

Data management

Francisco Rodriguez-Sanchez

@frod_san

<https://frodriguezsanchez.net>

See <https://dataoneorg.github.io/Education/bestpractices/>

- .
- 1. Planification (e.g. DMPTool)
- 2. Collection
- 3. Metadata description (dataspice, EML, Data Packages, DataPackageR)
- 4. Quality control (e.g. assertr, validate, pointblank)
- 5. Storage

Document your data

<https://docs.ropensci.org/dataspice/>

```
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
```

Check data before analysis

```
library("assertr")

dataset %>%
  assert(within_bounds(0, 0.20), fruit.weight) %>%
  assert(in_set("black", "red"), colour)
```

Check out also [pointblank](#)

Editorial expression of concern

IN THE 3 June issue, *Science* published the Report “Environmentally relevant concentrations of microplastic particles influence larval fish ecology” by Oona M. Lönnstedt and Peter Eklöv (*1*). The authors have notified *Science* of the theft of the computer on which the raw data for the paper were stored. These data were not backed up on any other device nor deposited in an appropriate repository. *Science* is publishing this Editorial Expression of Concern to alert our readers to the fact that no further data can be made available, beyond those already presented in the paper and its supplement, to enable readers to understand, assess, reproduce, or extend the conclusions of the paper.

Jeremy Berg

Editor in Chief

Use the **cloud**: safe, persistent, easy to share

- [Open Science Framework](#)
- GitHub
- Dropbox
- Figshare, Zenodo, etc
- See all data repositories in www.re3data.org