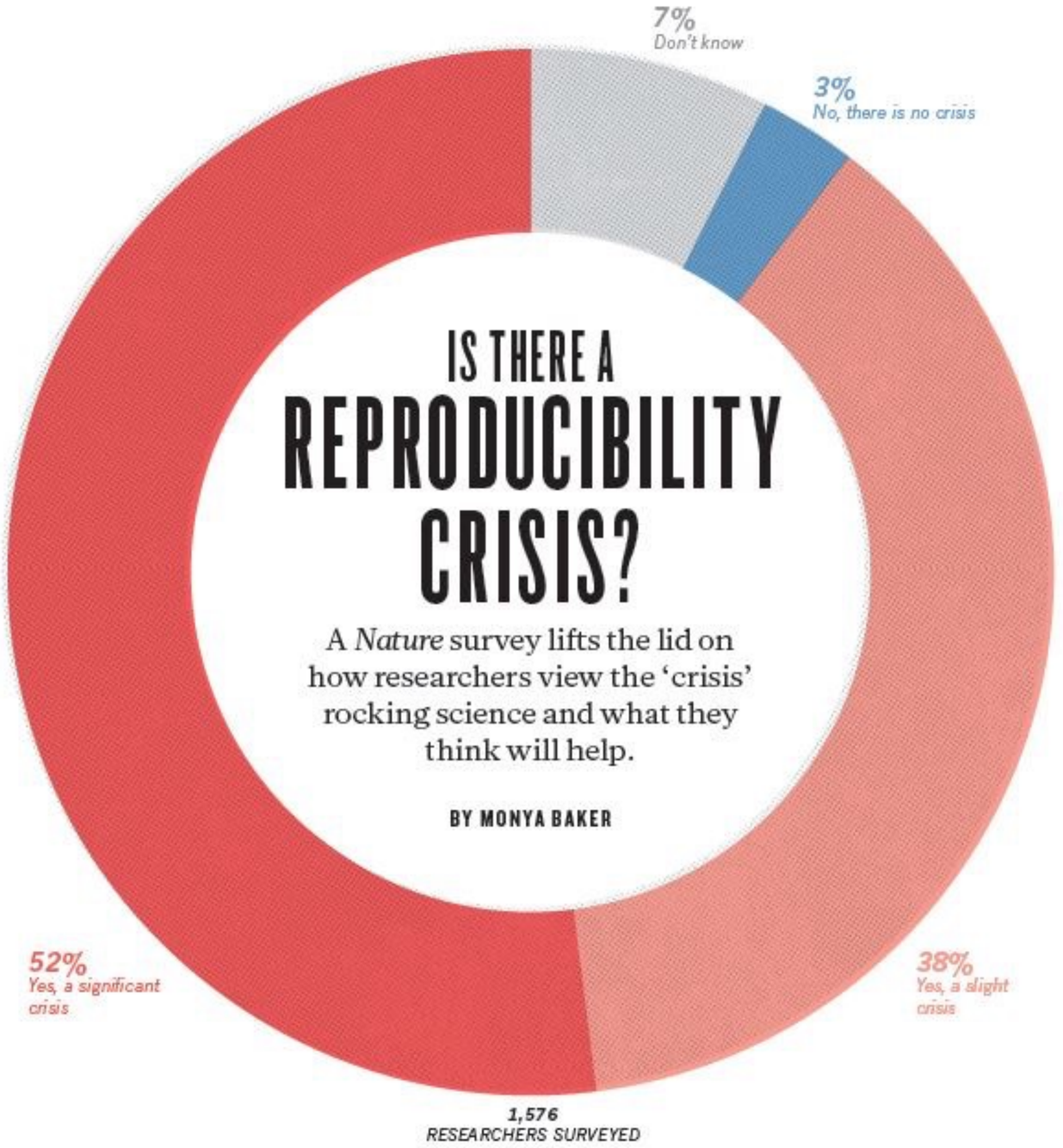


https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970

Reproducible research is still a challenge

Rich FitzJohn | Matt Pennell | Amy Zanne | Will Cornwell

June 9, 2014

https://ropensci.org/blog/2014/06/09/reproducibility

why is this still hard?

complexity

why is this still hard?

tooling

why is this still hard?

incentives + motivation

nobody cares

(Apart from people doing it out of the goodness of their hearts of course)

nobody cares?

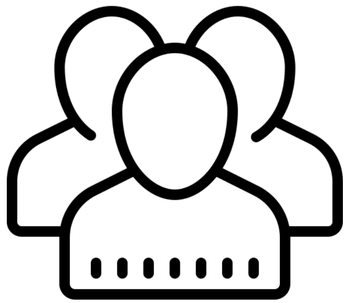
(Apart from people doing it out of the goodness of their hearts of course)

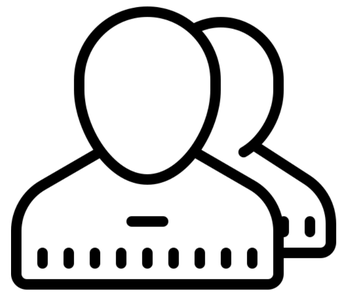
VACCINE IMPACT
MODELLING CONSORTIUM

BILL & MELINDA GATES *foundation*

Gavi
The Vaccine Alliance

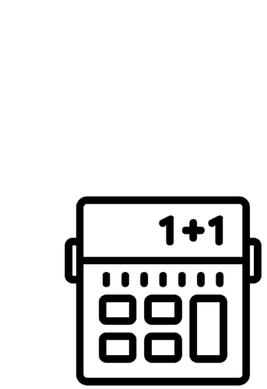
https://vaccineimpact.org

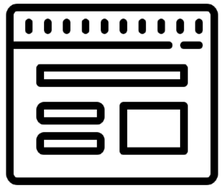


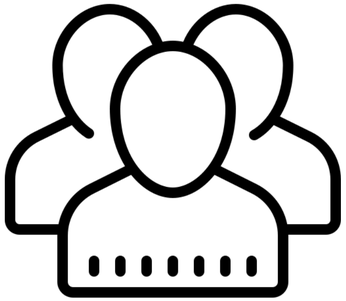


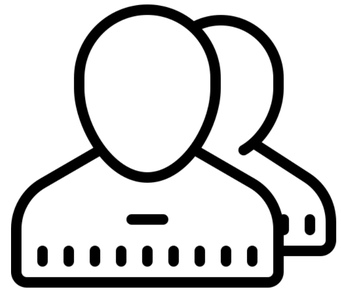
BILL & MELINDA GATES foundation

Gavi The Vaccine Alliance



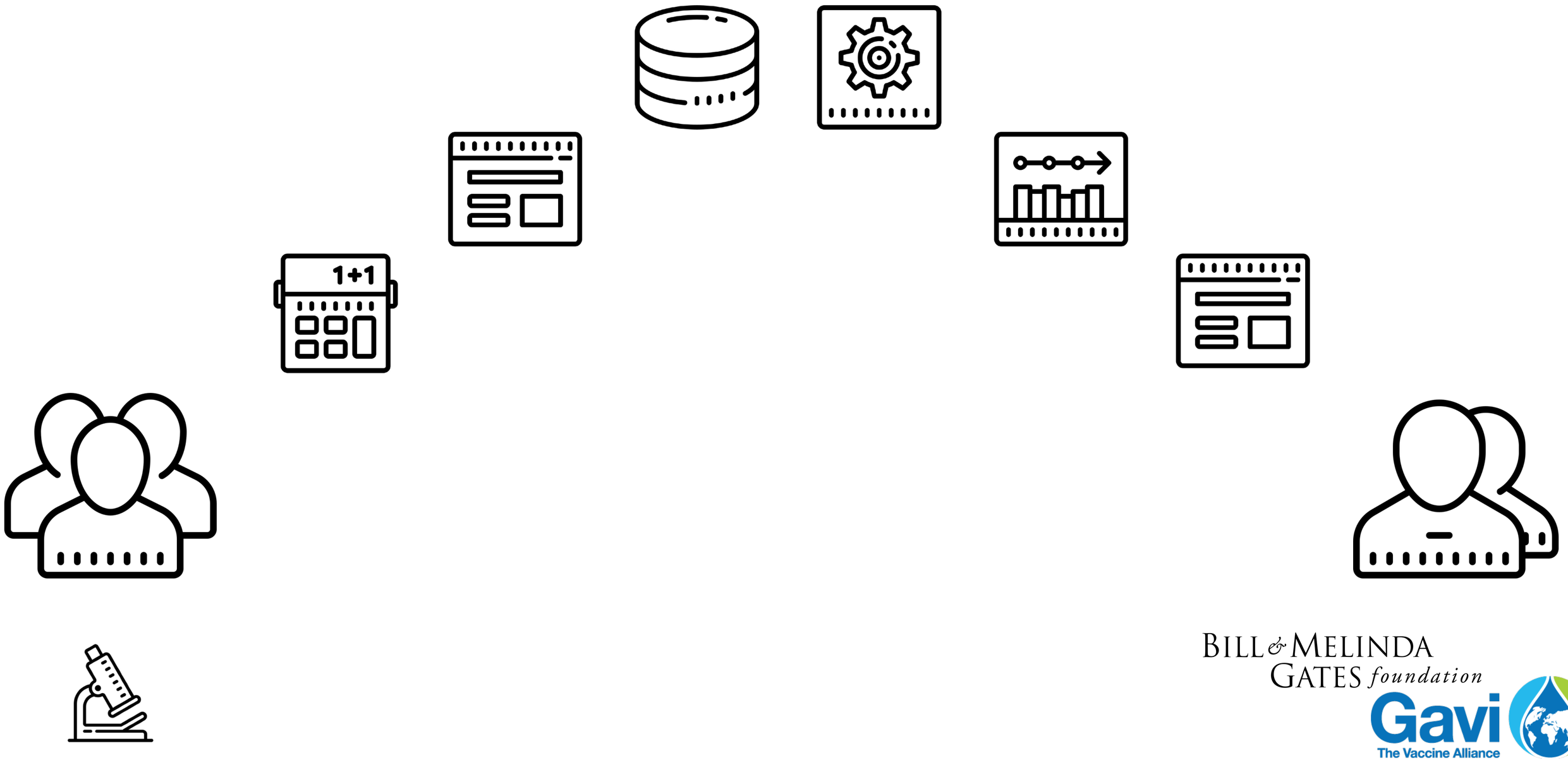






BILL & MELINDA GATES foundation

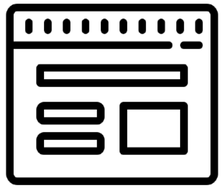
Gavi The Vaccine Alliance

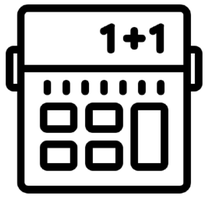


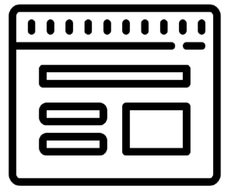
BILL & MELINDA GATES foundation

Gavi
The Vaccine Alliance

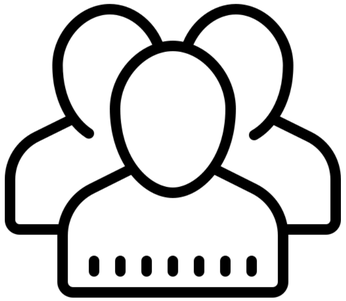
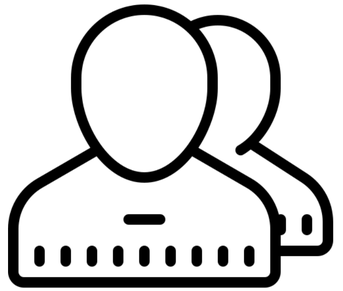











BILL & MELINDA GATES foundation

Gavi
The Vaccine Alliance

Our goal

Code written by domain scientists runs on multiple machines, first time

Trace back to origins of changes in any report

knitr

www.rstudio.com

rmarkdown

www.rstudio.com

RStudio

Go to file/function | Addins | Project

example.Rmd

Knit | Insert | Run

```
1  ---
2  title: "Example"
3  output: html_document
4  ---
5
6  ```{r setup, include=FALSE}
7  knitr::opts_chunk$set(echo = TRUE)
8  ```
9
10 ## R Markdown
11
12 This is an R Markdown document. Markdown is a simple formatting
   syntax for authoring HTML, PDF, and MS Word documents. For more
   details on using R Markdown see <http://rmarkdown.rstudio.com>.
13
14 When you click the **Knit** button a document will be generated
   that includes both content as well as the output of any
   embedded R code chunks within the document. You can embed an R
   code chunk like this:
15
16 ```{r cars}
17 summary(cars)
18 ```
19
20 ## Including Plots
21
22 You can also embed plots, for example:
23
24 ```{r pressure, echo=FALSE}
25 plot(pressure)
26 ```
27
28 Note that the `echo = FALSE` parameter was added to the code
   chunk to prevent printing of the R code that generated the
   plot.
29
```

2:16 | Example | R Markdown

Console

~/example.html

example.html | Open in Browser | Find | Publish

Example

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

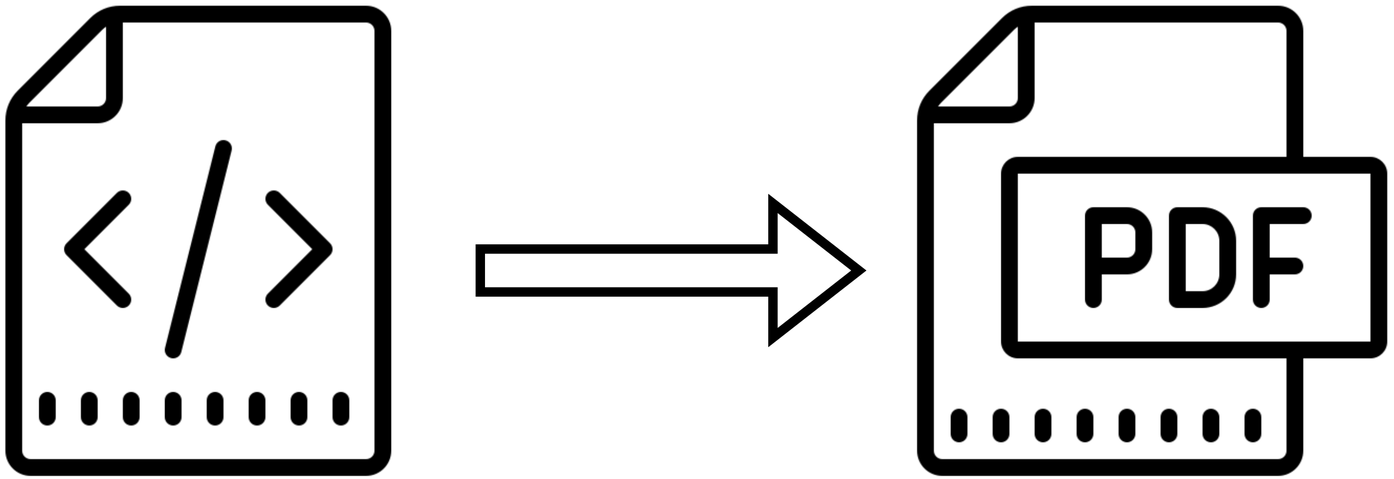
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

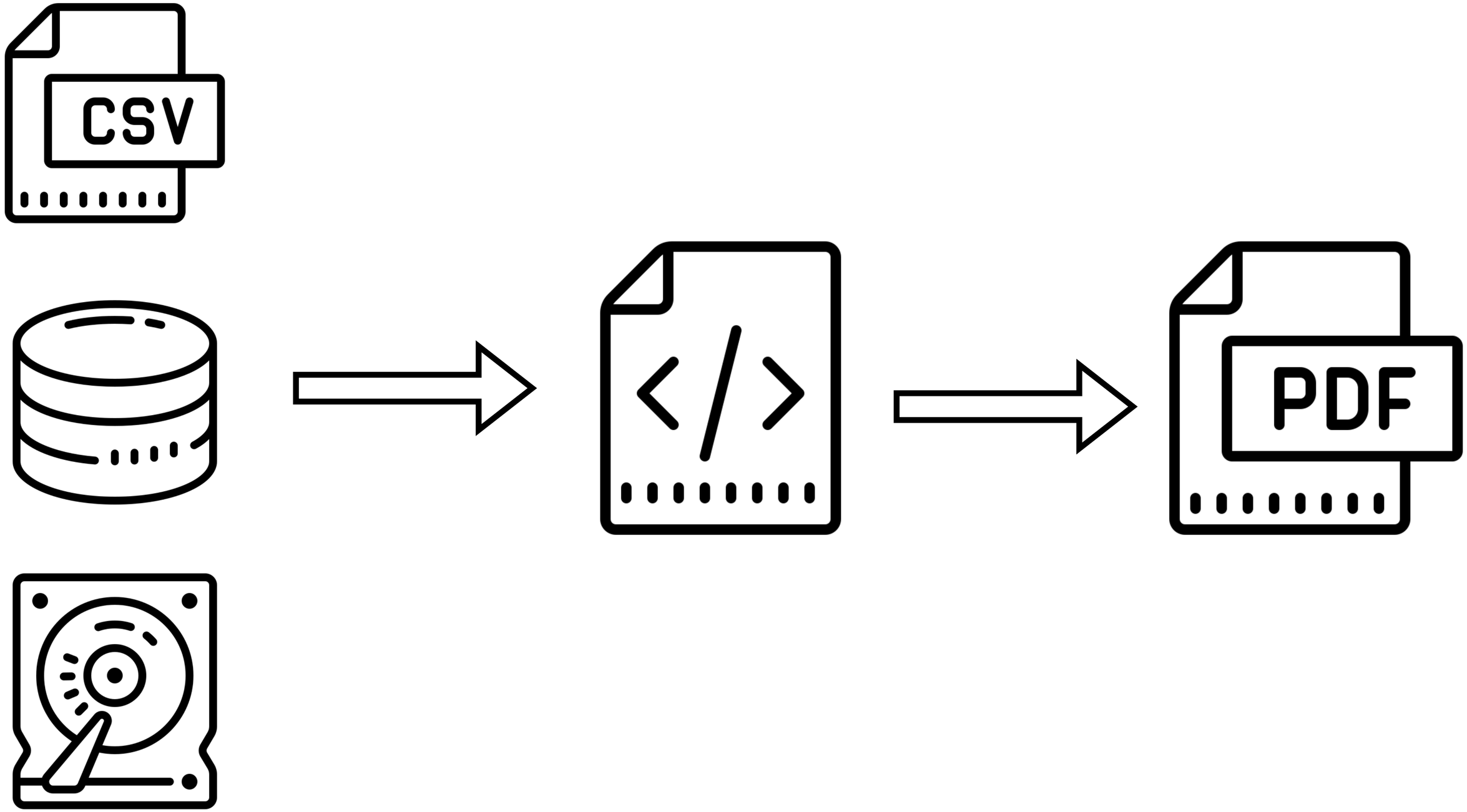
```
summary(cars)
```

```
##      speed           dist
##  Min.   : 4.0   Min.   :  2.00
##  1st Qu.:12.0   1st Qu.: 26.00
##  Median :15.0   Median : 36.00
##  Mean   :15.4   Mean   : 42.98
##  3rd Qu.:19.0   3rd Qu.: 56.00
##  Max.   :25.0   Max.   :120.00
```

Including Plots

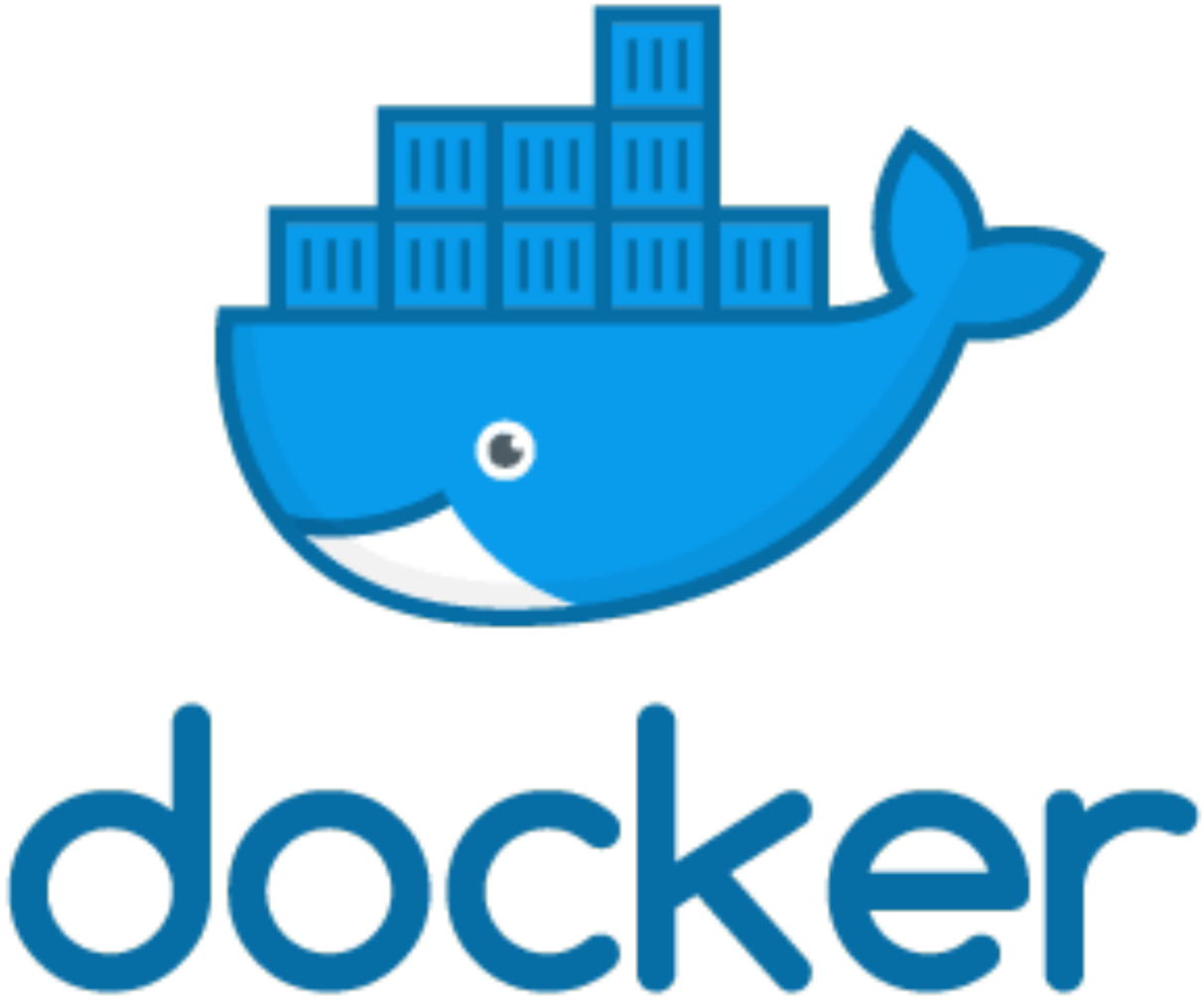
You can also embed plots, for example:

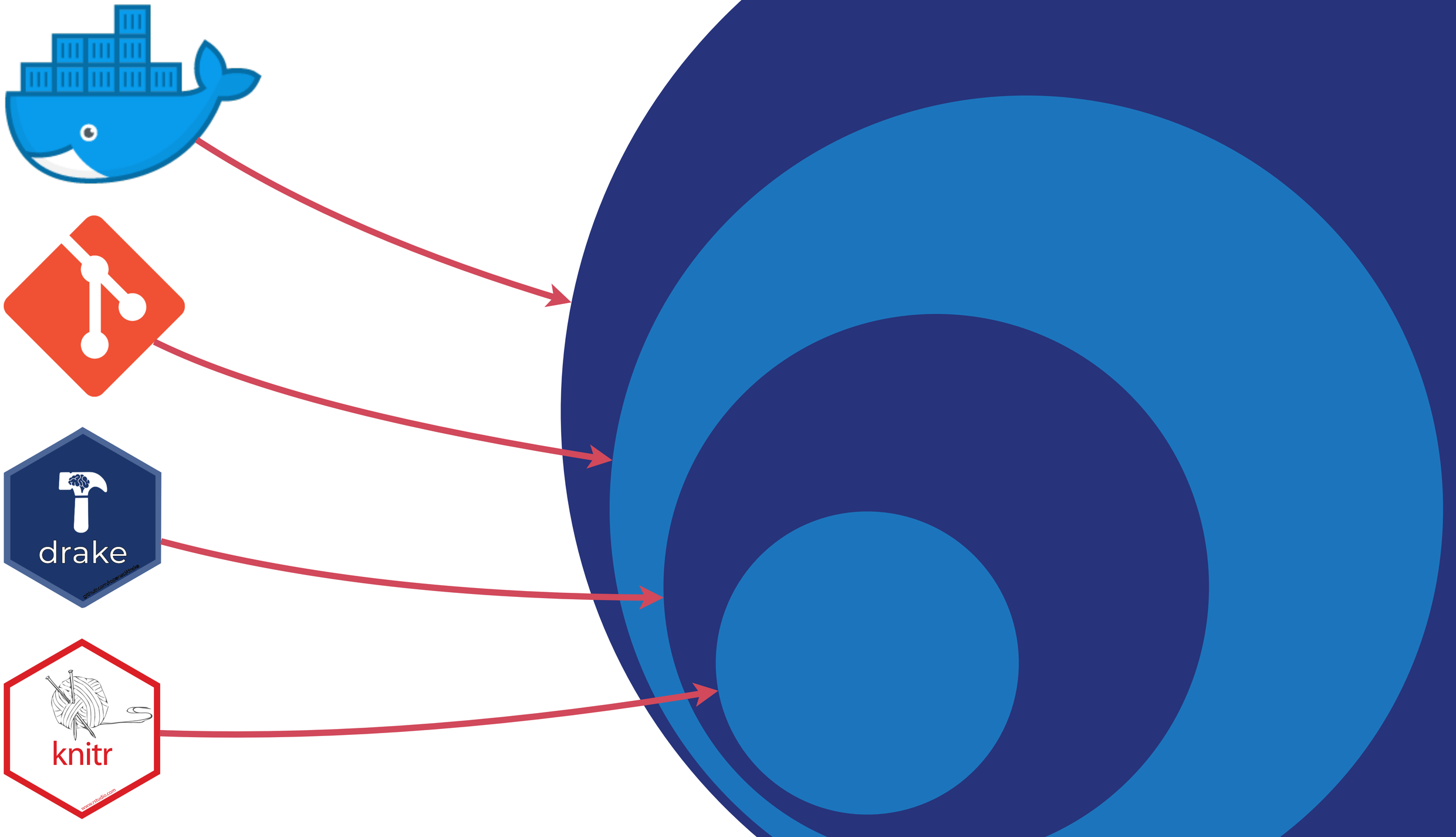

PDF


CSV
PDF


drake
github.com/ropensci/drake


git


docker


drake
knitr

Go To Statement Considered Harmful

Key Words and Phrases: go to statement, jump instruction, branch instruction, conditional clause, alternative clause, repetitive clause, program intelligibility, program sequencing
CR Categories: 4.22, 5.23, 5.24

EDITOR:

For a number of years I have been familiar with the observation that the quality of programmers is a decreasing function of the density of **go to** statements in the programs they produce. More recently I discovered why the use of the **go to** statement has such disastrous effects, and I became convinced that the **go to** statement should be abolished from all "higher level" programming languages (i.e. everything except, perhaps, plain machine code).

```
add <- function(a, b) {
  a + b
}
```

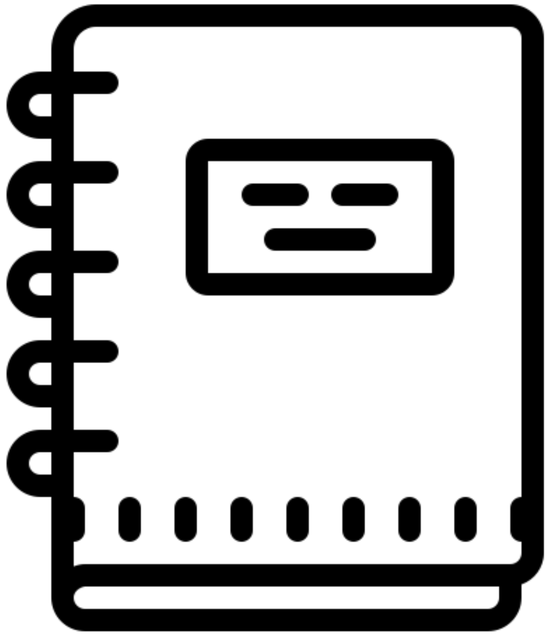
https://github.com/vimc/orderly

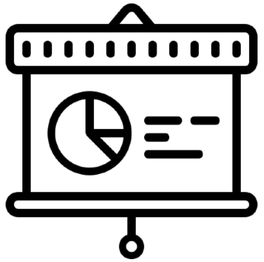
```
expect_equal(
  add(1, 3),
  4)
```

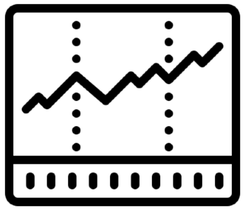
https://github.com/vimc/orderly

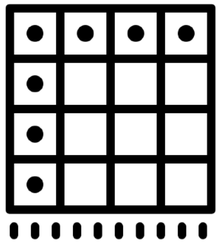
```
#' @param a,b inputs
#' @return the sum
#' @export
```

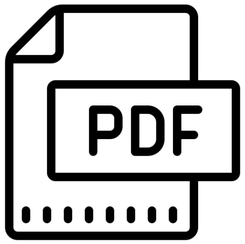
https://github.com/vimc/orderly

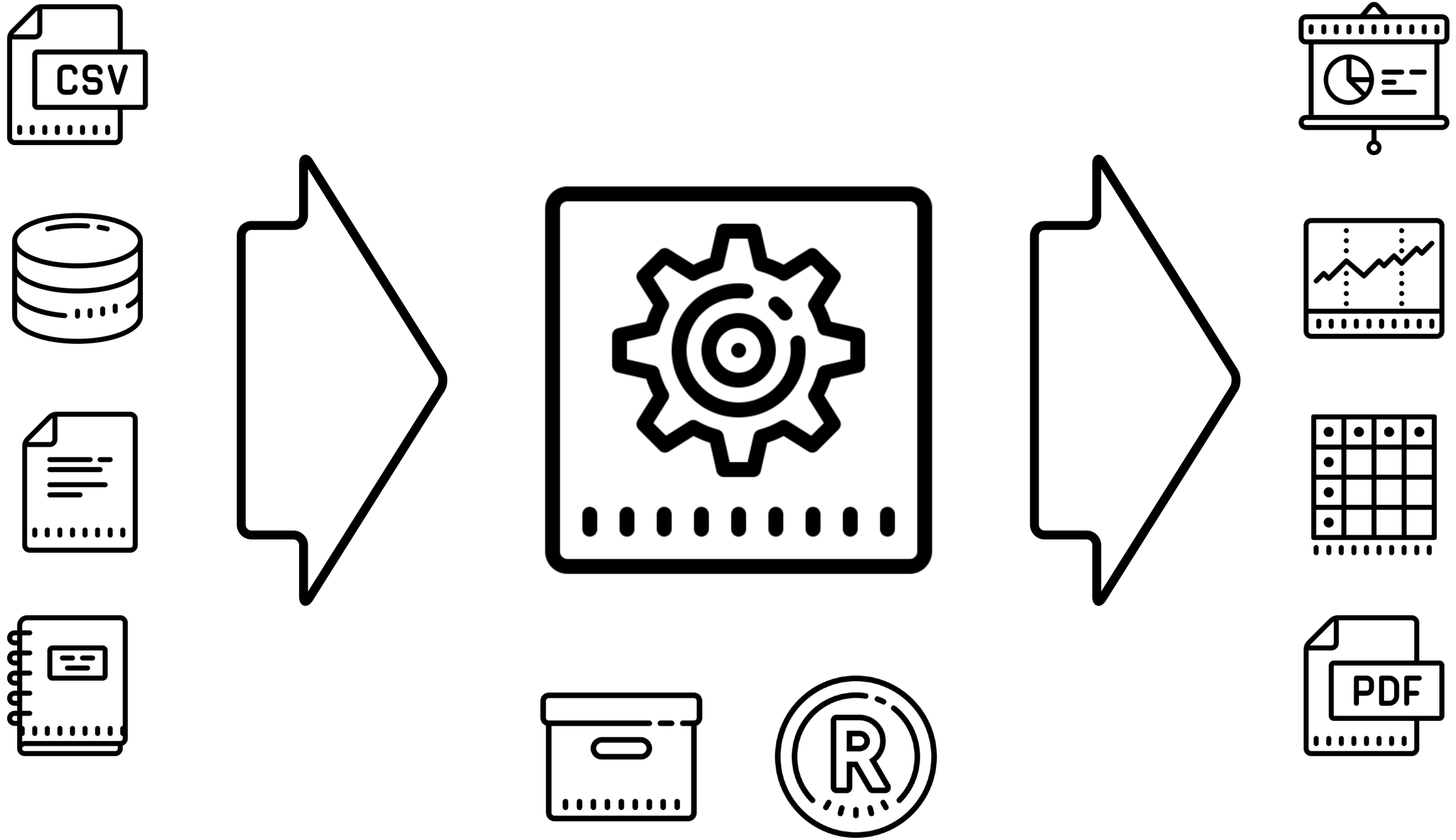









PDF


CSV
R
PDF

orderly

https://github.com/vimc/orderly

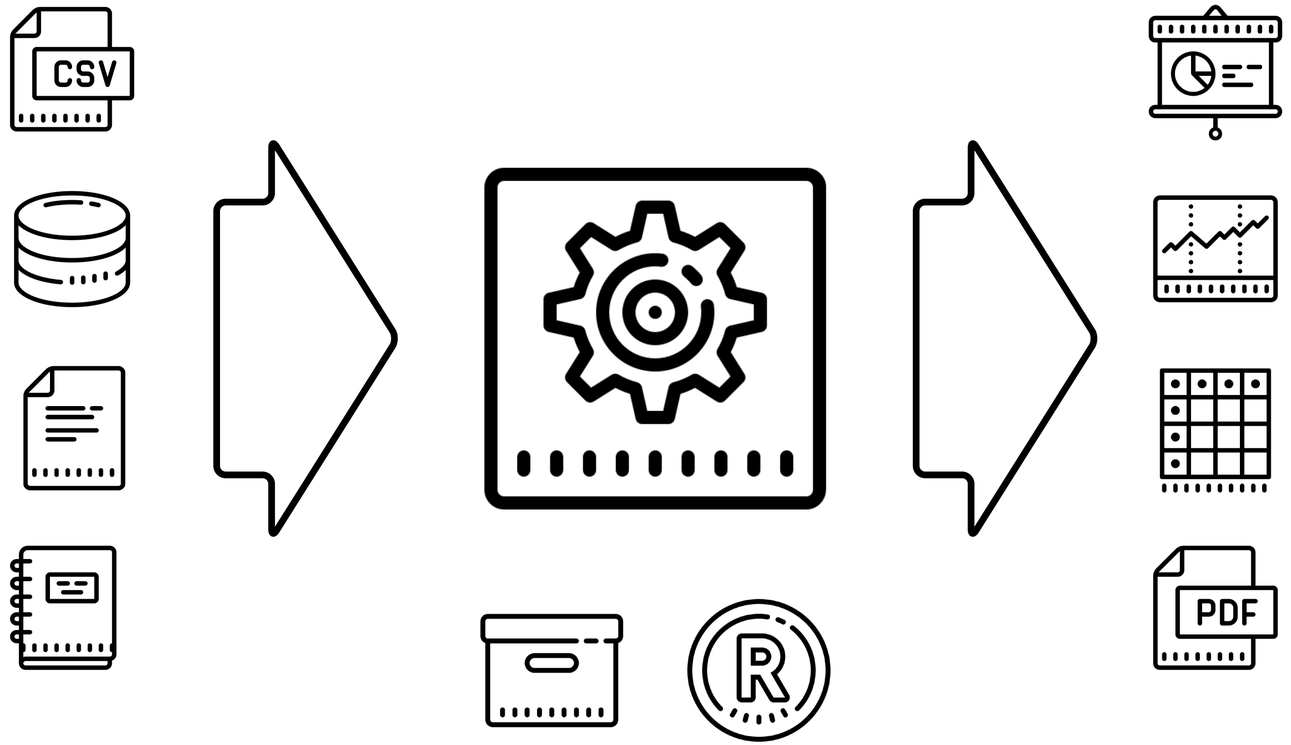
```
data:
  summary: SELECT * FROM ...

resources:
  - support.R
  - metadata.csv

packages:
  - ggplot2
  - knitr

script: script.R

artefacts:
  - report:
      description: Summary of results
      filenames: summary.pdf
  - data:
      description: Processed data for further use
      filenames: data.csv
```



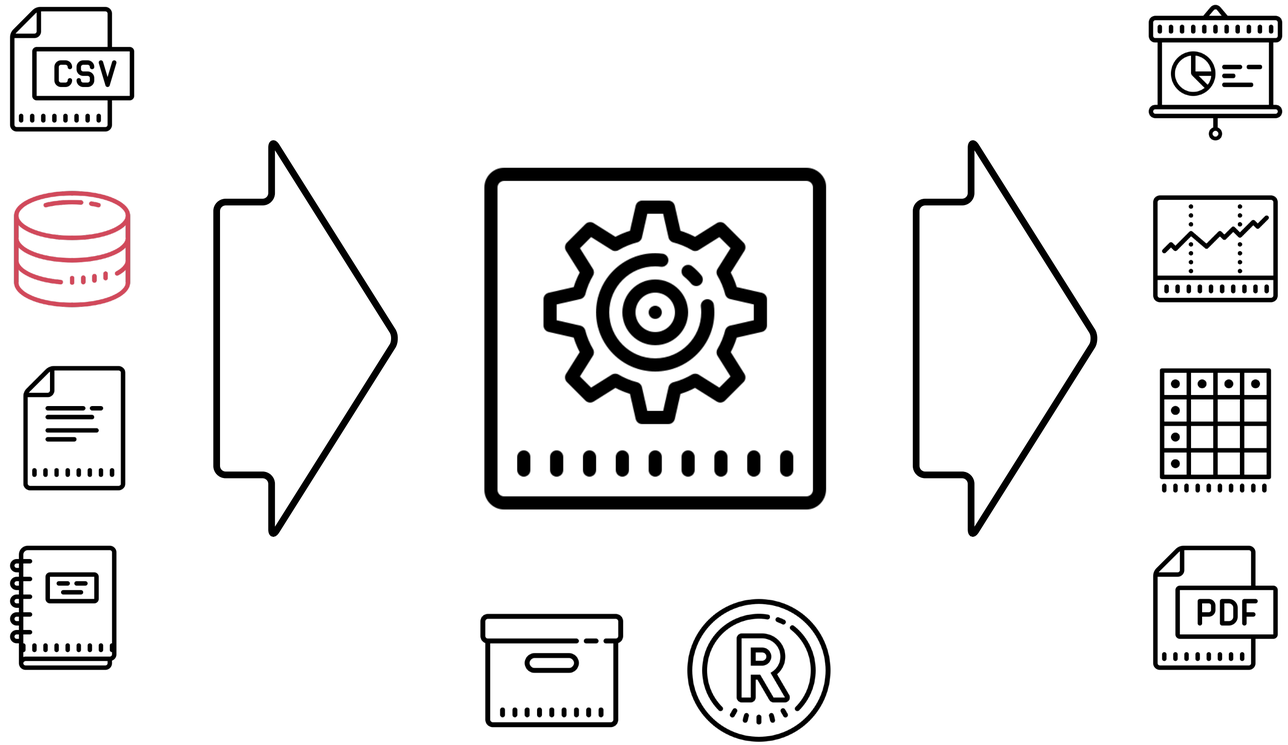
```
data:
  summary: SELECT * FROM ...

resources:
  - support.R
  - metadata.csv

packages:
  - ggplot2
  - knitr

script: script.R

artefacts:
  - report:
      description: Summary of results
      filenames: summary.pdf
  - data:
      description: Processed data for further use
      filenames: data.csv
```



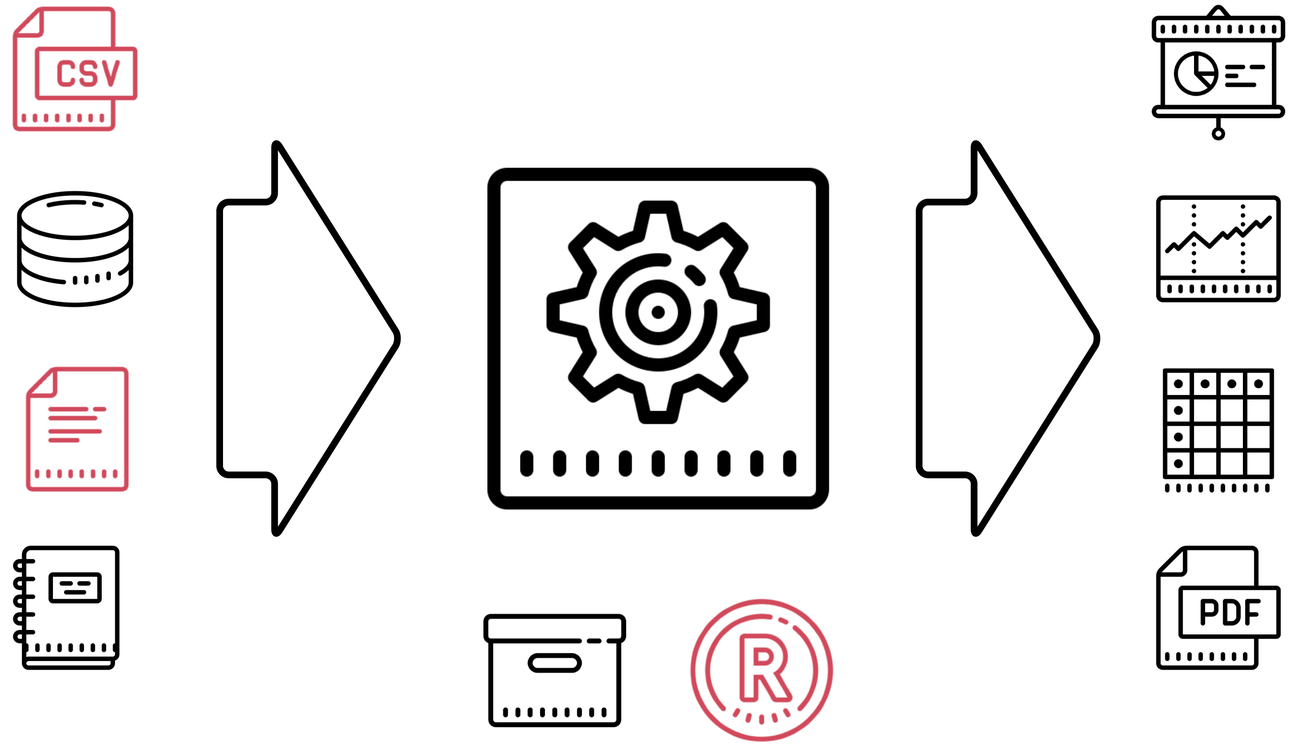
```yaml
data:
  summary: SELECT * FROM ...

resources:
  - support.R
  - metadata.csv

packages:
  - ggplot2
  - knitr

script: script.R

artefacts:
  - report:
      description: Summary of results
      filenames: summary.pdf
  - data:
      description: Processed data for further use
      filenames: data.csv
```



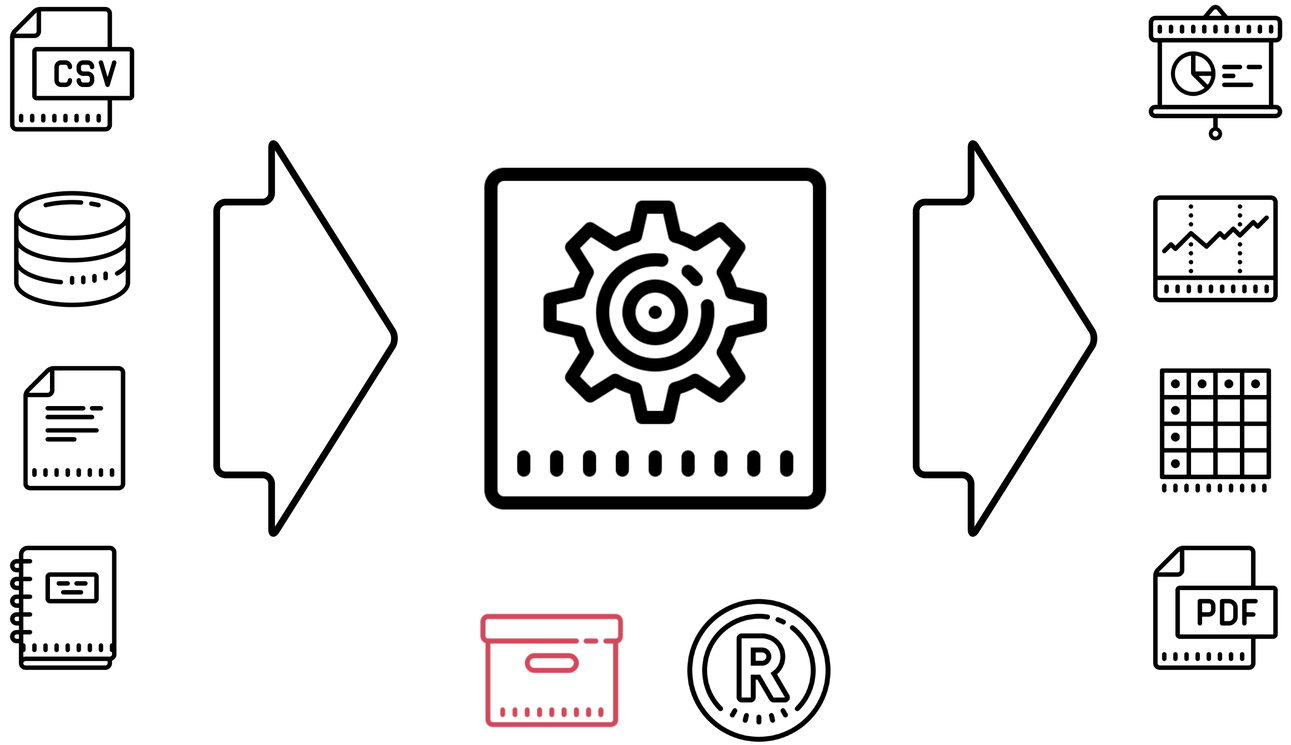
```
data:
  summary: SELECT * FROM ...

resources:
  - support.R
  - metadata.csv

packages:
  - ggplot2
  - knitr

script: script.R

artefacts:
  - report:
      description: Summary of results
      filenames: summary.pdf
  - data:
      description: Processed data for further use
      filenames: data.csv
```



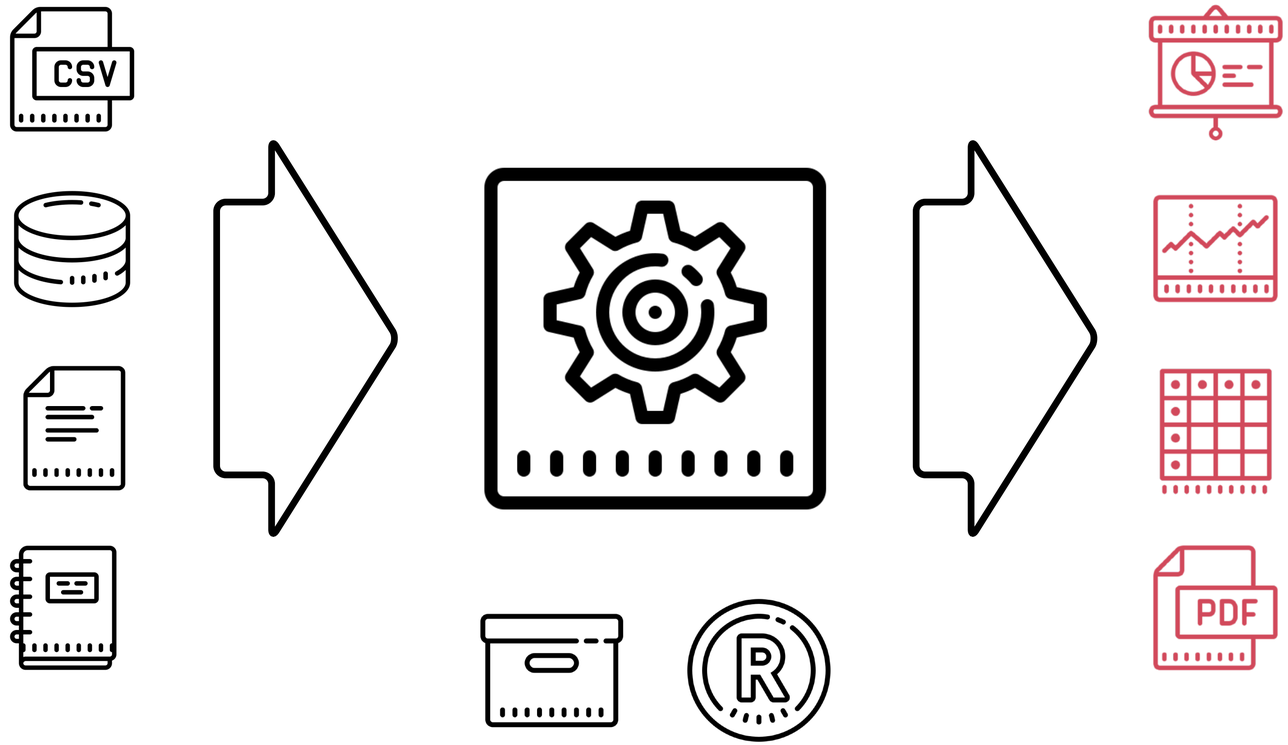
```yaml
data:
  summary: SELECT * FROM ...

resources:
  - support.R
  - metadata.csv

packages:
  - ggplot2
  - knitr

script: script.R

artefacts:
  - report:
      description: Summary of results
      filenames: summary.pdf
  - data:
      description: Processed data for further use
      filenames: data.csv
```



```
orderly_config.yml
src/
  myreport/
    orderly.yml
    script.R
    support.R
    metadata.csv
```

```
orderly_config.yml
src/
  myreport/
    orderly.yml
    script.R
    support.R
    metadata.csv
archive/
  myreport/
    20190204-143204-f5aa3bc9/
      orderly.yml
      script.R
      support.R
      metadata.csv
      summary.pdf
      data.csv
      orderly_run.rds
```

orderly run myreport

orderly run myreport

```
orderly_config.yml
src/
  myreport/
    orderly.yml
    script.R
    support.R
    metadata.csv
archive/
  myreport/
    20190204-143204-f5aa3bc9/
    20190204-192249-3bc9f5aa/
      orderly.yml
      script.R
      support.R
      metadata.csv
      summary.pdf
      data.csv
      orderly_run.rds
```

Automate the boring bits

This is embarrassingly simple

Interface for **Stakeholders**

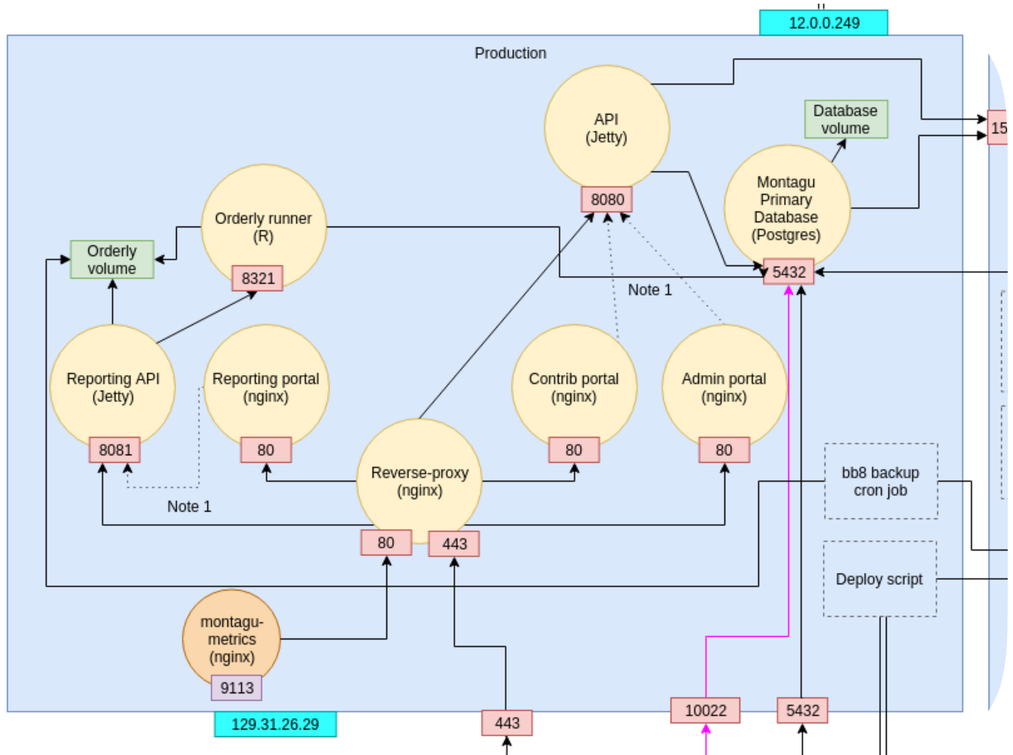
Find a report

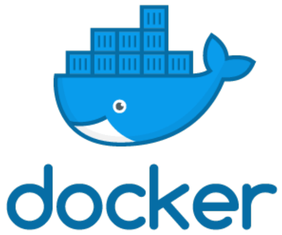
Click on a column heading to sort by that field. Hold shift to multi-sort.

Collapse all reports / Expand all reports

Name	Version	Status	Author	Requester
Type to filter...	Mar 1, to Feb 6, Type to filter by id...	All	Type to filter...	Type to filter...
▼ native-diagnostics-burden-report-drafts 5 versions: view latest			Science Team	VIMC
	Tue Feb 05 2019 latest (20190205-151702-1ba5e47a)	published	Science Team	VIMC
	Thu Jan 31 2019 out-dated (20190131-162935-86a83d30)	published	Science Team	VIMC
	Thu Jan 31 2019 out-dated (20190131-123847-53fe189e)	published	Science Team	VIMC
	Mon Jan 28 2019 out-dated (20190128-151914-a2c1edce)	published	Science Team	VIMC
	Mon Jan 28 2019 out-dated (20190128-125432-31d88275)	internal	Science Team	VIMC

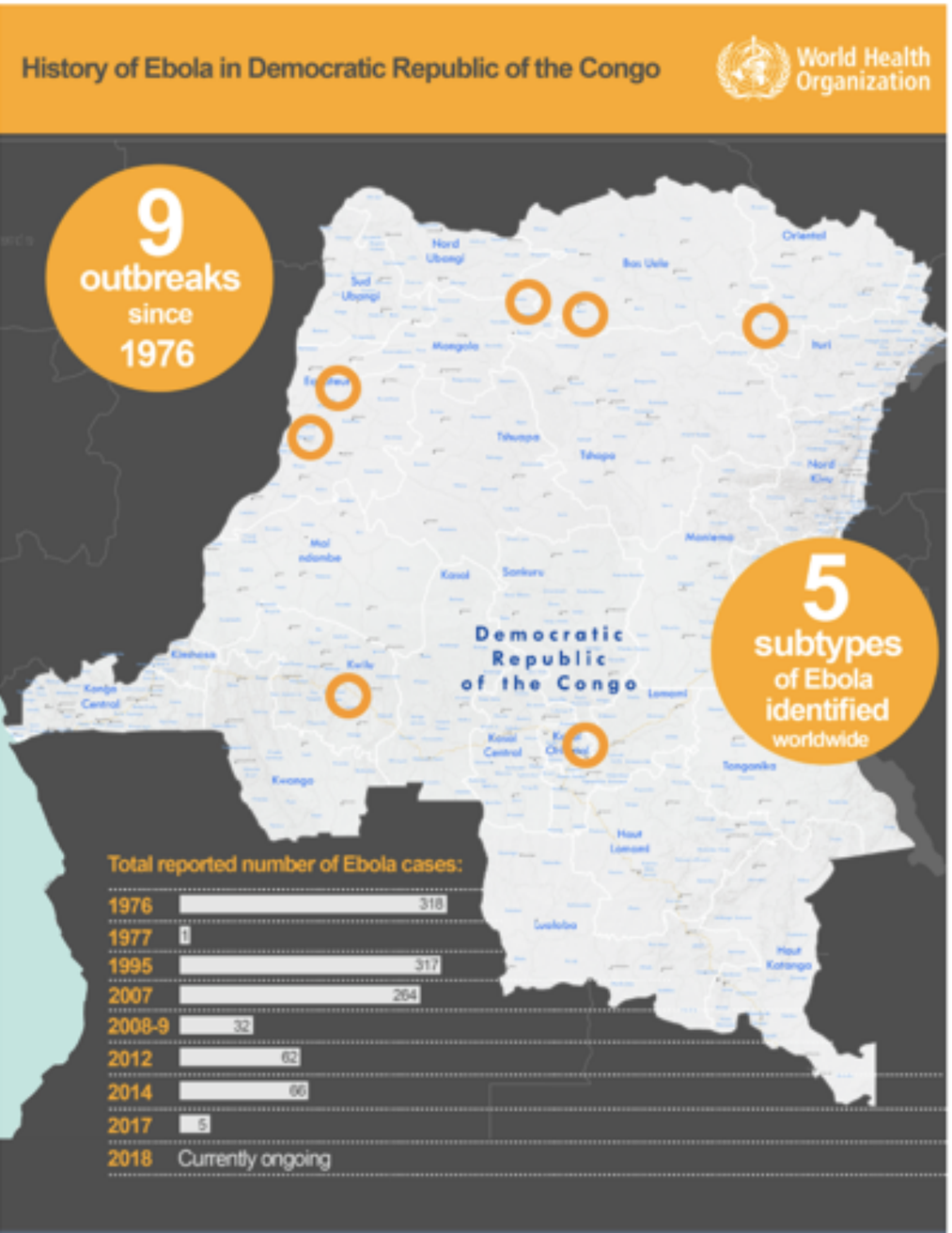
Interface for **Engineering team**





Interface for **Science team**

```
install.packages("orderly")
orderly::orderly_new("myreport")
```



Interface for **Other groups**

Lessons learnt

you can blackbox too much

code reuse is really hard

reproducibility can be easy

problems are as social as technical

Work the way people want to work

Our team












Vaccine Impact Modelling Consortium
MRC Centre for Global Infectious Disease Analysis
Department of Infectious Disease Epidemiology, Imperial College London