10-20 minutes to perform a storage cluster upgrade.
- Published:** March 14, 2025 | Version v0.1
- Buttons:** Software, Open
- Title:** Docker and Apptainer container images for Schöley, J., 2021, Demographic Research 44, 443–458
- Producers:** Kotov, Egor (Producer)<sup>1</sup> ; Schöley, Jonas (Project leader)<sup>1</sup>
- Affiliations:** Show affiliations
- Description:** This repository contains R+RStudio computational environment preserved in Docker and Apptainer container images to reproduce the code in **Schöley, J., 2021. The centered ternary balance scheme: A technique to visualize surfaces of unbalanced three-part compositions. Demographic Research 44, 443–458.** <https://doi.org/10.4054/DemRes.2021.44.19>. These containerized environments contain all R package and system level dependencies required to reproduce the code and do not rely on any external online repositories to install anything else at runtime.
- GitHub Registry:** The Docker image is also hosted in the GitHub Container registry at <https://github.com/e-kotov/demographic-research.44-19-containerized/pkgs/container/demographic-research.44-19-containerized> and the live preview in a web browser can be launched from the repository at <https://github.com/e-kotov/demographic-research.44-19-containerized>.

Schöley, J., 2021. The centered ternary balance scheme: A technique to visualize surfaces of unbalanced three-part compositions. *Demographic Research* 44, 443–458.

<https://doi.org/10.4054/DemRes.2021.44.19>

Repo link: <https://github.com/e-kotov/demographic-research.44-19-containerized>

# PRESERVING THE COMPUTATIONAL ENVIRONMENT



The screenshot shows a web browser window with the title "GitHub - e-kotov/demographic-research" and the tab "Docker and Apptainer containerized". The URL is <https://zenodo.org/records/15024145>. The page content includes a note about launching a web browser from the repository at <https://github.com/e-kotov/demographic-research.44-19-containerized>. Below this, there's a "Files" section with two items:

- r422scholey2021-apptainer-container-image.sif** (md5:580972eeeb876534a054a492980068cff) - 1.6 GB
- r422scholey2021-docker-container-image.tar** (md5:3e17a2116735d41d71e843dd2cbb81f3) - 5.0 GB

Each item has a "Download" button. To the right of the files, there are logos for Docker (blue whale icon), SingularityCE (S icon), and Apptainer (A icon).

**Additional details**

**Related works**    **Compiles**

Schöley, J., 2021. The centered ternary balance scheme: A technique to visualize surfaces of unbalanced three-part compositions. *Demographic Research* 44, 443–458.  
<https://doi.org/10.4054/DemRes.2021.44.19>

Repo link: <https://github.com/e-kotov/demographic-research.44-19-containerized> ; Logos by respective copyright holders

WITH CONTAINERS YOU ARE FREE TO RUN ANY  
SOFTWARE, NO NEED TO BOTHER YOUR IT



*Image from The I.T Crowd C4 TV Show*

*For the record, this is not in any way representative of the IT team at MPI IDR or GWDG HPC*



# WHERE TO GET CONTAINERS?



*Icons and search bar by Google, but feel free to use other search engines*



# WHERE TO GET CONTAINERS?

The screenshot shows a web browser window with the URL <https://hub.docker.com/r/jupyter/datasience-notebook>. The page is titled "jupyter/datasience-notebook" and is described as a "Sponsored OSS" project with 1.1K stars. It is a Data Science Jupyter Notebook Python Stack from <https://github.com/jupyter/docker-stacks>. The page includes tabs for "IMAGE" and "ARTIFACT", and categories for "DATA SCIENCE", "LANGUAGES & FRAMEWORKS", and "MACHINE LEARNING & AI". Below the main title, there is a section for the "Jupyter Notebook Data Science Stack" which states that images are no longer updated and points to [quay.io](https://quay.io/image/jupyter/datasience-notebook). It also shows statistics: 33M docker pulls, 1.1k docker stars, and an image size of 1.9 GB. A "Docker Pull Command" box contains the command `docker pull jupyter/datasience-notebook`, with a "Copy" button next to it.

<https://hub.docker.com/r/jupyter/datasience-notebook>



# WHERE TO GET CONTAINERS?

Rocker Project

The Rocker Project

Docker Containers for the R Environment

Getting Started

Home   Code of Conduct   Images ▾

The Rocker Images

VERSIONED STACK

- r-ver
- rstudio, tidyverse, verse, geospatial
- binder
- shiny, shiny-verse
- cuda, ml, ml-verse

BASE STACK

- r-base ↗
- rocker/r-base
- r-devel, drd, drp, r-devel-san, r-devel-ubsan-clang

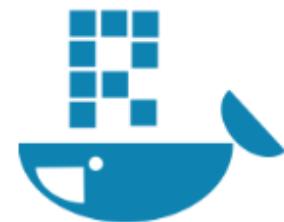
ADDITIONAL IMAGES

Source: <https://rocker-project.org/>



# Vanilla R

 ROCKER PROJECT



## The Rocker Project

Docker Containers for the R Environment

 **Getting Started**

Home Code of Conduct Images ▾

### The Rocker Images

#### VERSIONED STACK

r-ver

rstudio, tidyverse, verse, geospatial

binder

shiny, shiny-verse

cuda, ml, ml-verse

#### BASE STACK

r-base 

rocker/r-base

r-devel, drd, drp, r-devel-san, r-devel-ubsan-clang

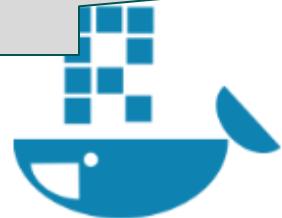
#### ADDITIONAL IMAGES



# Vanilla R

 ROCKER PROJECT

Various packages  
preinstalled



## The Rocker Project

Docker Containers for the R Environment

 **Getting Started**

Home Code of Conduct Images ▾

### The Rocker Images

#### VERSIONED STACK

r-ver

rstudio, tidyverse, verse, geospatial

binder

shiny, shiny-verse

cuda, ml, ml-verse

#### BASE STACK

r-base 

rocker/r-base

r-devel, drd, drp, r-devel-san, r-devel-ubsan-clang

#### ADDITIONAL IMAGES



# Vanilla R

 ROCKER PROJECT

Various packages  
preinstalled

geospatial + specially  
designed for  
[mybinder.org](https://mybinder.org), has most  
popular packages

**Project**  
the R Environment  
started

Home Code of Conduct Images ▾

The Rocker Images

VERSIONED STACK

r-ver

rstudio, tidyverse, verse, geospatial

binder

shiny, shiny-verse

cuda, ml, ml-verse

BASE STACK

r-base 

rocker/r-base

r-devel, drd, drp, r-devel-san, r-devel-ubsan-clang

ADDITIONAL IMAGES



## WHERE TO GET CONTAINERS?

The screenshot shows the Docker Hub page for the 'rocker/binder' repository. The 'Tags' tab is selected, showing two entries:

- Tag: 3.5**  
Last pushed 6 years ago by [rocker](#)  
Digest: [4f22a9c1b130](#)  
OS/ARCH: linux/amd64  
Compressed Size: 1.49 GB  
Actions: docker pull rocker/binder:3.5 (with a Copy button), Compressed Size (with a Copy button)
- Tag: 3**  
Last pushed 6 years ago by [rocker](#)  
Digest: [4f22a9c1b130](#)  
OS/ARCH: linux/amd64  
Compressed Size: 1.49 GB  
Actions: docker pull rocker/binder:3 (with a Copy button), Compressed Size (with a Copy button)

A large callout box on the right side of the screenshot contains the text:

Find the R version you need,  
perhaps considering the  
release date of this specific  
version



## 4. EXAMPLE: DESIGN AND BUILD CONTAINERS

Which pre-built  
container to use

```
FROM rocker/binder:4.0.1
```



## 4. EXAMPLE: DESIGN AND BUILD CONTAINERS

```
📦 Dockerfile M X
📦 Dockerfile
1  FROM rocker/binder:4.0.1
2
3  COPY --chown=${NB_USER} . ${HOME}
```

Which pre-built container to use

Copy all files in the same directory into the container



## 4. EXAMPLE: DESIGN AND BUILD CONTAINERS

```
📦 Dockerfile M X
📦 Dockerfile
1  FROM rocker/binder:4.0.1
2
3  COPY --chown=${NB_USER} . ${HOME}
4
5  RUN if [ -f install.R ]; then R --quiet -f install.R; fi
```

Which pre-built container to use

Copy all files in the same directory into the container

Run R script to install packages



## 4. EXAMPLE: DESIGN AND BUILD CONTAINERS

```
install.packages("MortalitySmooth")
install.packages("gridsample")
```



## Egor Kotov

Ph.D. Student  
Laboratory of Migration and Mobility

[kotov@demogr.mpg.de](mailto:kotov@demogr.mpg.de)

## Jonas Schöley

Research Scientist  
Laboratory of Population Health

[schoeley@demogr.mpg.de](mailto:schoeley@demogr.mpg.de)

