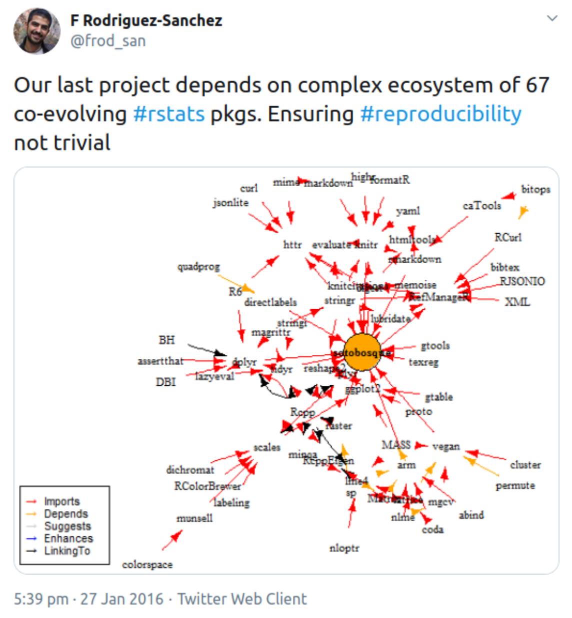
Controlling dependencies

Francisco Rodriguez-Sanchez

@frod_san

https://frodriguezsanchez.net



F Rodriguez-Sanchez
@frod_san

Our last project depends on complex ecosystem of 67 co-evolving #rstats pkgs. Ensuring #reproducibility not trivial

5:39 pm · 27 Jan 2016 · Twitter Web Client

Package changes can break your analysis

How to reproduce your analysis in a year,
or different computer?

sessionInfo records OS & used packages

```r
sessionInfo()
```

```
R version 4.2.0 (2022-04-22)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Ubuntu 20.04.4 LTS

Matrix products: default
BLAS:   /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.9.0
LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.9.0

locale:
 [1] LC_CTYPE=en_GB.UTF-8       LC_NUMERIC=C
 [3] LC_TIME=es_ES.UTF-8        LC_COLLATE=en_GB.UTF-8
 [5] LC_MONETARY=es_ES.UTF-8    LC_MESSAGES=en_GB.UTF-8
 [7] LC_PAPER=es_ES.UTF-8       LC_NAME=C
 [9] LC_ADDRESS=C               LC_TELEPHONE=C
[11] LC_MEASUREMENT=es_ES.UTF-8 LC_IDENTIFICATION=C

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base

other attached packages:
[1] knitr_1.40

loaded via a namespace (and not attached):
 [1] compiler_4.2.0  magrittr_2.0.3  fastmap_1.1.0   cli_3.3.0
 [5] htmltools_0.5.3 tools_4.2.0     rstudioapi_0.14 yaml_2.3.5
 [9] stringi_1.7.8   rmarkdown_2.16  binb_0.0.6      stringr_1.4.1
[13] xfun_0.32       digest_0.6.29   rlang_1.0.5     evaluate_0.16
```

checkpoint reconstructs packages in given date

```
library('checkpoint')

checkpoint(”2019-10-08”)

source(”analysis.R”)
```

1. Detects packages used
2. Installs version from given date (only CRAN)
3. Independent install (not messing w/ main library)

automagic records & install packages (CRAN + GitHub)

```
automagic::make_deps_file()
```

File `deps.yaml` records dependencies:

```
- Package: equatiomatic
  Repository: CRAN
  Version: 0.1.0

- Package: report
  GithubUsername: easystats
  GithubRepo: report
  GithubRef: HEAD
  GithubSHA1: c48a4bb0a40df7116bc502aa3ce2cbbc9d70b7e2
```

To install all those dependencies:

```
automagic()
```

renv also controls dependencies

```r
renv::init()
# Create private package library for project

renv::snapshot()
# Capture dependencies in lockfile

renv::restore()
# Regenerate dependencies from lockfile
```

https://environments.rstudio.com/

To ensure reproducibility,

besides R packages

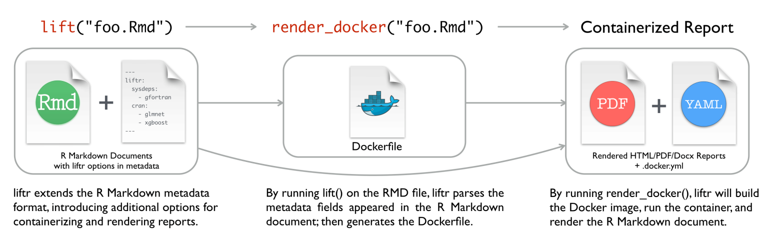
we also need to control

computational environment

`Docker` recreates virtual systems

from a `Dockerfile`

liftr: process Rmd in Docker container



```
---
liftr:
  sysdeps:
    - gfortran
  cran:
    - glmnet
    - xgboost
---
```

liftr extends the R Markdown metadata format, introducing additional options for containerizing and rendering reports.

By running lift() on the RMD file, liftr parses the metadata fields appeared in the R Markdown document; then generates the Dockerfile.

By running render_docker(), liftr will build the Docker image, run the container, and render the R Markdown document.

https://liftr.me/

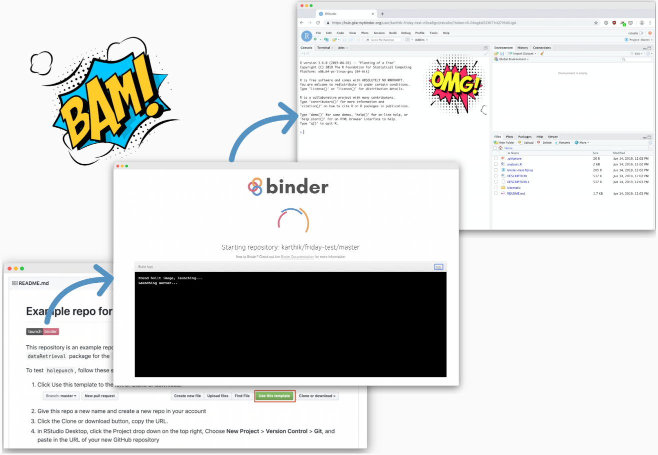
containerit creates Dockerfile

```
library(”containerit”)

dockfile <- dockerfile(from = ”mypaper.Rmd”)
```

https://o2r.info/containerit

holepunch: reproduce analysis in the cloud (Binder)



https://hub.gke.mybinder.org/user/karthik-friday-test-n9ca6gci/rstudio/?token=6-0AsgkAS2WT1vQ7VNXUgA

```
R version 3.6.0 (2019-04-26) -- "Planting of a Tree"
Copyright (C) 2019 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

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.

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.

> |
```

Name	Size	Modified
.gitignore	26 B	Jun 14, 2019, 12:02 PM
analysis.R	2 KB	Jun 14, 2019, 12:02 PM
binder-test.Rproj	205 B	Jun 14, 2019, 12:02 PM
DESCRIPTION	517 B	Jun 14, 2019, 12:02 PM
DESCRIPTION.1	517 B	Jun 14, 2019, 12:03 PM
kitematic		
README.md	1.7 KB	Jun 14, 2019, 12:02 PM

Starting repository: karthik/friday-test/master

New to Binder? Check out the Binder Documentation for more information

```
Found built image, launching...
Launching server...
```

Example repo for

This repository is an example repo
dataRetrieval package for the

To test holepunch, follow these s

1. Click Use this template to the
2. Give this repo a new name and create a new repo in your account
3. Click the Clone or download button, copy the URL.
4. in RStudio Desktop, click the Project drop down on the top right, Choose **New Project** > **Version Control** > **Git**, and paste in the URL of your new GitHub repository

https://karthik.github.io/holepunch/

Your turn

Checkpoint

- Create script/Rmd using different packages
- Call `checkpoint` on former date

automagic/renv

- Record dependencies:
    - `automagic::make_deps_file()`
    - `renv::snapshot`
- Recreate packages
    - `automagic()`
    - `restore()`