

57TH ONLINE-SEMINAR

Practical Steps Towards Open and Reproducible Research

Dr. Heidi Seibold

Helmholtz Association – Helmholtz AI

WWW, February 10, 2021, from 4:00 p.m. to 5:00 p.m. CEST

Practical Steps Towards Open and Reproducible Research

Heidi Seibold / @HeidiBaya

Open AI in Health

HELMHOLTZAI | ARTIFICIAL INTELLIGENCE
COOPERATION UNIT

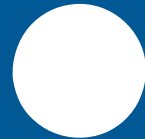
Helmholtz Open Science
[Seminar #57](#)

Slides and drawings are
licensed under CC-BY 4.0





Please ask questions in the
cryptpad



The Economist

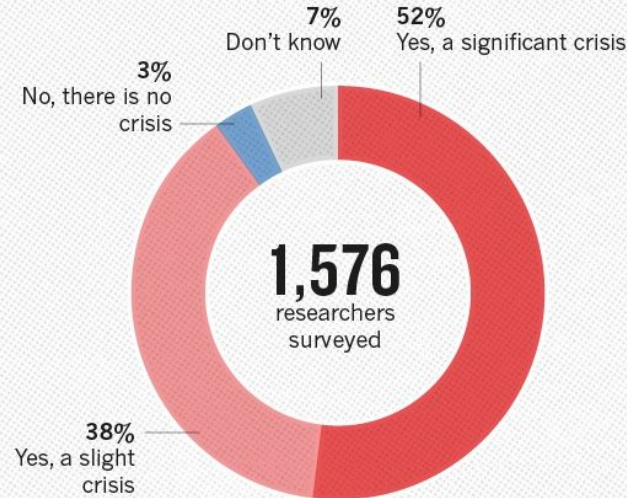
Washington's lawyer surplus
How to do a nuclear deal with Iran
Investment tips from Nobel economists
Junk bonds are back
The meaning of Sachin Tendulkar

OCTOBER 19TH 2014 2014 2013 economist.com

HOW
SCIENCE
GOES
WRONG

99
Einsteinium

IS THERE A REPRODUCIBILITY CRISIS?



Why Most Published Research Findings Are False

John P. A. Ioannidis

[DOI: 10.1371/journal.pmed.0020124](https://doi.org/10.1371/journal.pmed.0020124)

A computational reproducibility study featuring longitudinal data analyses

Heidi Seibold^{*,1,2,3,4}, Severin Czerny¹, Siona Decke¹, R Steffen Fohr¹, Nico Hahn¹, Rabea Hartmann¹, Christo Dario Lepke¹, Verena Loidl¹, Maximilian Mandl¹, Sara Alexander Piehler¹, Elio Rojas¹, Stefanie Schmid¹, Hanns-Lennart Schneider¹, Xiao-Yin To¹, Viet Tran¹, Antje Wagner¹, Maria Waize¹, Hannah Wecker¹, Rui Yang¹,

1 LMU Munich, Germany

2 University of Bielefeld, Germany

3 Helmholtz Zentrum München, Germany

4 LMU Open Science Center

Is this what we want?

**How to
become a
science
champion?**



		Data	
		Same	Different
Analysis	Same	Reproducible	Replicable
	Different	Robust	Generalisable

From [The Turing Way](#)

Practical steps towards computational reproducibility



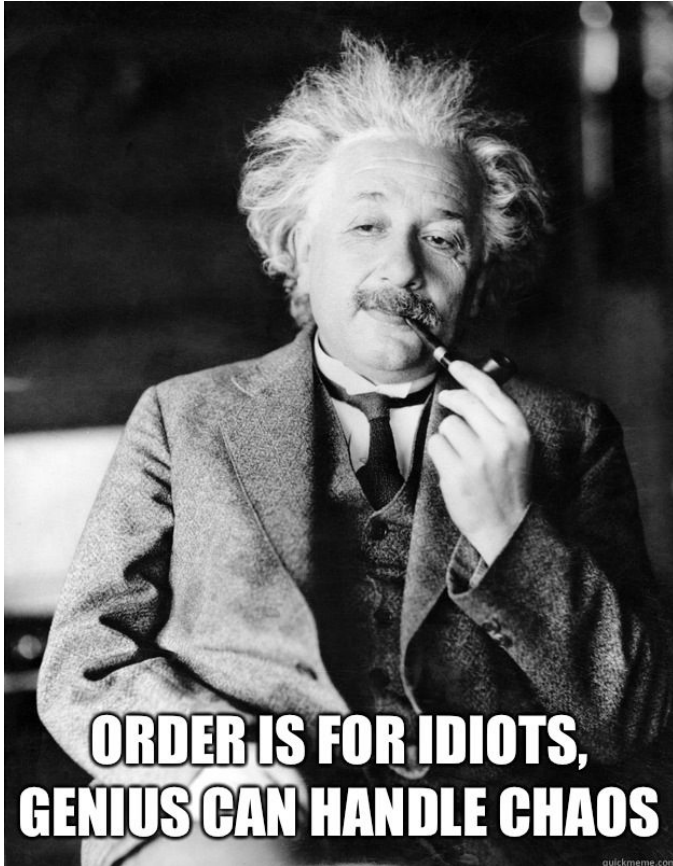
Practical steps towards reproducible research

- Get organized!
- Use Open Source Software
- Use Version Control
- Make your work available online

Get organized!



Good organisation



**ORDER IS FOR IDIOTS,
GENIUS CAN HANDLE CHAOS**

<http://www.quickmeme.com/meme/3r98zx>



I'm an extremely stable genius.

Let's not pretend: we're not geniuses ;P

Good organisation

... starts simple

- No sending emails with files
- Nice file organisation
- Good naming

Dear colleagues,

attached you find the first public version of the [REDACTED] protocol. Please have a look and do comment. We can also meet to aggregate our reviews.

▶ 📎 1 attachment: StudyProposa[REDACTED]_Validation_V1_250918docx.docx

```
— 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
```

Naming

NO

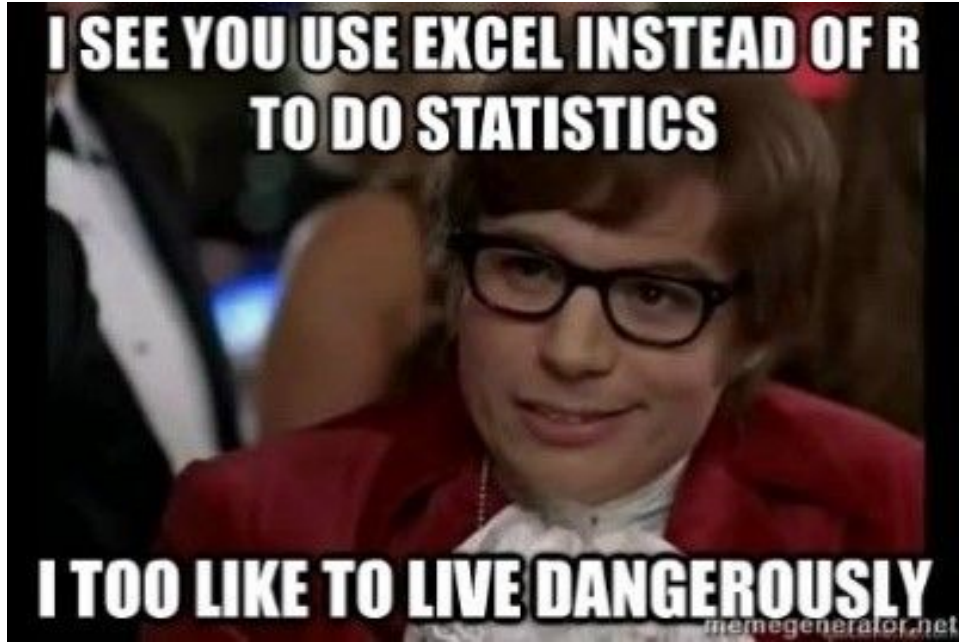
- Myabstract.docx
- Joe's Filenames Use Spaces and Punctuation.xlsx
- figure 1.png
- fig 2.png
- JW7d^(2sl@deletethisandyourcareerisoverWx2*.txt

YES

- 2014-06-08_abstract-for-sla.docx
- Joes-filenames-are-getting-better.xlsx
- Fig01_scatterplot-talk-length-vs-interest.png
- Fig02_histogram-talk-attendance.png
- 1986-01-28_raw-data-from-challenger-o-rings.txt

Use Open Source Software

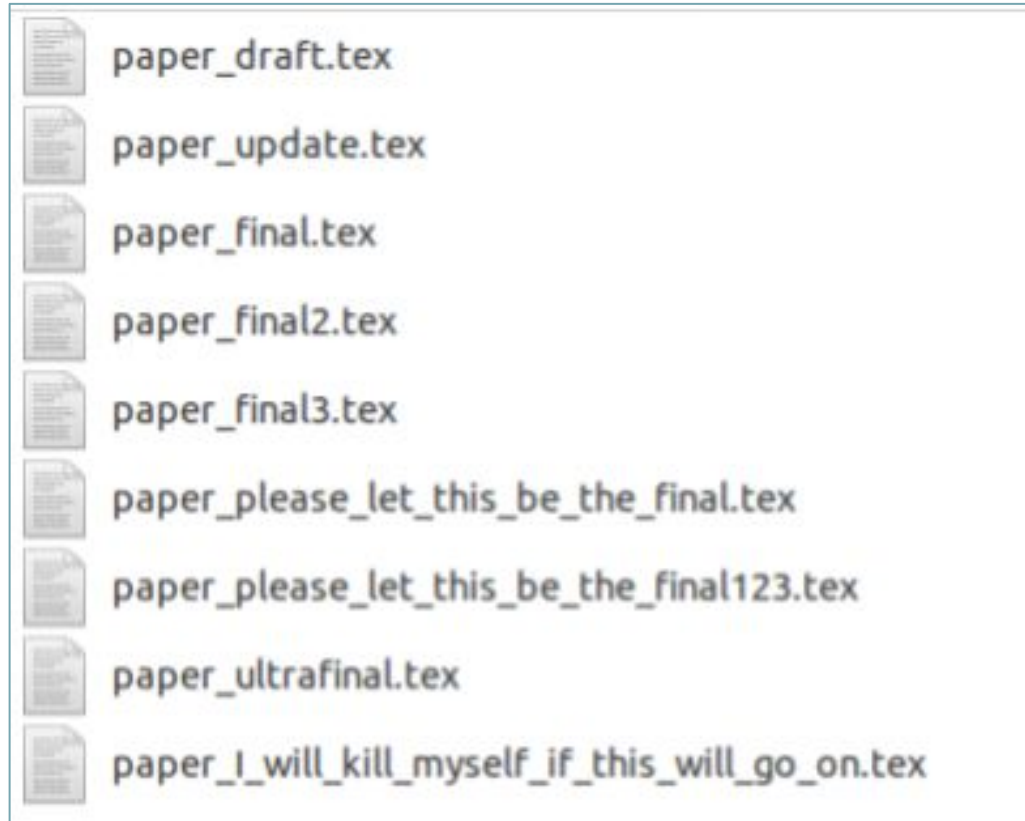




Use Version Control



Version Control in the olden days



Real Version Control

(Including Backup)


Commit `fe6b5538` authored 10 months ago by  HeidiSeibold

Browse files

Options ▾

fix typos

parent `e8a0da8a`  master

 No related merge requests found

Showing 1 **changed file** ▾ with 3 **additions** and 3 **deletions**

Hide whitespace changes

Inline

Side-by-side

▼  2018_swisscore.tex 



View file @ fe6b5538

...	...	@@ -56,8 +56,8 @@ In my research I develop algorithms that detect which patient characteristics
56	56	lead to a positive or negative reaction to a therapy. The algorithms can also
57	57	detect and predict which patients are likely to have side effects from a
58	58	therapy. It is important to acknowledge, that <code>\textbf{personalised medicine does not</code>
59		- necessarily mean, that each patient receives a personalised treatment, but it
60		- can also mean that we find that treatments work similarly across patients and
	59	+ necessarily mean, that each patient receives a personalised treatment, but it
	60	+ can also mean that we find that treatments work similarly across patients and

Prerequisites for Version Control

- **Text**
 - Documents/Papers: LaTeX, Markdown
 - Analyses: R, Python

- **Willingness to learn something new: Git**

**Make your work available
online**



Open/FAIR Data

FAIR DATA PRINCIPLES

AH!



FINDABLE



ACCESIBLE

HOW DO YOU
OPEN A .XZQ FILE?



INTEROPERABLE



REUSABLE

Open Code

Steps

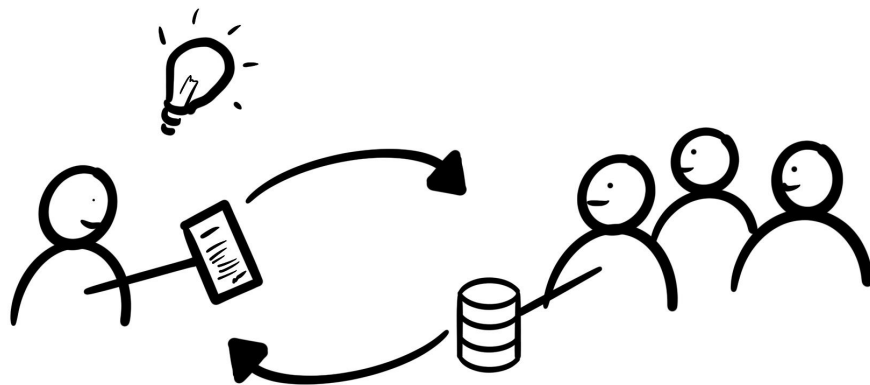
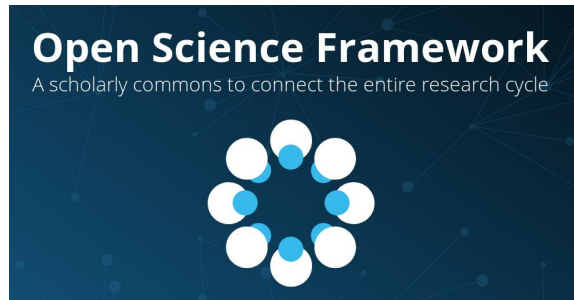
s1 <- "Use scripts (no clicking!)"

s2 <- "Publish code online"

Pro tip

s3 <- "User version control"

Open Materials



GitLab

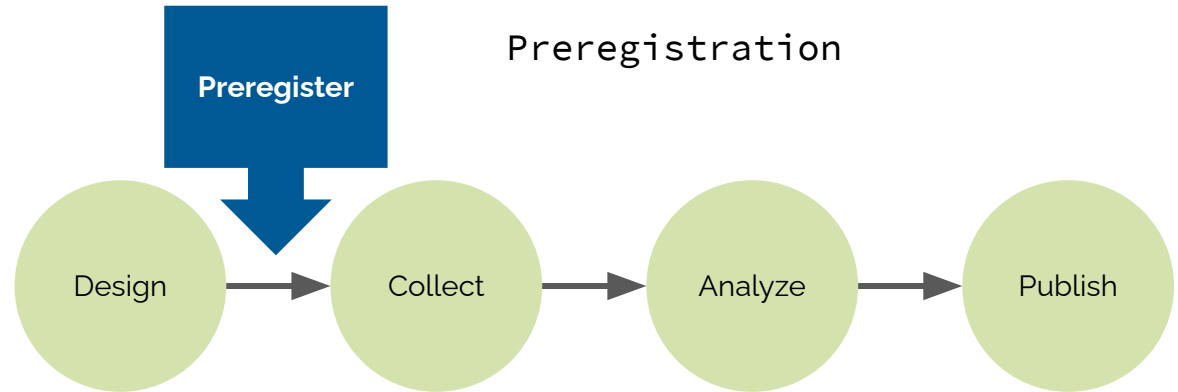
Practical steps towards reproducible research

- Get organized!
- Use Open Source Software
- Use Version Control
- Make your work available online

Other things to consider



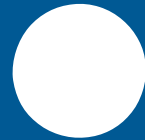
Science communication



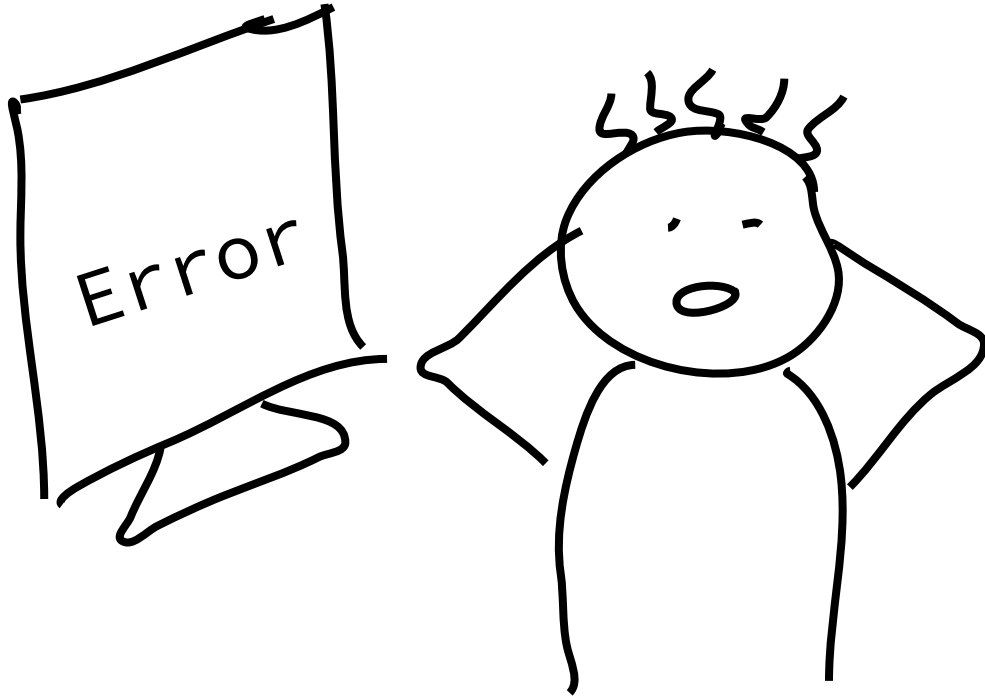
cos.io/prereg



What is stopping you?
→ **Answer in cryptpad**

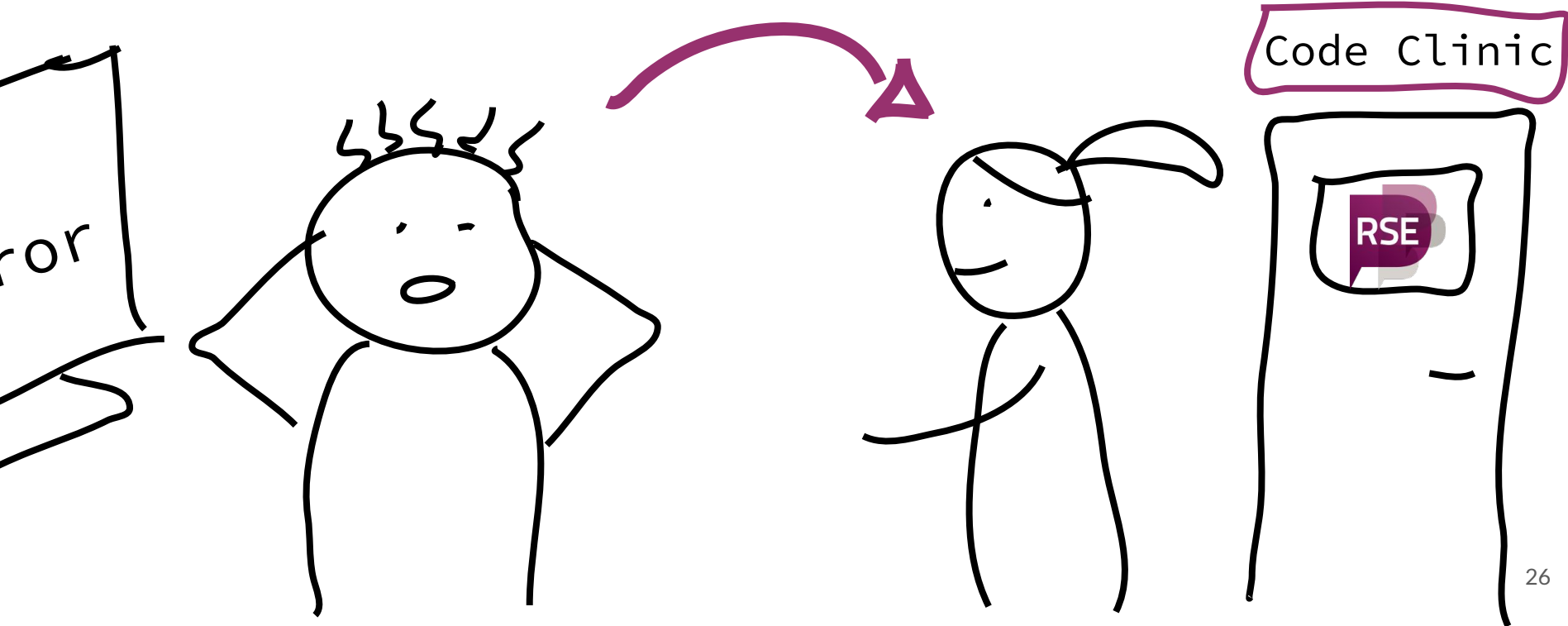


Overwhelmed?



Reproducible
Research
requires
Software Skills

Research Software Engineers can help

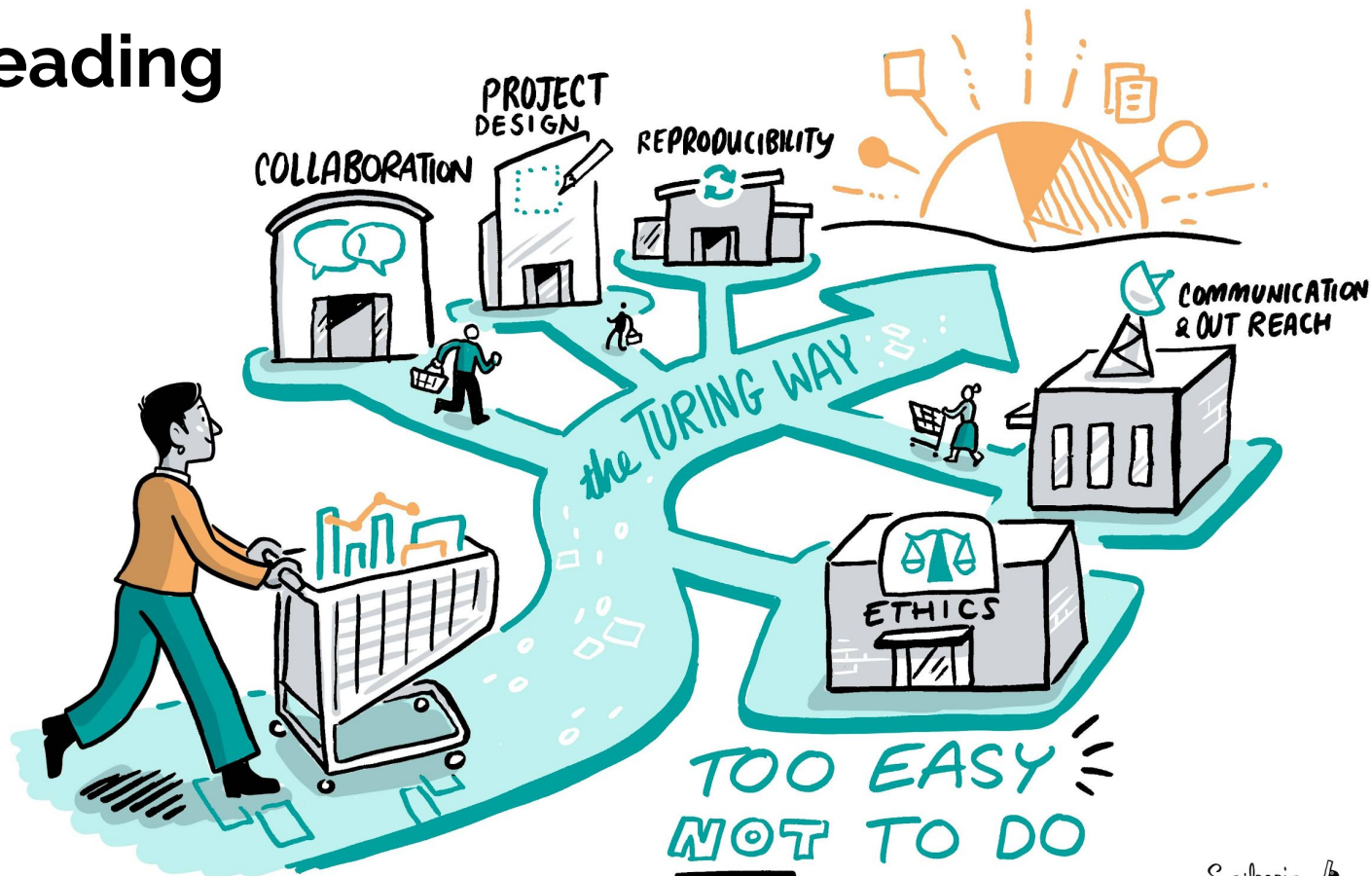


Help us help you



**BETTER
SOFTWARE
BETTER
RESEARCH**

Further reading



the-turing-way.netlify.app

Scriberia 

Can I ask a favor?

→ Help me become **AI newcomer of the year** (vote before March 7th)



kicamp.org/en/ai-newcomers

Voting

Your mail address

I hereby confirm that I have read and agree with the data privacy information of the Gesellschaft für Informatik e.V.

[Vote](#)

<https://anchor.fm/opensciencestories>

Open Science Stories

Let's discuss!

@HeidiBaya

HELMHOLTZ **AI** | ARTIFICIAL INTELLIGENCE
COOPERATION UNIT

OPEN DISCUSSION

KEEP IN TOUCH AND STAY INFORMED

- Mail – open-science@helmholtz.de
- Website – www.os.helmholtz.de
- Make sure to sign up for our [Open Science Newsletter](#)
 - Now both in German & English!
- Twitter – [@helmholtz_os](https://twitter.com/helmholtz_os)

KEEP IN TOUCH AND STAY INFO

- Mail – open-science@helmholtz.de
- Website – www.os.helmholtz.de
- Make sure to sign up for our [Open Science News](#)
 - Now both in German & English!
- Twitter – [@helmholtz_os](https://twitter.com/helmholtz_os)

**RDA DE
Conference 2021**

February 22 – 26

Program
Registration

MANY THANKS FOR YOUR PARTICIPATION!

heidi.seibold@helmholtz-muenchen.de

 <https://orcid.org/0000-0002-8960-9642>

All texts in this presentation, except citations, are licenced under
Attribution 4.0 International (CC BY 4.0):

<https://creativecommons.org/licenses/by/4.0/deed.de>