

# Reproducible Research

block course @ifgi

Jakub Krukar  
<http://krukar.staff.ifgi.de>  
krukar@uni-muenster.de

# Our plan

## Day 1

- (*before lunch*) - we learn the basics

- (*after lunch*) - we reproduce

## Day 2

- (*before lunch*) - own projects

- (*if any time left*) - reviewing each other's work

Note: I'm not your guru

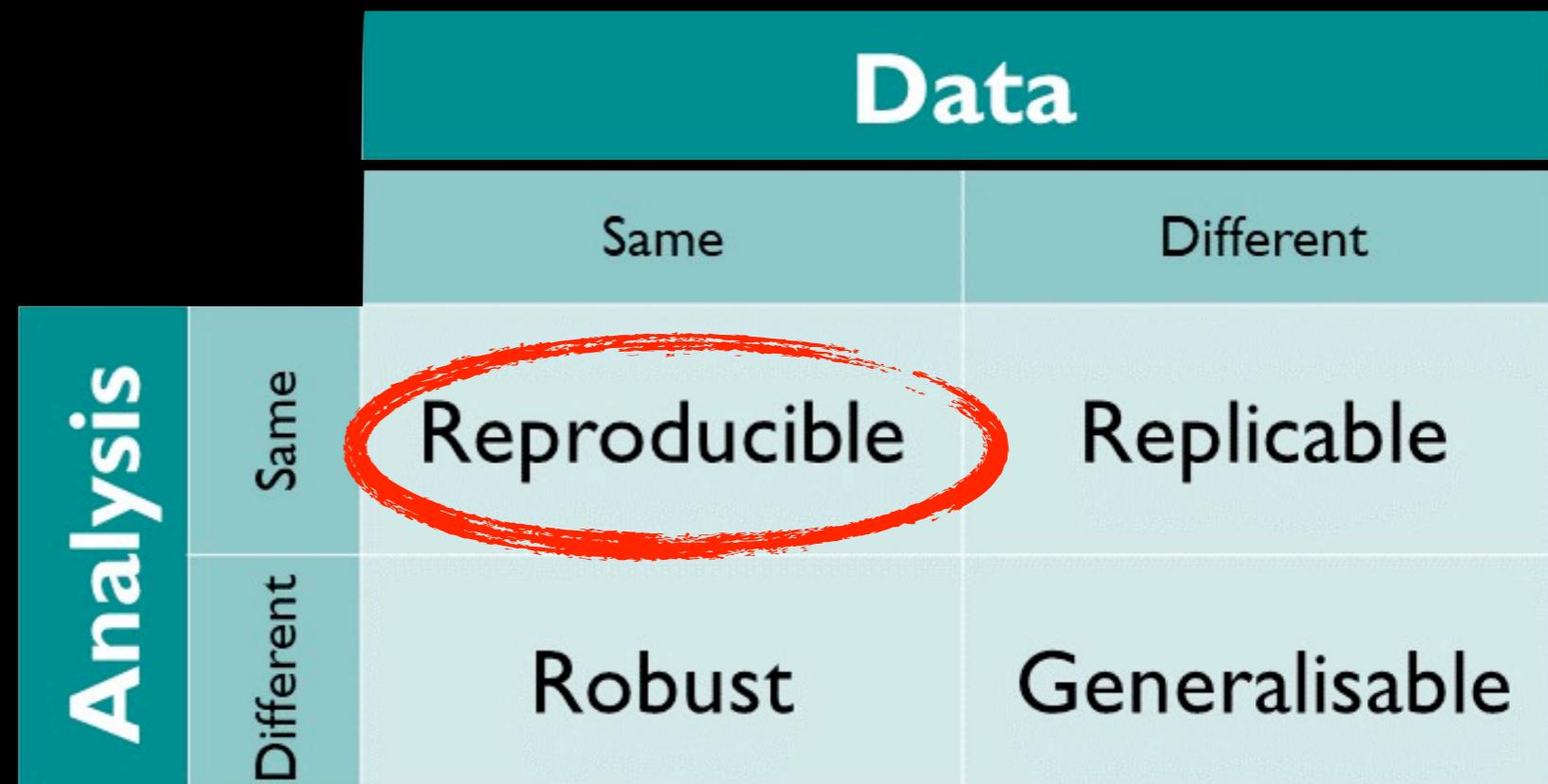


# Reproducibility: the basics

- **What** is it?
- **Why** do it?
- **How** to do it?

# Reproducibility: the basics

- **What** is it?
- **Why** do it?
- **How** to do it?



Source: *The Turing Way*. DOI: 10.5281/zenodo.3233853

# Reproducibility: the basics

- **What** is it?
- **Why** do it?
- **How** to do it?

## 4.4 Software and Data Availability

Data used in this paper can be accessed with the public Flickr API.<sup>8</sup> The query used to access the data, code and interactive data visualization (Fig. 2) are available on GitHub.<sup>9</sup>

# Reproducibility: the basics

## ***Research data/code supporting this publication ...***

*... is available in [name of the repository(-ies)] and is accessible via the following DOI [DOI link(s)]*

*... was accessed on [date of dataset access/download] with the following [query parameters, if applicable] under the license [dataset license].*

## ***The computational workflow supporting this publication ...***

*... is executed via [choose, e.g., a single command/file, a workflow management software, a set of numbered scripts] published under license [the license] at [DOI of repository].*

*... is published in a [language] module/package at [link of software project]. The used version is archived at [DOI of repository].*

*... is provided as a [container/VM] published at [DOI of repository] with instructions included in the file README.md in the repository.*

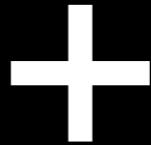
# Reproducibility: the basics

- What is it?
- Why do it?
- How to do it?



# Reproducibility: the basics

- What is it?
- Why do it?
- How to do it?



OSFHOME ▾ My Quick Files My Projects Search Support Donate Mathias Harrer ▾

Meta-Analysis Project Files Wiki Analytics Registrations Contributors Add-ons Settings

Notice: The site will undergo maintenance between Oct 17, 2019 3:00 AM and Oct 17, 2019 4:00 AM (+0200 UTC). Thank you for your patience.

Private Make Public ⚡ ...

## Meta-Analysis Project

Contributors: Mathias Harrer

Date created: 2019-10-16 05:55 PM | Last Updated: 2019-10-16 05:55 PM

Category: Project

Description: Add a brief description to your project

License: Add a license

Wiki

Add important information, links, or images here to describe your project.

Files

Click on a storage provider or drag and drop to upload

Name ▾ ▾ Modified ▾ ▾

- Meta-Analysis Project
- OSF Storage (United States)

Citation

Components Add Component Link Projects

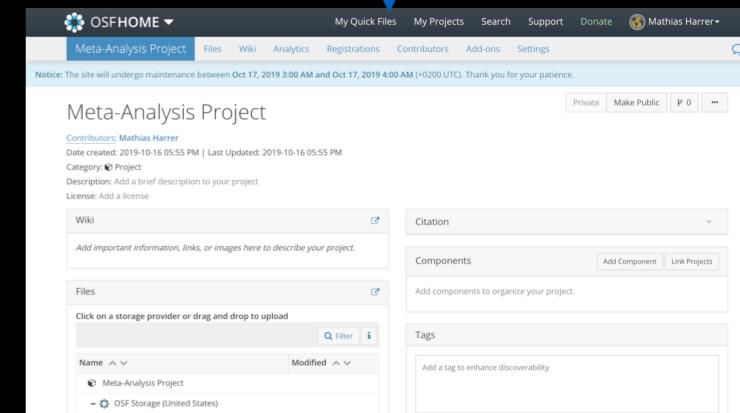
Add components to organize your project.

Tags

Add a tag to enhance discoverability

# Reproducibility: the basics

- **What** is it?
- **Why** do it?
- **How** to do it?



# Reproducibility: the basics

- **What is it?**
- **Why do it?**
- **How to do it?**



joint note-taking: <http://go.wwu.de/nh5zv>

# Reproducibility: the basics

- **What is it?**
- **Why do it?**
- **How to do it?**

Open science challenges, benefits and tips in early career and beyond

Christopher Allen<sup>1</sup>\*, David M. A. Mehler<sup>1,2</sup>\*

1 Cardiff University Brain Research Imaging Centre (CUBRIC), Wales, United Kingdom, 2 Department of

## Reproducibility and Replicability in a Fast-Paced Methodological World



Sacha Epskamp

Department of Psychology, University of Amsterdam

## Reproducibility Starts from You Today

Yasemin Turkyilmaz-van der Velden,<sup>1,\*</sup> Nicolas Dintzner,<sup>1</sup> and Marta Teperek<sup>1</sup>

<sup>1</sup>Delft University of Technology, Mekelweg 5, 2628 CD Delft, the Netherlands

\*Correspondence: [y.turkyilmaz-vandervelden@tudelft.nl](mailto:y.turkyilmaz-vandervelden@tudelft.nl)

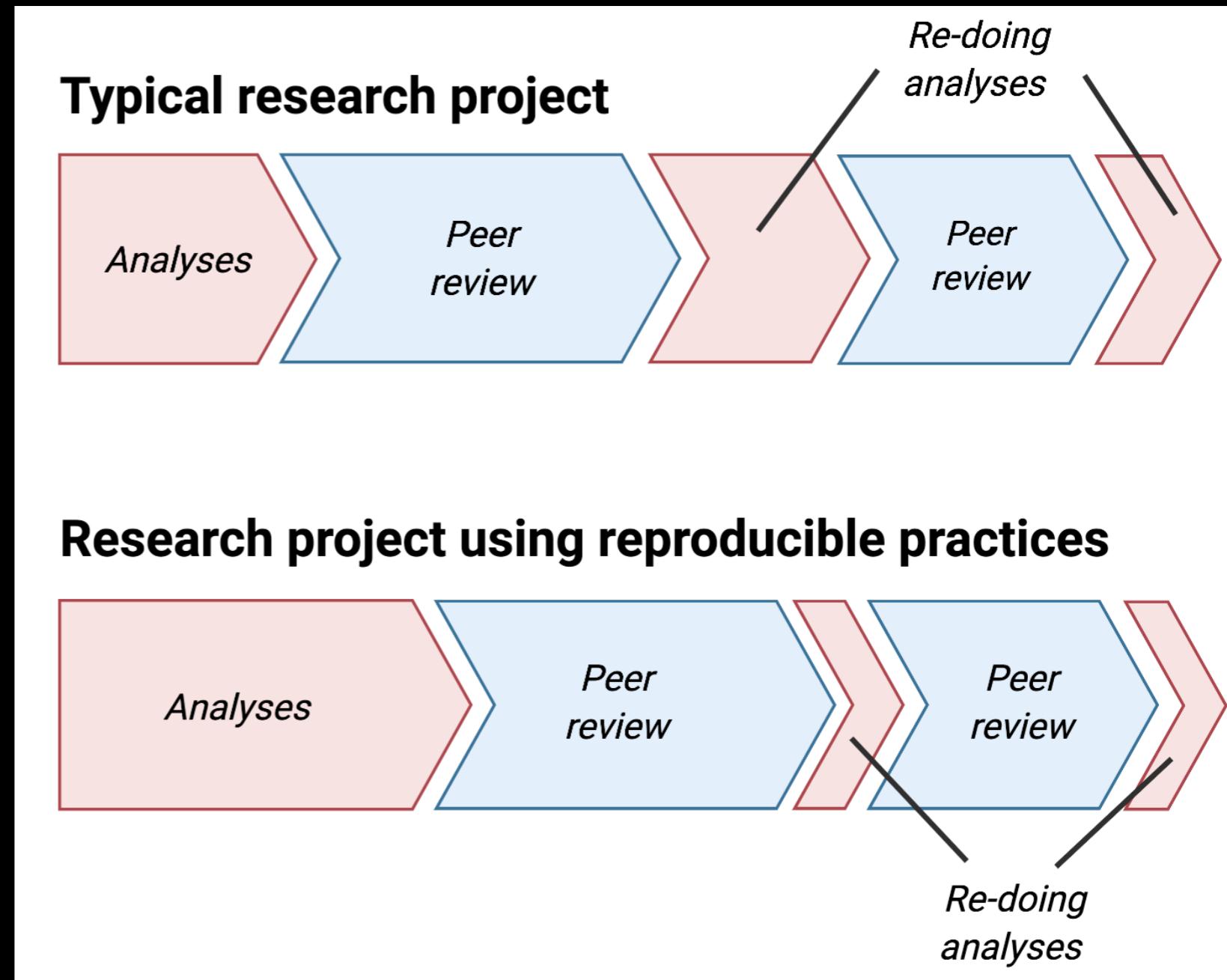
<https://doi.org/10.1016/j.patter.2020.100099>

# Reproducibility: the basics

- **What** is it?
  - easy entry (just upload your data)
  - more citations
  - longevity / long term impact
- **Why** do it?
  - (...of your major work output)
- **How** to do it?
  - takes pressure of your decisions:
    - researcher degrees of freedom
    - future methods will be better

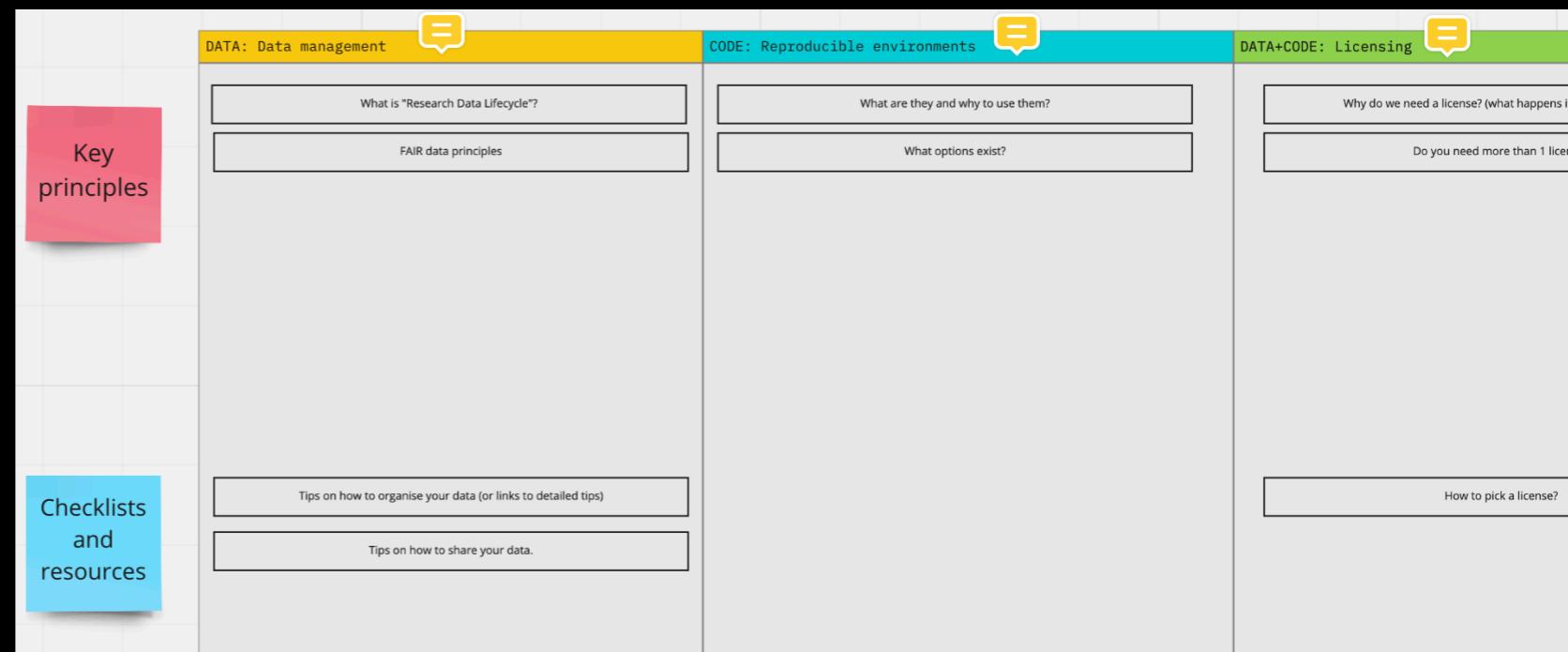
# Reproducibility: the basics

- **What** is it?
- **Why** do it?
- **How** to do it?



# Reproducibility: the basics

- **What is it?**
- **Why do it?**
- **How to do it?**



The handbook:  
Miro board with notes:

<http://the-turing-way.netlify.app>  
<http://go.wwu.de/4mb5u>

# Our plan

## Day 1

- (*before lunch*) - we learn the basics
- (*after lunch*) - we reproduce

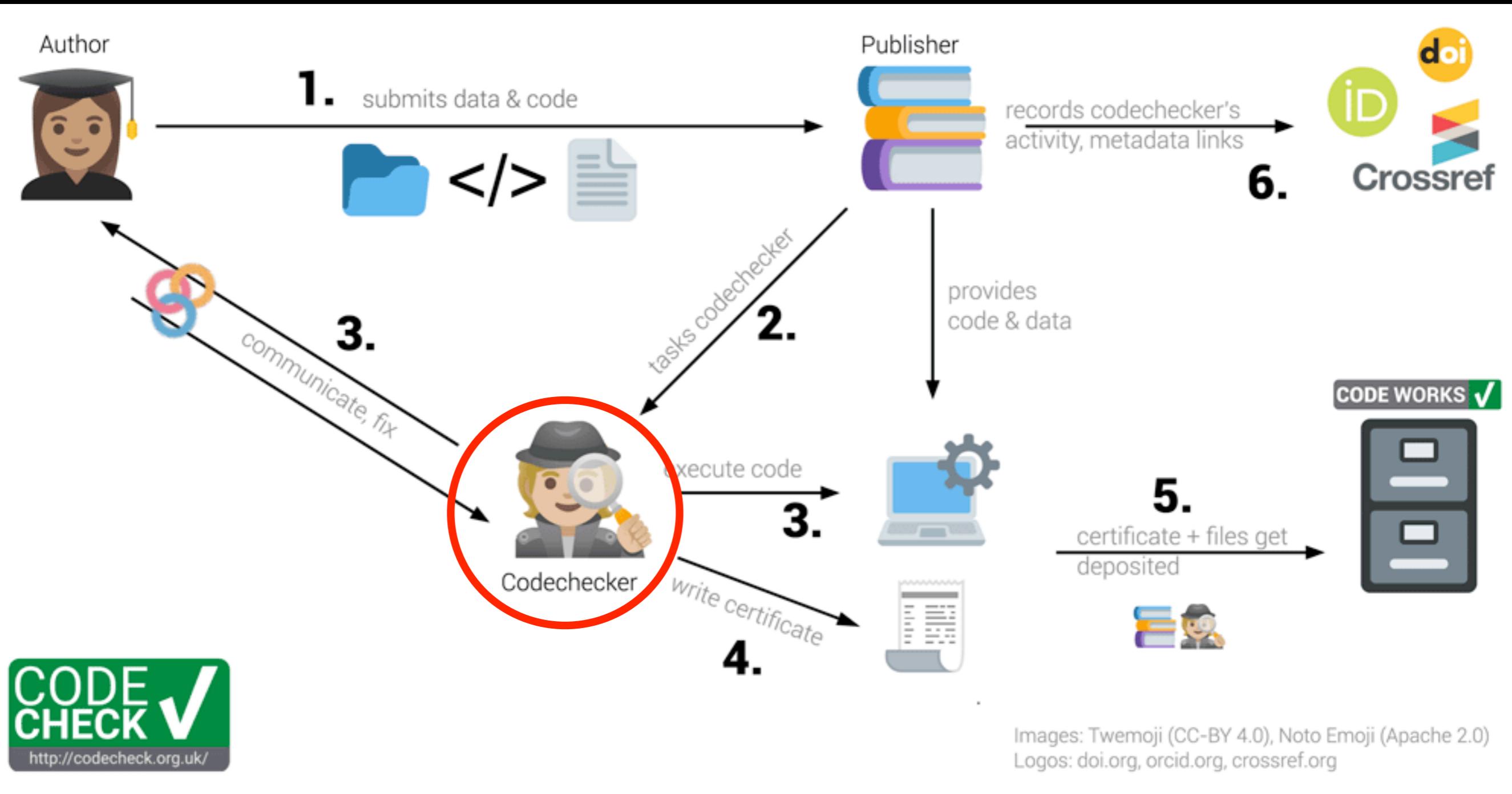
## Day 2

- (*before lunch*) - own projects
- (*if any time left*) - reviewing each other's work

# Become a reproducibility reviewer



# Become a reproducibility reviewer



# Your task

1. Pick a paper:

<https://agile-giss.copernicus.org/articles/>

2. Reproduce results (tables? figures?).

3. Prepare two slides:

<http://go.wwu.de/xhtf9>

4. Repeat.

5. Present at 15:00.

# Your task

1. Pick a paper:

<https://agile-giss.copernicus.org/articles/>

2. Reproduce results (tables? figures?).

3. Prepare two slides:

<http://go.wwu.de/xhtf9>

4. Repeat.

5. Present at 15:00.

Key tips:

- don't fix anything for more than few minutes
- be fast - if it doesn't work it's not your fault
- for R, you can try the checkpoint package:  
`checkpoint(  
 "2014-09-17",  
 r_version="3.6.2",  
 project_dir="defaults to current working dir")`

# For tomorrow

- Rmd / Jupyter notebook data analysis project

# Our plan

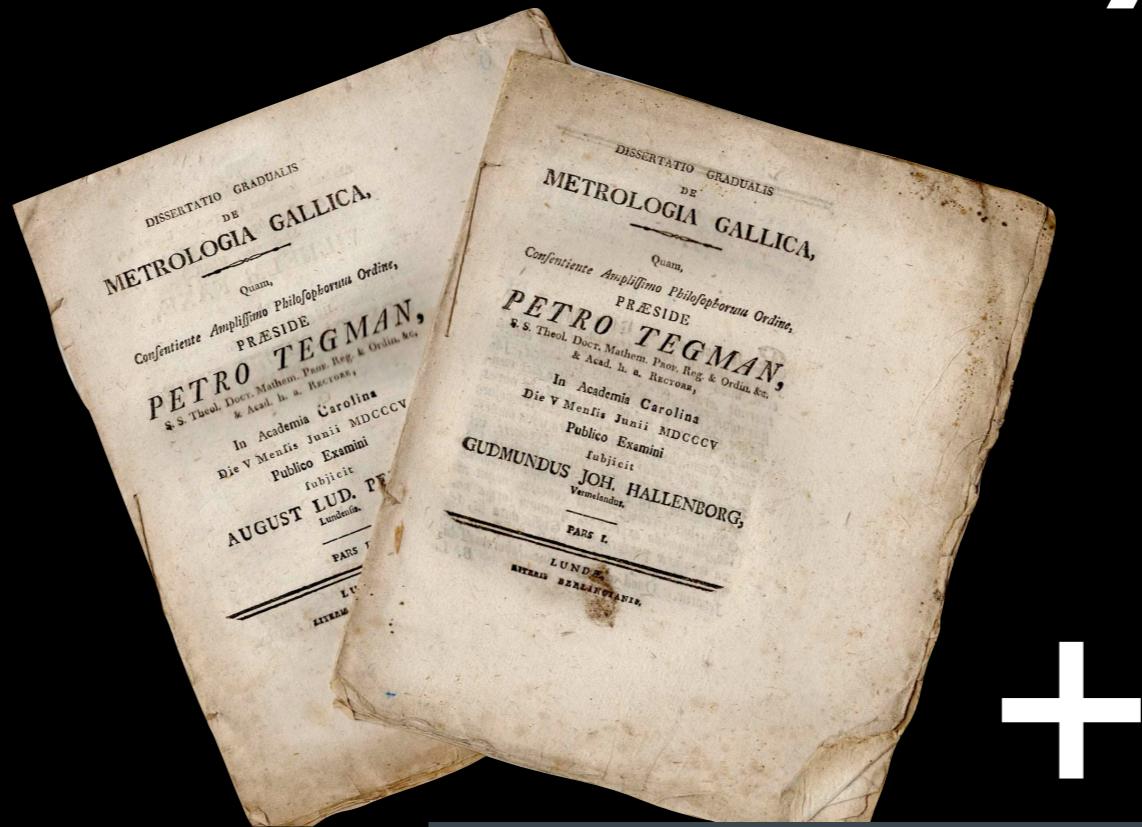
## Day 1

- (*before lunch*) - we learn the basics
- (*after lunch*) - we reproduce

## Day 2

- (*before lunch*) - own projects
- (*if any time left*) - reviewing each other's work

# Our aim today



OSFHOME ▾ My Quick Files My Projects Search Support Donate Mathias Harrer ▾

Meta-Analysis Project Files Wiki Analytics Registrations Contributors Add-ons Settings

Notice: The site will undergo maintenance between Oct 17, 2019 3:00 AM and Oct 17, 2019 4:00 AM (+0200 UTC). Thank you for your patience.

Private Make Public ⚡ 0 ...

## Meta-Analysis Project

Contributors: Mathias Harrer

Date created: 2019-10-16 05:55 PM | Last Updated: 2019-10-16 05:55 PM

Category: Project

Description: Add a brief description to your project

License: Add a license

Wiki

Add important information, links, or images here to describe your project.

Files

Click on a storage provider or drag and drop to upload

Name ▲▼ Modified ▲▼

- Meta-Analysis Project
- OSF Storage (United States)

Citation

Components Add Component Link Projects

Add components to organize your project.

Tags

Add a tag to enhance discoverability

# Our aim today



OSFHOME ▾ My Quick Files My Projects Search Support Donate Mathias Harrer ▾

Meta-Analysis Project Files Wiki Analytics Registrations Contributors Add-ons Settings

Notice: The site will undergo maintenance between Oct 17, 2019 3:00 AM and Oct 17, 2019 4:00 AM (+0200 UTC). Thank you for your patience.

Private Make Public ⚡ ...

## Meta-Analysis Project

Contributors: Mathias Harrer Date created: 2019-10-16 05:55 PM | Last Updated: 2019-10-16 05:55 PM Category: Project Description: Add a brief description to your project License: Add a license

Wiki Add important information, links, or images here to describe your project.

Files Click on a storage provider or drag and drop to upload Filter i

Name	Modified
Meta-Analysis Project	
OSF Storage (United States)	

Citation

Components Add Component Link Projects

Add components to organize your project.

Tags

Add a tag to enhance discoverability

research stuff

documents and data (questionnaires,  
config files for hardware, VR models...)



# Our aim today

MyBinder.org



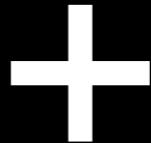
GitHub repo



research stuff



(questionnaires,  
config files for hardware, VR models...)



OSFHOME ▾

My Quick Files My Projects Search Support Donate Mathias Harrer

Meta-Analysis Project Files Wiki Analytics Registrations Contributors Add-ons Settings

Notice: The site will undergo maintenance between Oct 17, 2019 3:00 AM and Oct 17, 2019 4:00 AM (+0200 UTC). Thank you for your patience.

Private Make Public ⚡ ...

## Meta-Analysis Project

Contributors: Mathias Harrer

Date created: 2019-10-16 05:55 PM | Last Updated: 2019-10-16 05:55 PM

Category: Project

Description: Add a brief description to your project

License: Add a license

Wiki

Add important information, links, or images here to describe your project.

Files

Click on a storage provider or drag and drop to upload

Name ▾ ▾ Modified ▾ ▾

- Meta-Analysis Project
- OSF Storage (United States)

Citation

Components Add Component Link Projects

Add components to organize your project.

Tags

Add a tag to enhance discoverability

# Make your project reproducible today

1. Open an [OSF.io](#) account.
2. "Create new project".
3. Upload your paper and supporting material. (+set LICENSE)
4. Open a GitHub account.
5. Initialize a new repository (+with README +with LICENSE).  
Upload your code\*.
6. On OSF: "Add-ons" GitHub -> link to your GitHub account and select the repository from the drop-down menu.
7. Go to [MyBinder.org](#) and follow the tutorial.  
End the link with: [?urlpath=rstudio](#)

# Our plan

## Day 1

- (*before lunch*) - we learn the basics
- (*after lunch*) - we reproduce

## Day 2

- (*before lunch*) - own projects
- (*if any time left*) - reviewing each other's work

# Feedback time

- Go to:  
<http://go.wwu.de/lp4jt>
- Add your name and your project's link
- Pick a project from another person  
(and insert your name next to it)
- Reproduce
- Provide feedback via the form. Be nice :)

# What's next?

## From now on:

- always share data and license
- always add DASA section to your papers



Source: dailysabah.com

# What's next?

INCLUDING DATA IN RESEARCH PAPERS		
	Minimum requirements	Recommended practices
<b>What?</b>	<ul style="list-style-type: none"><li>• All input data and configuration</li><li>• Data description/documentation, including provenance, field or column types, etc.</li><li>• If data is retrieved from an external source, documentation on collection queries and download steps</li></ul>	<ul style="list-style-type: none"><li>• Standardised, discipline-specific metadata<sup>8</sup> and ontologies to describe your data</li><li>• Data download scripts</li></ul>
<b>Where?</b>	<ul style="list-style-type: none"><li>• Publish data in a public repository providing a DOI</li><li>• Cite data (including date and version) in the paper</li></ul>	<ul style="list-style-type: none"><li>• Discipline- or data type-specific repository<sup>9</sup></li><li>• Include recommended citation in dataset description (unless already provided by repository)</li><li>• Create a registration for OSF projects<sup>10</sup> and use the DOI to cite it</li></ul>
<b>How?</b>	<ul style="list-style-type: none"><li>• Use open data formats; export from proprietary format for publication</li><li>• Specify the license</li></ul>	<ul style="list-style-type: none"><li>• Use plain text-based file formats</li></ul>

# What's next?

## From now on:

- always share data and license
- always add DASA section to your papers



Source: dailysabah.com

## Tomorrow:

- always upload your code and README
- use version control (= code already online!)

<https://the-turing-way.netlify.app/reproducible-research/vcs.html> + [HappyGitWithR.com](https://HappyGitWithR.com)

# What's next?

## The day after tomorrow:

- **single all-in workflow:**

Peikert, A., & Brandmaier, A. M. (2021). A Reproducible Data Analysis Workflow With R Markdown, Git, Make, and Docker. *Quantitative and Computational Methods in Behavioral Sciences*, 1, 1–27.  
<https://doi.org/10.5964/qcmb.3763>

- **simplified version with WORCS R package:**

Lissa, C. J. V., Brandmaier, A. M., Brinkman, L., Lamprecht, A.-L., Peikert, A., Struiksma, M. E., & Vreede, B. M. I. (2021). WORCS: A workflow for open reproducible code in science. *Data Science*, 4(1), 29–49.  
<https://doi.org/10.3233/ds-210031>

- **MyBinder.org**

# What else to check out:

- Support at WWU:

<https://confluence.uni-muenster.de/pages/viewpage.action?pageId=47787915>

The screenshot shows a GitHub repository page for 'kubakrukukar/ifgi-GSGI'. The 'Wiki' tab is selected. The main content area has a heading 'Reproducible research' and a note that 'kubakrukukar edited this page 1 minute ago · 8 revisions'. Below this is a section titled 'Useful overview material' with a bulleted list of resources. To the right is a sidebar with a list of 10 pages.

Search or jump to... Pull requests Issues Marketplace Explore

kubakrukukar / ifgi-GSGI Public Unwatch 3 Star 0 Fork 1

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

## Reproducible research

kubakrukukar edited this page 1 minute ago · 8 revisions

### Useful overview material

- Reproducibility knowledge base at WWU has tones of curated resources.
- RIOT youtube channel
- The Turing Way ebook
- Reproducibility policy at the AGILE conference
  - ...and conference proceedings with reproducible-paper badges

### Reproducibility checklists

### Reproducible research GSGI block course in 2021

Pages 10

- Find a Page...
- Home
- Find a summer school
- Getting feedback
- GSGI Requirements
- Managing your personal brand
- Presentation tips
- Productivity techniques

Thanks!