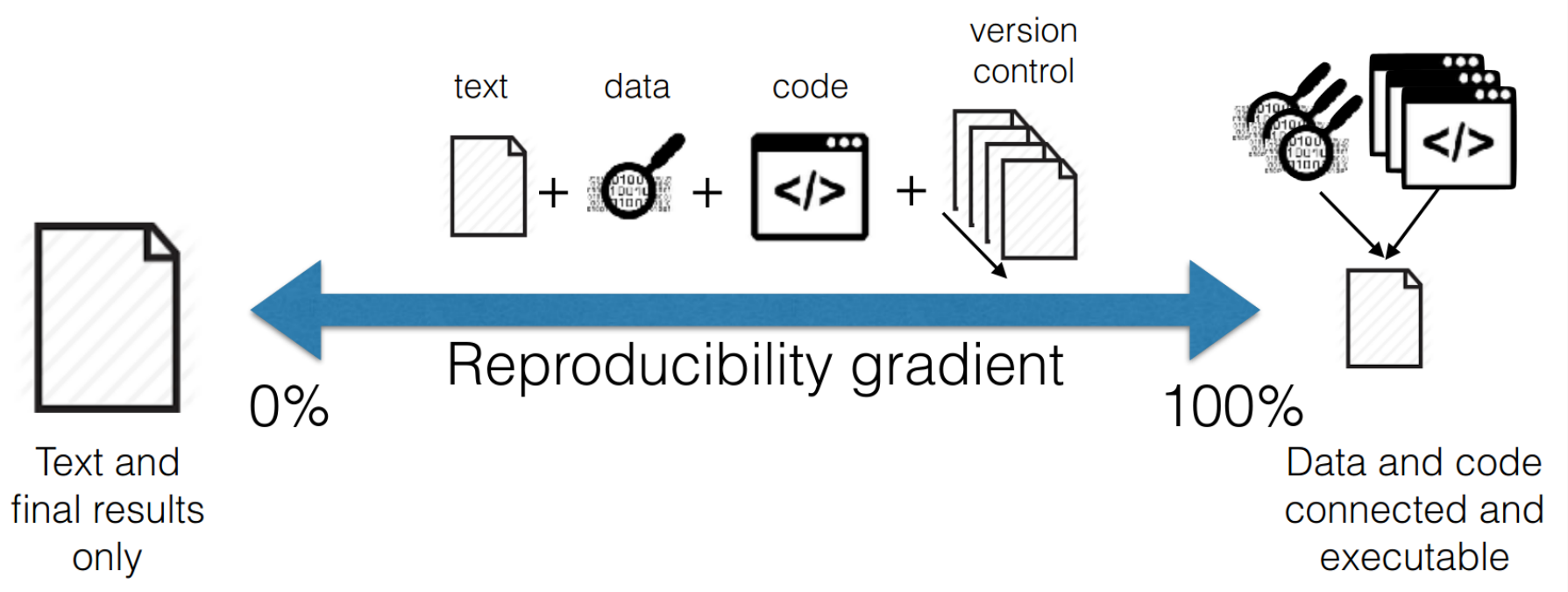
Basic reproducibility

Francisco Rodríguez-Sánchez

https://frodriguezsanchez.net

Reproducibility is a gradient



Rodríguez-Sánchez et al. 2016 (modif. Peng 2011)

Basic reproducibility

Basic reproducibility

- **MANUSCRIPT** (Text + Tables + Figures)
- **DATA** in permanent archive (see Tierney & Ram 2020)
- **CODE** in permanent archive (see Eglen et al 2016)

Permanent archive:

- Zenodo, Dryad, OSF, Figshare, Data Paper...
- NOT GitHub, website...

How to share data

- **Open** format (csv, txt)
- **README** (who, what, when, where, why, how)
- **Describe variables**
- **Licence** (CC0, CC-BY, ODbL)
- **Citation** (DOI)
- **Metadata** standardised (JSON, XML)

Tierney & Ram 2020

Document your data

```
library('dataspice')
create_spice()   # create CSV templates for metadata

edit_creators()  # open Shiny apps to edit the CSVs
prep_access()
edit_access()
prep_attributes()
edit_attributes()
edit_biblio()

write_spice()  # write machine-readable metadata

build_site()  # build human-readable metadata report
```

How to share code

- Scripts: **plain text** (**.R**)
- **Permanent archive** (eg. Zenodo) with **DOI (citable)**
- Licence
- README
- Computational **environment** (session info)

Eglen et al 2016

`sessionInfo` records OS & used packages

```
## R version 4.4.2 (2024-10-31)
## Platform: x86_64-pc-linux-gnu
## Running under: Ubuntu 20.04.6 LTS
##
## Matrix products: default
## BLAS:   /usr/lib/x86_64-linux-gnu/openblas-pthread/libblas.so.3
## LAPACK: /usr/lib/x86_64-linux-gnu/openblas-pthread/liblapack.so.3;  LAPACK version 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
##
## time zone: Europe/Madrid
## tzcode source: system (glibc)
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] knitr_1.49
##
## loaded via a namespace (and not attached):
##  [1] compiler_4.4.2    fastmap_1.2.0     cli_3.6.3         htmltools_0.5.8.1
##  [5] tools_4.4.2       rstudioapi_0.17.1 yaml_2.3.10       codetools_0.2-20
##  [9] rmarkdown_2.29    binb_0.0.7        xfun_0.50         digest_0.6.37
```

renv also records packages used

```
renv::snapshot()
```

creates **renv.lock**file recording dependencies.

Can use **renv::restore()** to restore packages later or in different computer.

Basic reproducibility

DATA + CODE

- analysis fully **traceable**
- results can be **regenerated**