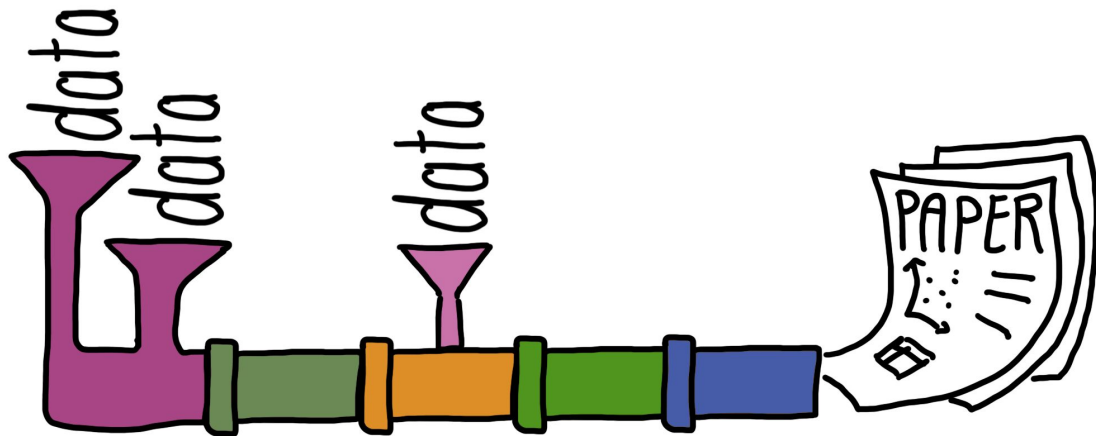


Reproducible Data Analysis

Heidi Seibold,
Open Science trainer / consultant / moderator
[@HeidiBaya](#)



Reproducible data analysis

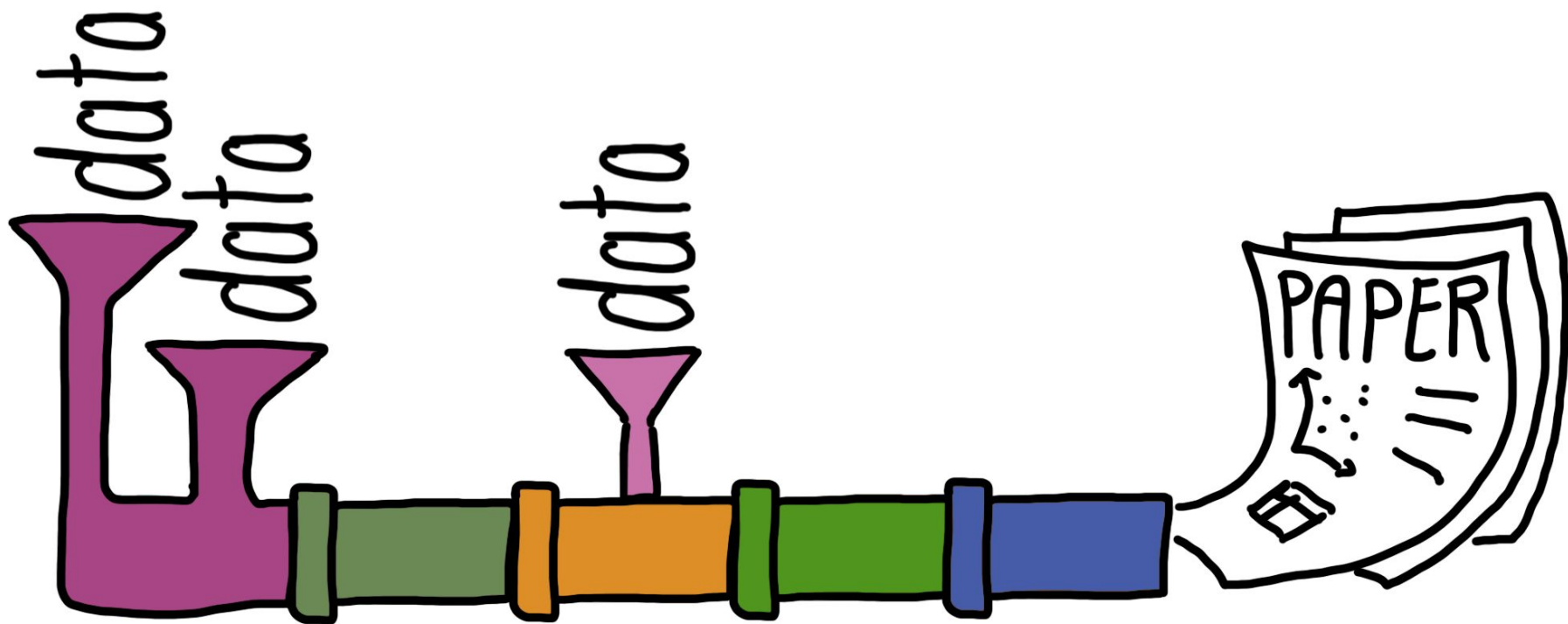


Obligatory:

1. Use code (no clicking)
2. Document / automate your steps
3. Stay organized

Recommended:

- Use version control



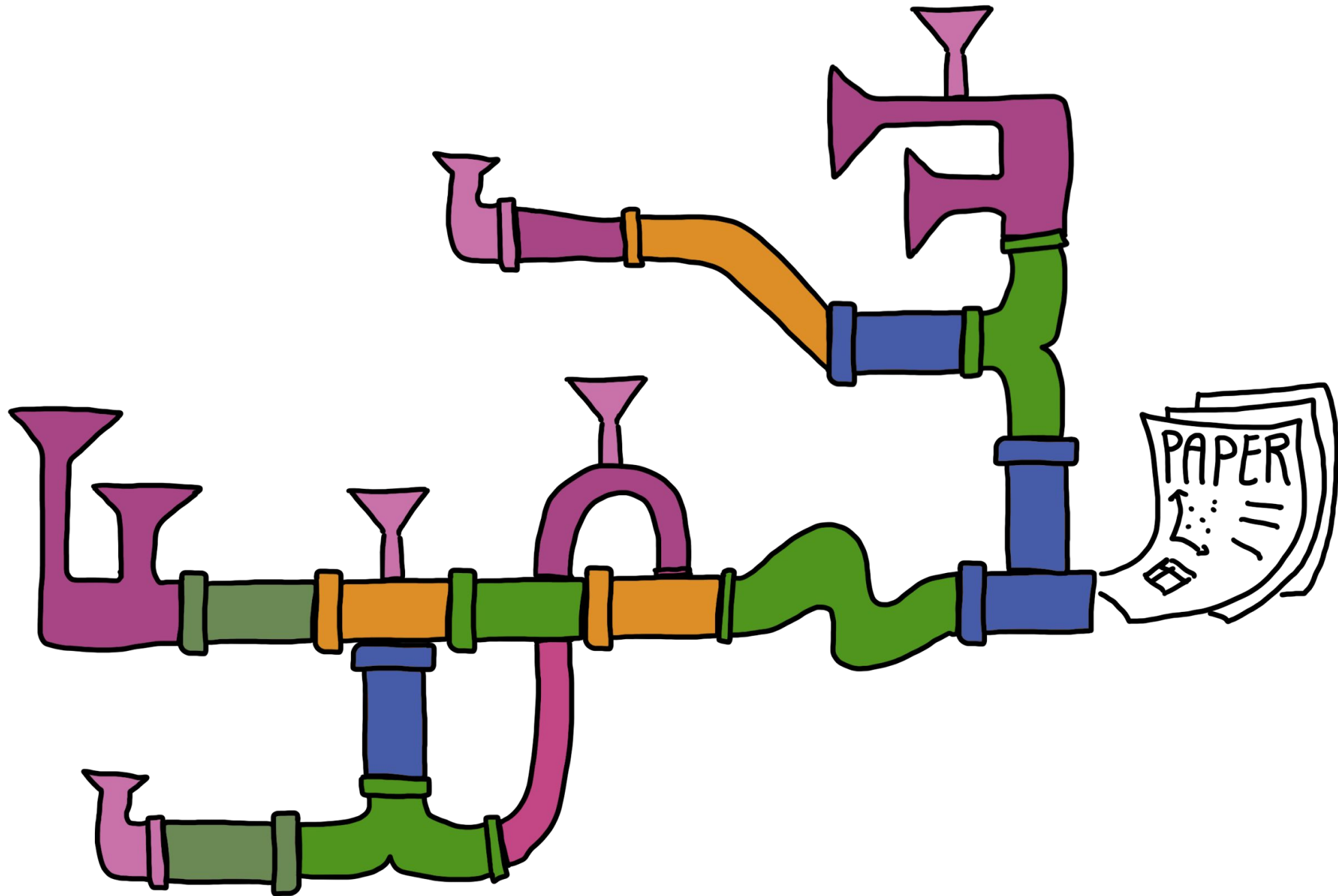
data
cleaning

overview

figures

modelling

text





Scriberia 

Illustration from The Turing Way
<https://doi.org/10.5281/zenodo.6560477>

Example toolkit



L^AT_EX

Make



```
├─ Makefile
├─ references_topic.bib
├─ manuscript_topic.Rnw
├─ abstract.tex
├─ introduction.tex
├─ methods.tex
├─ application.Rnw
├─ simulation.Rnw
├─ discussion.tex
├─ journal_correspondence
│ ├─ journal2
│ │ ├─ response1.txt
│ │ └─ p2p_1.tex
│ └─ journal1
│   └─ response.txt
├─ analysis
│ ├─ Makefile
│ ├─ try_stuff.Rmd
│ ├─ simulation.Rnw
│ ├─ simulation.R
│ ├─ analysis_study1.R
│ └─ analysis_study2.R
├─ data
│ ├─ clean_data.R
│ ├─ data_study2.csv
│ ├─ data_study1.csv
│ └─ simulation_results.RData
├─ data_original
│ ├─ data_study1_original.csv
│ └─ data_study2_original.csv
├─ basis
│ ├─ functions_simulation.R
│ └─ functions_analyses.R
```

LATEX + R = knitr

Example: my paper “model4you”

DOI: [10.5334/jors.219](https://doi.org/10.5334/jors.219)

The estimates and confidence intervals of this model can be computed via

```
<<>>=  
cbind(estimate = coef(bmod_math), confint(bmod_math))  
@
```

The model can be visualised by plotting the estimated densities (see Figure~\ref{fig:math_bmod_vis}):

```
<<math_bmod_vis, fig.height=2.5, fig.width=5,  
out.width="0.7\\textwidth", fig.cap = "Density  
estimates of base model for the Mathematics Exam  
data.">>=  
lm_plot(bmod_math)  
@
```

The estimates and confidence intervals of this model can be computed via

```
cbind(estimate = coef(bmod_math), confint(bmod_math))
```

```
##           estimate      2.5%      97.5%  
## (Intercept) 57.600184 55.122708 60.07766  
## group2      -2.332414 -5.698108 1.03328
```

The model can be visualised by plotting the estimated densities (see Figure 1):

```
lm_plot(bmod_math)
```

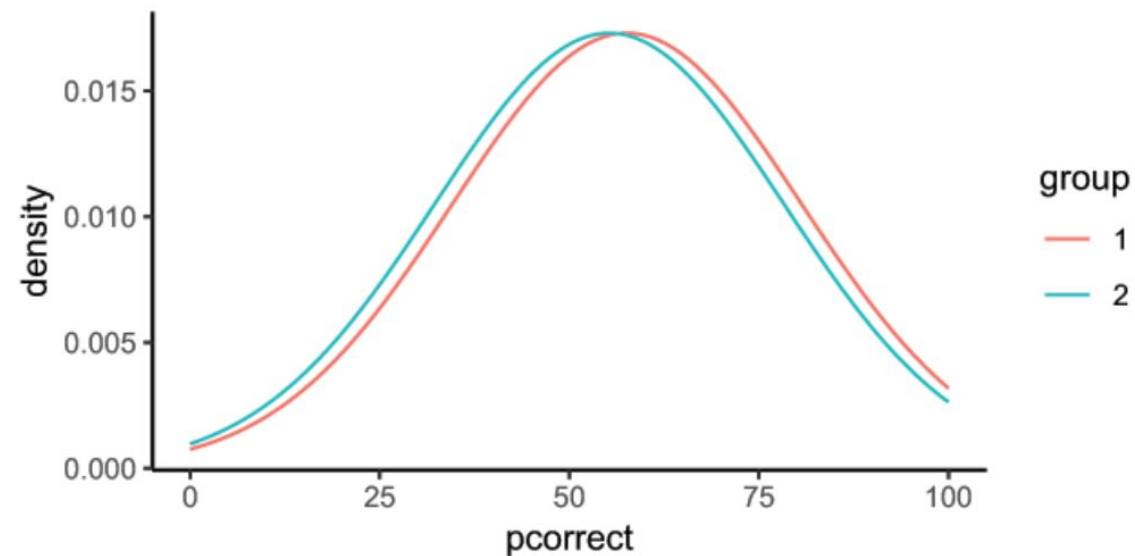
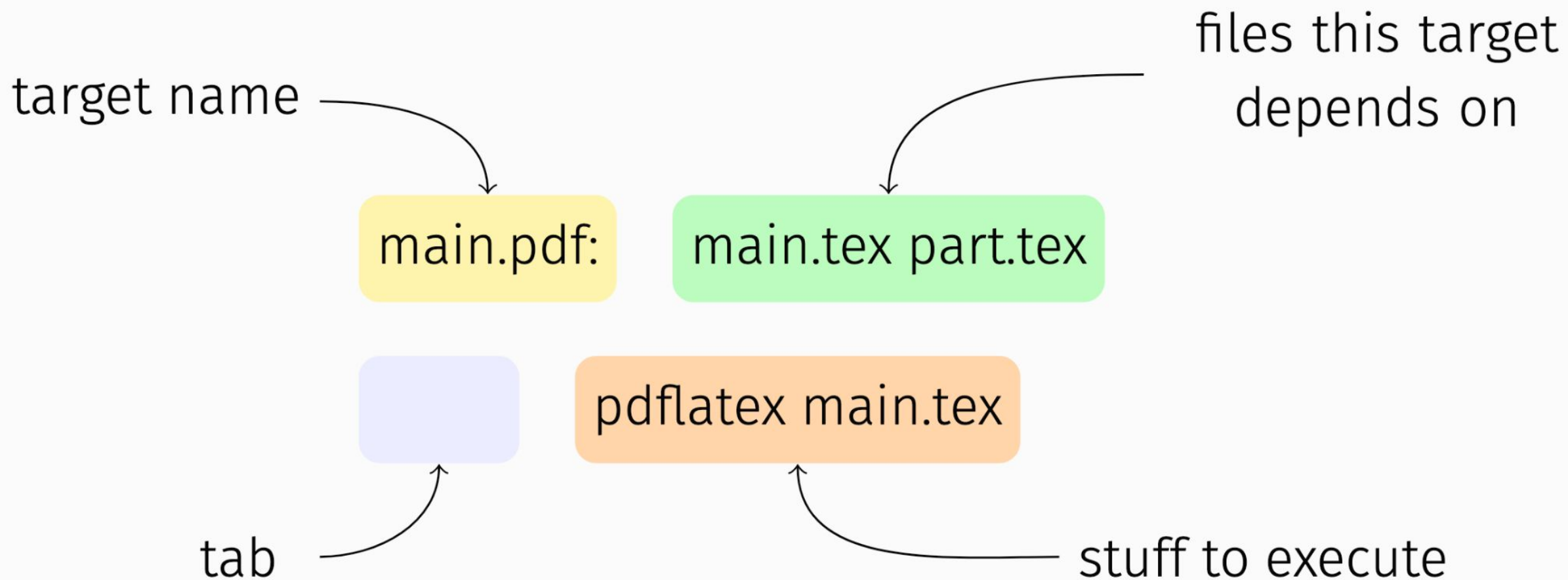


Figure 1

Density estimates of base model for the Mathematics Exam data.

Make



Make

Example: my paper “model4you”

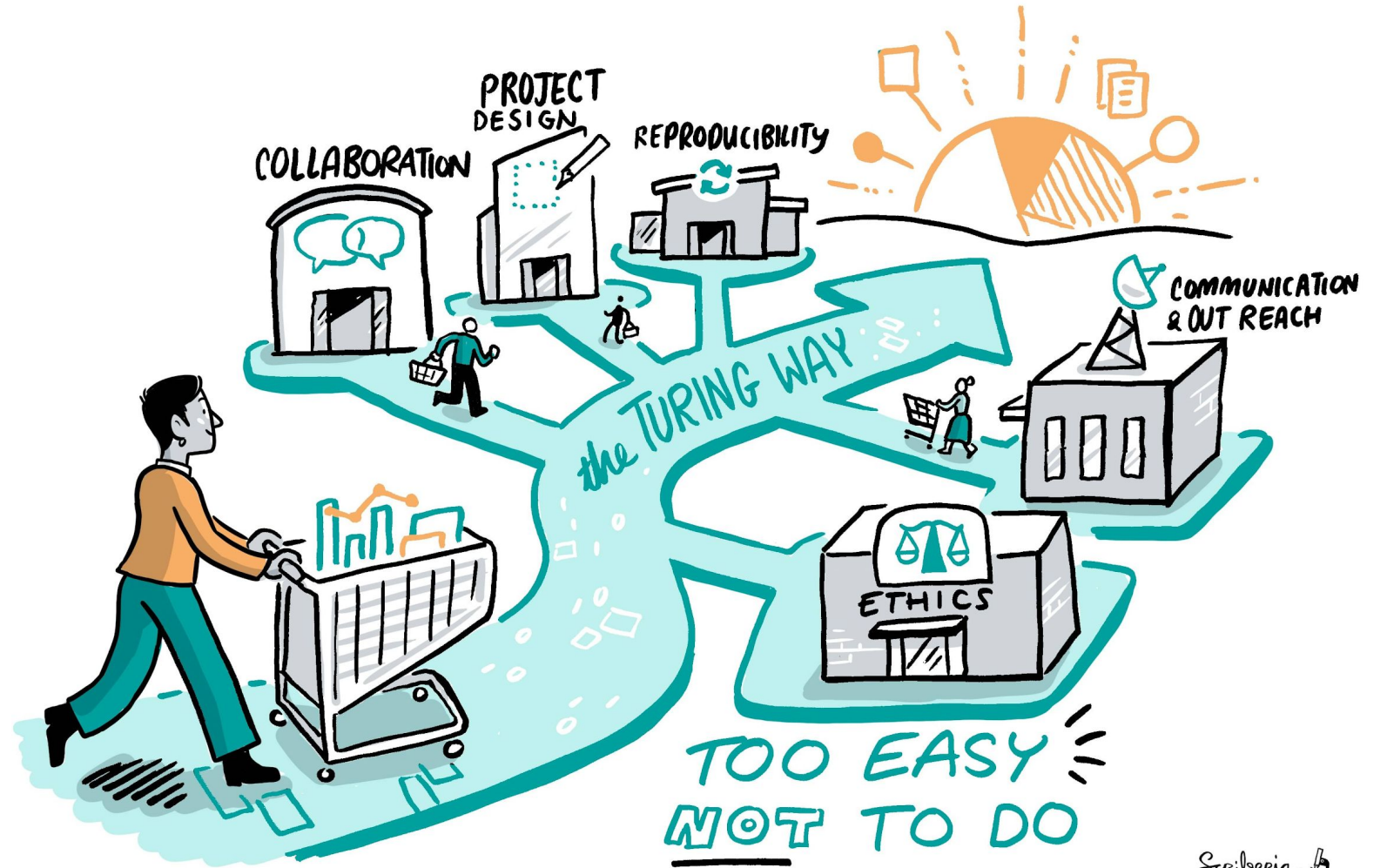
DOI: [10.5334/jors.219](https://doi.org/10.5334/jors.219)

```
man_model4you.tex: man_model4you.Rnw  
    Rscript -e "knitr::knit('man_model4you.Rnw')"  
  
man_model4you.pdf: man_model4you.tex ref_model4you.bib  
    pdflatex man_model4you.tex  
    bibtex man_model4you  
    pdflatex man_model4you.tex  
    pdflatex man_model4you.tex  
  
all: man_model4you.pdf
```

Further reading

[The Turing Way](#)

handbook to reproducible,
ethical and collaborative
data science





Thanks!

<https://heidiseibold.gitlab.io>

Heidi Seibold @HeidiBaya