



# PROD IS NOT A FOUR-LETTER WORD

How to confidently send R to production.

COLIN FAY



> whoami

# Colin FAY

Data Scientist & R-Hacker at ThinkR, a french company focused on Data Science & R Engineering.

Hyperactive open source developer.

<http://rtask.thinkr.fr>

[http://twitter.com/\\_colinfay](http://twitter.com/_colinfay)

<http://github.com/colinfay>



**INKR**



**R & DATA SCIENCE  
ENGINEERING**



**TRAINING**




**CONSULTING**



# PROD IS NOT A FOUR-LETTER WORD

How to confidently send R to production.





What are we  
going to talk  
about today?

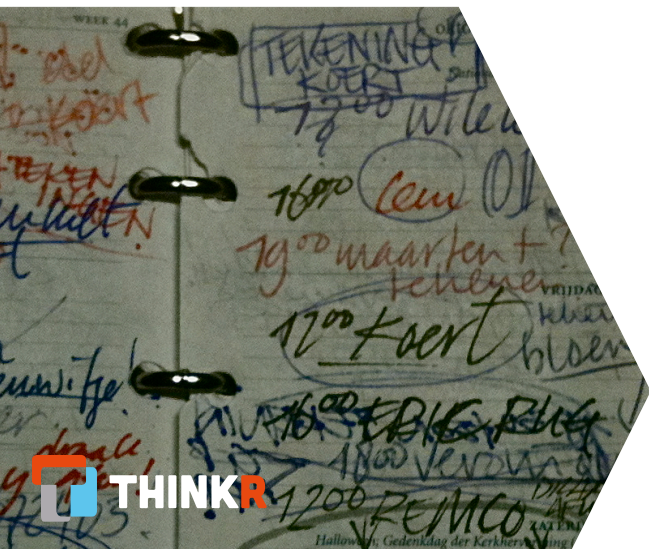
# SCHEDULE

01.  What does production even mean?

02.  About R in production

03.  Let's talk about tools

04.  R in production, the challenges



# Q & A


Have you ever heard?

"That's good that we have this prototype in R, now let's rewrite this with a production language"

# Q & A

We've come to a point that we've heard this so much that we convinced ourselves that this is true.



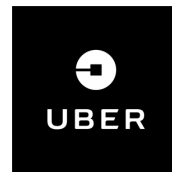


“ All programming languages are basically just bundles of trade-offs, and so no single language is going to be globally superior to another. **There really are no production languages – only production engineers.** ”

Gordon Shotwell

[https://blog.shotwell.ca/posts/why\\_i\\_use\\_r/](https://blog.shotwell.ca/posts/why_i_use_r/)

THEY  
USE R  
IN PROD



<https://github.com/ThinkR-open/companies-using-r>



DON'T TELL ME  
R IS NOT PRODUCTION R READY



# 01

DID YOU SAY 'PRODUCTION'?





**John Detlefs**  
@jdetle



Question for machine learning people in industry: what does it mean for your work to be "in production" what qualifies your models to be "production ready"? See this a lot and I think the meanings change across different domains pretty quickly.

8:26 PM · 2 mars 2020 · [Twitter for iPhone](#)

<https://twitter.com/jdetle>

“

Production is anything that is run repeatedly (or continuously) and is **relied upon**.

”

Mark Sellors

“

Software environments that are **used and relied on by real users with real consequences** if things go wrong

”

Joe Cheng



Fred v3.0.0 🤖🎸

@ProbNotABot

En réponse à @jdetle

As an engineer: Versioned, resilient, modular so it can be easily changed/iterated over and I can switch environments by configuration settings. More than business logic/prediction. Available/Integrated into a system

As anyone else: when it is producing value that makes impact.

6:42 AM · 3 mars 2020 · Twitter for Android

<https://twitter.com/ProbNotABot>

USERS



THE PRODUCT  
USER(S)

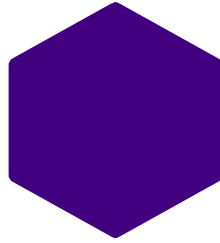


THE PRODUCT  
ENGINEER(S)

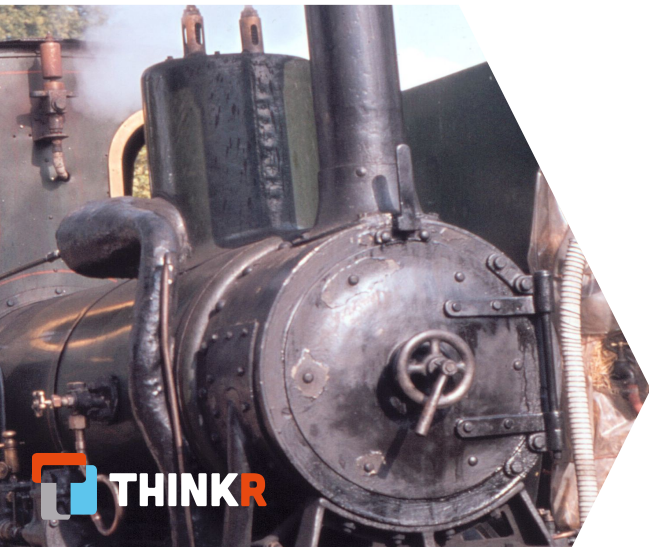




IN PROD



IS USED





**Ethan Rosenthal**

@eprosenthal



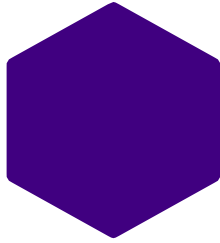
If nobody's using it, it's not in production

8:46 PM · 13 janv. 2020 · [Twitter Web App](#)

3 Retweets   13 J'aime

<https://twitter.com/eprosenthal>

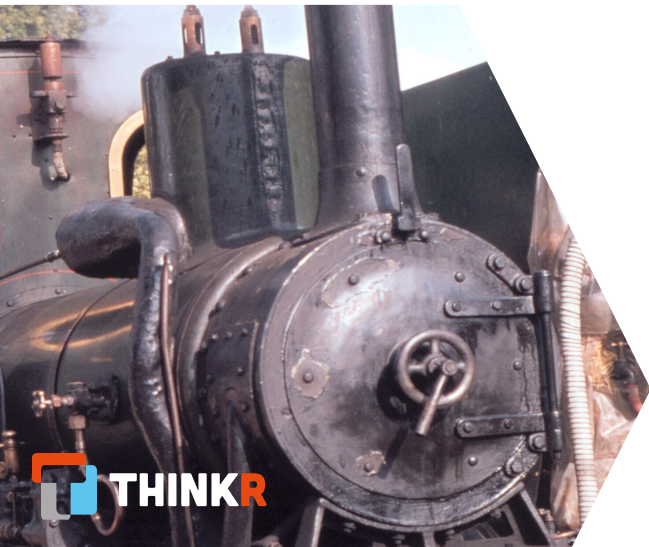
IN PROD



IS USED



IS RELIED UPON





**dierken**  
@dierken



En réponse à @jdetle

If turning it off breaks your business, it's in production

7:13 AM · 3 mars 2020 · Twitter for iPad

<https://twitter.com/dierken>



USERS  
TRUST  
THE  
SOFT



TO BE AVAILABLE



TO MAKE DECISION



TO DO THEIR DAILY JOB

DEV  
TRUST  
THE  
SOFT



TO KEEP RUNNING



TO BE BUG FREE

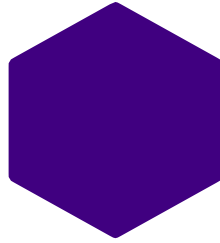


TO BE CUSTOMISABLE



TO BE SAFE TO UPDATE

IN PROD



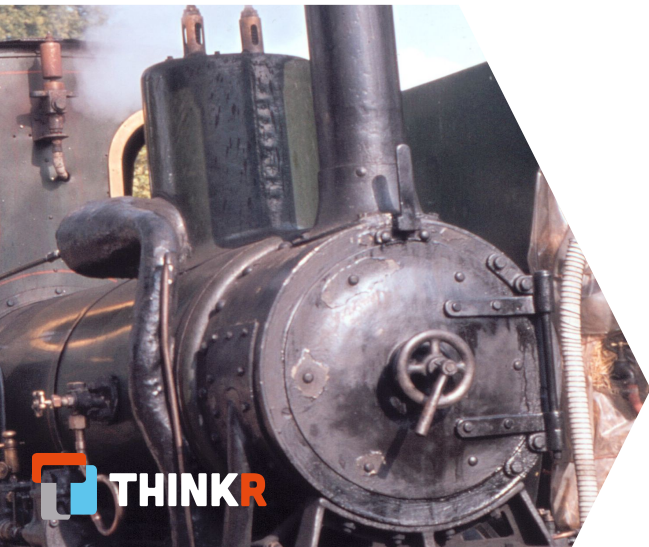
IS USED



IS RELIED UPON



HAS REAL LIFE  
IMPACT





**Sean J. Taylor**

@seanjtaylor

En réponse à @peteskomoroch @jdetle et 9 autres

I'd explain the main difference as one of accountability. If someone can blame you when something goes wrong and you're responsible for fixing in a timely manner, then it's in production. The proxy I'd use is if you have an oncall rotation.

6:59 AM · 3 mars 2020 · [Twitter for iPhone](#)

<https://twitter.com/seanjtaylor>

# WHAT CAN GO WRONG

❌ SOFTWARE IS UNAVAILABLE

💣 UPDATE BREAKS THE CODE-BASE

🚫 UPDATE BREAKS ANOTHER  
SERVICE

🤪 RESULTS ARE NOT RELIABLE

# REAL LIFE IMPACT



DATA LOSS




MONEY LOSS



TIME LOSS



BAD DECISIONS



```
> PRODUCTION_TOOLS == RELIABLE_TOOLS  
[1] TRUE
```





# 02

## ABOUT R AND PRODUCTION



# Q & A



How many ways are there  
to build a package in R?

# BUILDING RELIABLE TOOLS



RELY ON FRAMEWORKS



DON'T REINVENT THE WHEEL



TUTORIALS ARE EVERYWHERE



COMMENT, TEST, AND MONITOR



CONVENTIONS ARE AWESOME

(And so is the R community)



# 03

LET'S TALK ABOUT  
TOOLS

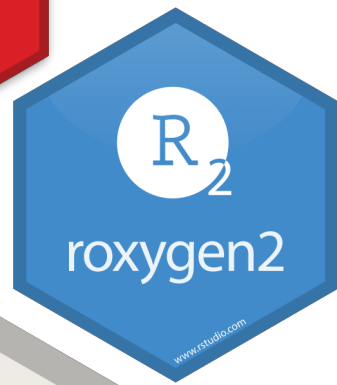
# R PACKAGES

🤝 MAIN GOAL IS TO BE  
SHARED AND REUSED

🔧 YEARS OF TOOLKITS

🍷 JUST A LIBRARY() AWAY

DEVELOP



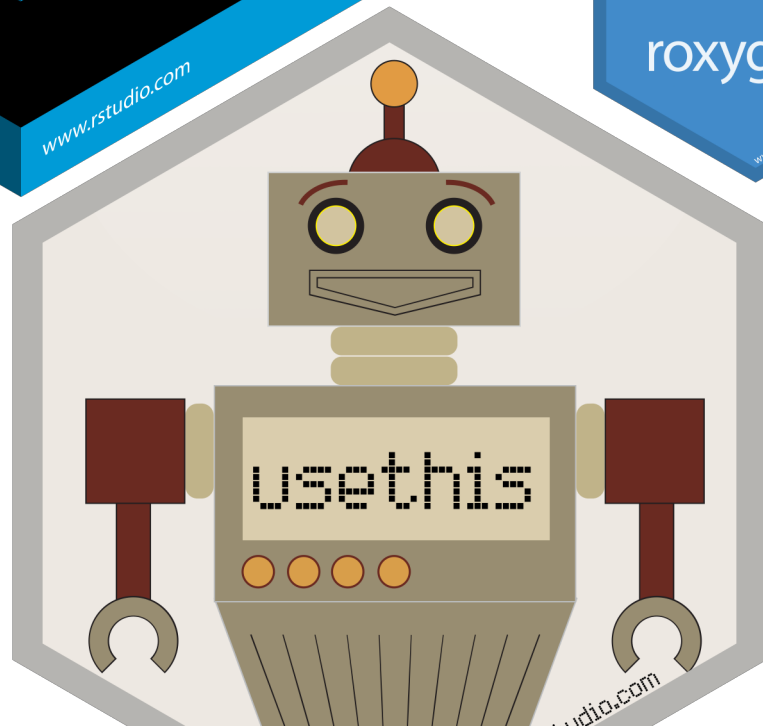
AUTOMATE



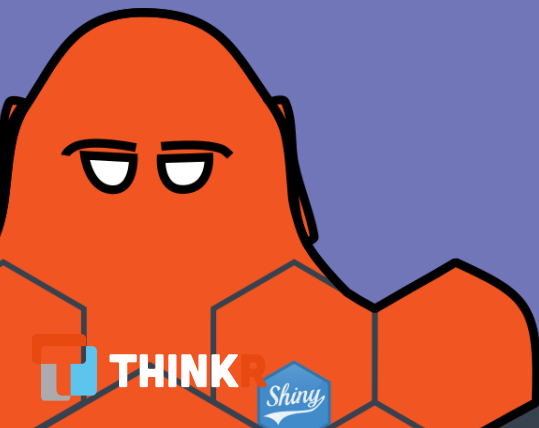
DOCUMENT



TEST



golem



# SHINY APPS WITH {GOLEM}



FRAMEWORK



AUTOMATION



COLLECTIVE  
INTELLIGENCE

<https://www.golemverse.org/>



Latest release

0.2.1

bf9d041

Verified

Compare ▼

## golem 0.2.1



ColinFay released this yesterday

This release matches the CRAN release of {golem} v 0.2.1

### ▼ Assets 2



Source code (zip)



Source code (tar.gz)

<https://cran.r-project.org/package=golem>



# DEPLOY



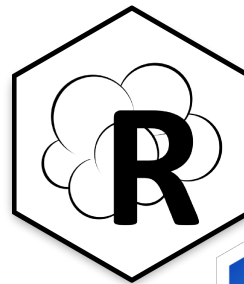
RSTUDIO



DOCKER / KUBE



AWS





04

R IN PRODUCTION  
THE CHALLENGES

# CHALLENGES



NOT THE  
LANGUAGE



NOT THE  
TOOLS



# CHALLENGES



CULTURAL



ORGANISATIONAL





**Sean J. Taylor**

@seanjtaylor

En réponse à @sarahcat21 @peteskomoroch et 9 autres

Believe it or not, I think the main problem is culture, not tech. People love building models but hate paying the interest payments on their tech debt.

7:11 AM · 3 mars 2020 · Twitter for iPhone

<https://twitter.com/seanjtaylor>

# Q & A

Who has ever experienced?

"Woaw, your PoC is awesome, let's put that on the production server and give access to 100 users!"



# R DEV



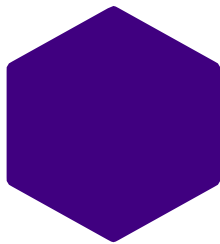
The good thing about R is that one can start using it and get results in a couple of hours.



🤔 The bad thing about R is that one can start using it and get results in a couple of hours.



# R DEV



R developers don't learn R as software engineers.



R developers don't know what's needed for productions.



IT might not be receptive.





# ADOPT A PROD MINDSET



NO MORE “THAT WORKS ON MY MACHINE”



EVERYTHING IS A PACKAGE



TEST AND DOCUMENT



DON'T START WITH TECHNICAL DEBT

ADOPT  
A  
PROD  
MINDSET



IT ARE YOUR FRIENDS



INFRASTRUCTURE IS CRUCIAL



RELIABLE CODE MATTERS



YOU CAN BE A GOOD  
SOFTWARE ENGINEER



# R WORKS IN PRODUCTION



# THANK YOU!

Questions?

COLIN@THINKR.FR

[HTTPS://TWITTER.COM/\\_COLINFAY](https://twitter.com/_COLINFAY)

[HTTPS://TWITTER.COM/THINKR\\_FR](https://twitter.com/THINKR_FR)

[HTTPS://RTASK.THINKR.FR](https://rtask.thinkr.fr)

[HTTPS://COLINFAY.ME](https://colinfay.me)

