

Swiss Reproducibility Conference

Plan for tomorrow today: why you need a data steward

Lars Schöbitz

Global Health Engineering - ETH Zurich



slides at:

ghe-open.ch

Prof. Elizabeth Tilley

Global Health Engineering - ETH Zurich

June 10, 2024

 hands up 

Who has an ORCID iD?



Who has published a scientific article in
a journal?

Meet a data steward

I have:

- 10+ years work experience (5 in research)
- empathy, compassion, patience, persistance
- an affinity for IT
- teaching experience

I don't have:

- a doctoral degree
- a qualification in computer science
- a qualification in statistics
- a lot of time

Your turn: Think & Note

For 1 minute, think about these two questions and take some notes for later:

1. How should I be rewarded scientifically?
2. Which career paths are there for data stewards?

Research Data Management

Three terms for three stages

term	explanation	file format
unprocessed raw data	data that is not processed and remains in its original form and file type	often XLSX, also CSV and others

Three terms for three stages

term	explanation	file format
unprocessed raw data	data that is not processed and remains in its original form and file type	often XLSX, also CSV and others
processed analysis-ready data	data that is processed to prepare for an analysis and is exported in its new form as a new file	CSV, R data package

Three terms for three stages

term	explanation	file format
unprocessed raw data	data that is not processed and remains in its original form and file type	often XLSX, also CSV and others
processed analysis-ready data	data that is processed to prepare for an analysis and is exported in its new form as a new file	CSV, R data package
final data underlying a publication	data that is the result of an analysis (e.g descriptive statistics or data visualization) and shown in a publication, but then also exported in its new form as a new file	CSV

Data Management Strategy

research
questions

experimental
design

data
collection

data
processing

data
analysis

reproducible
manuscript

published
article

~ 3 to 9 months



research
questions

experimental
design

data
collection

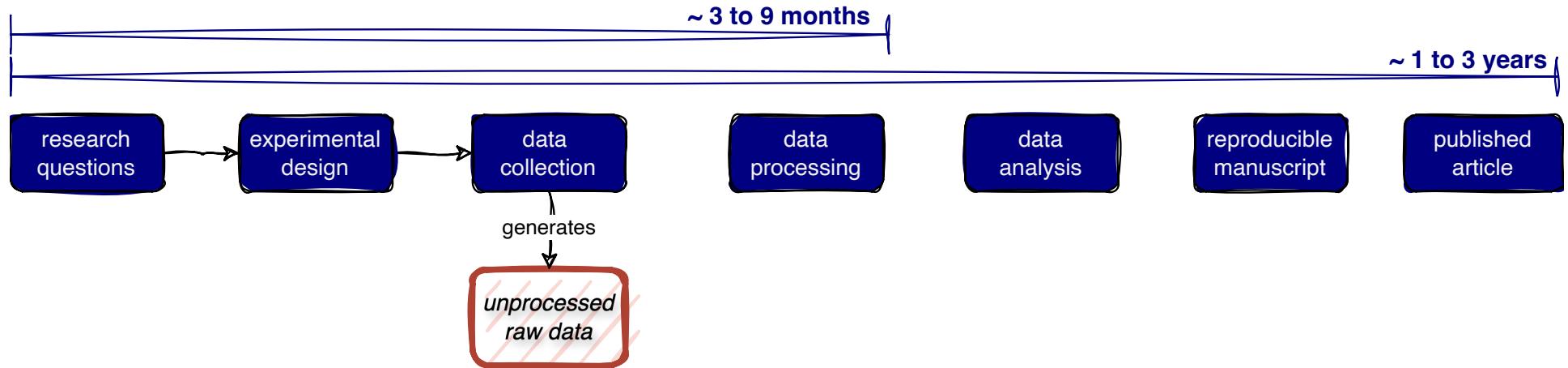
data
processing

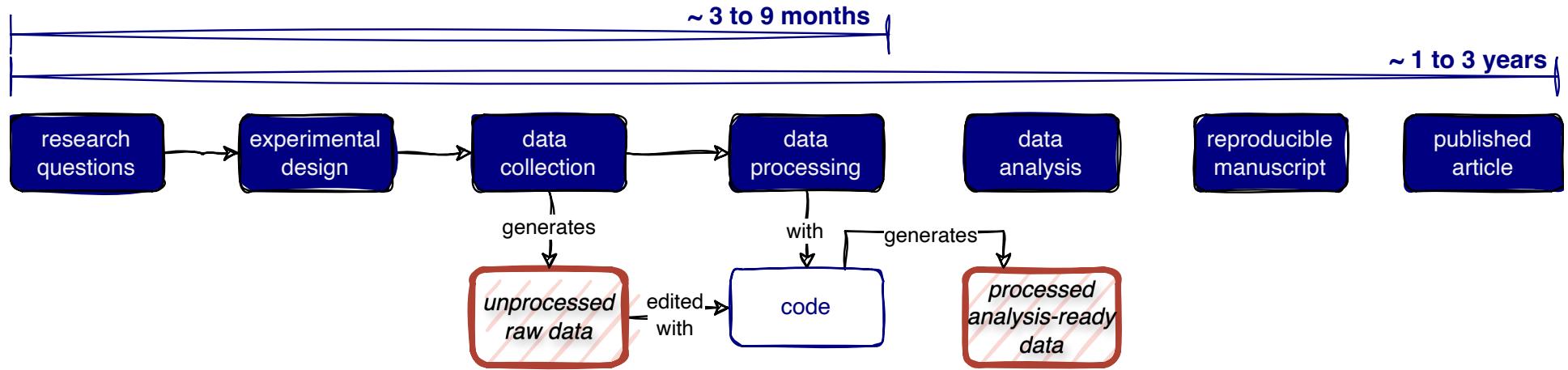
data
analysis

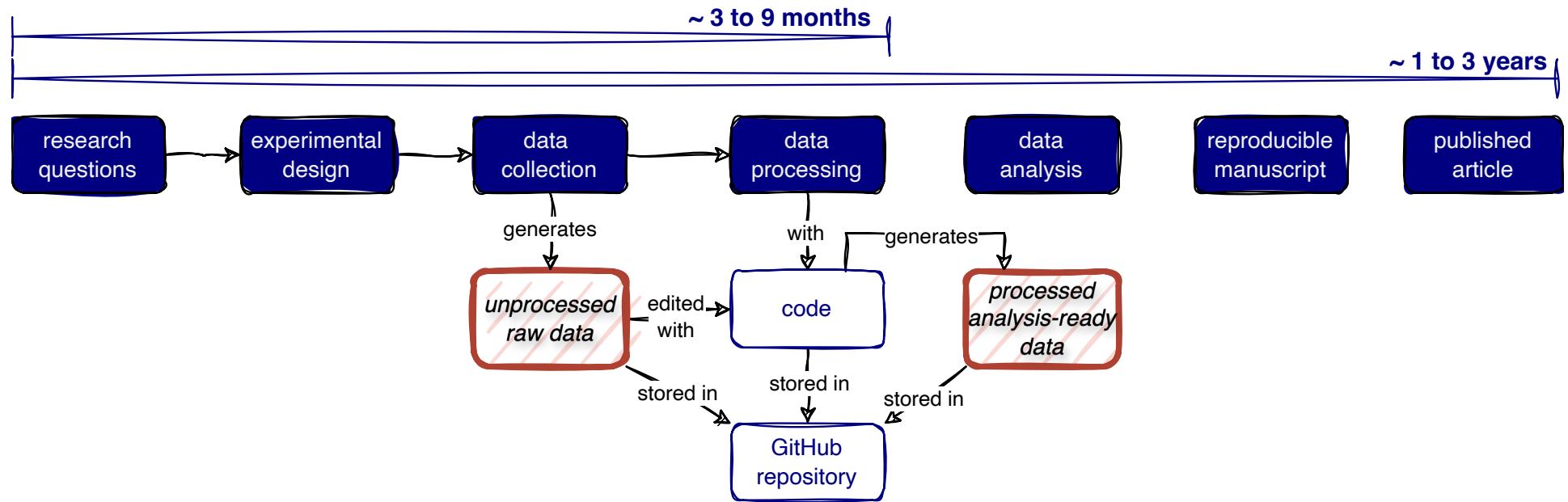
reproducible
manuscript

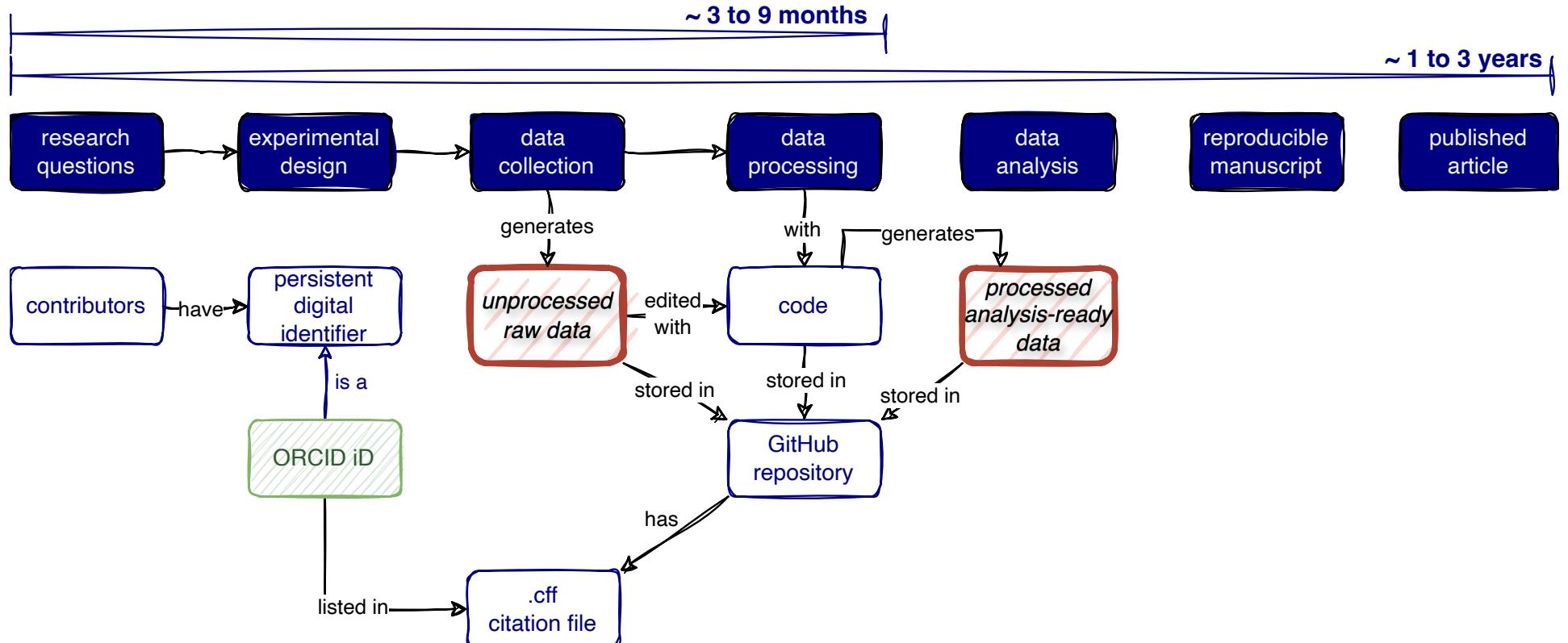
published
article

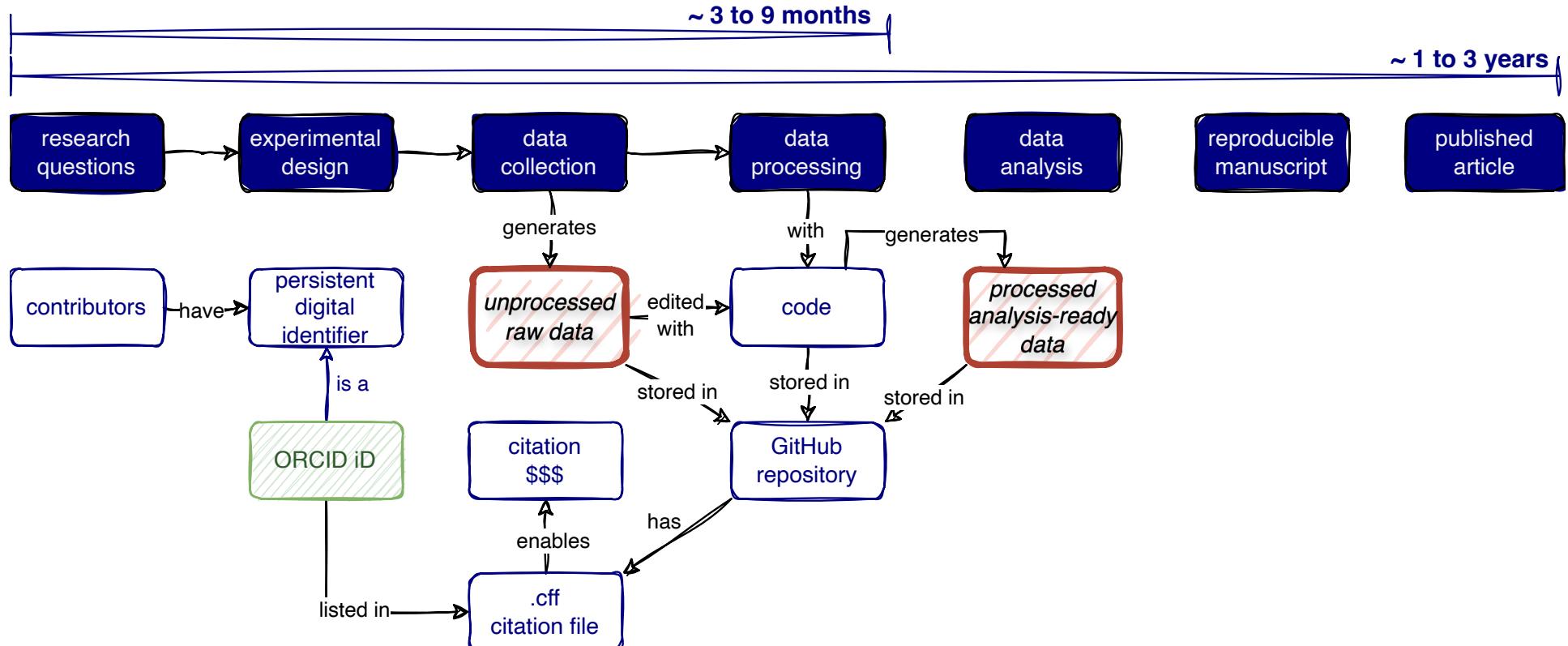
~ 1 to 3 years

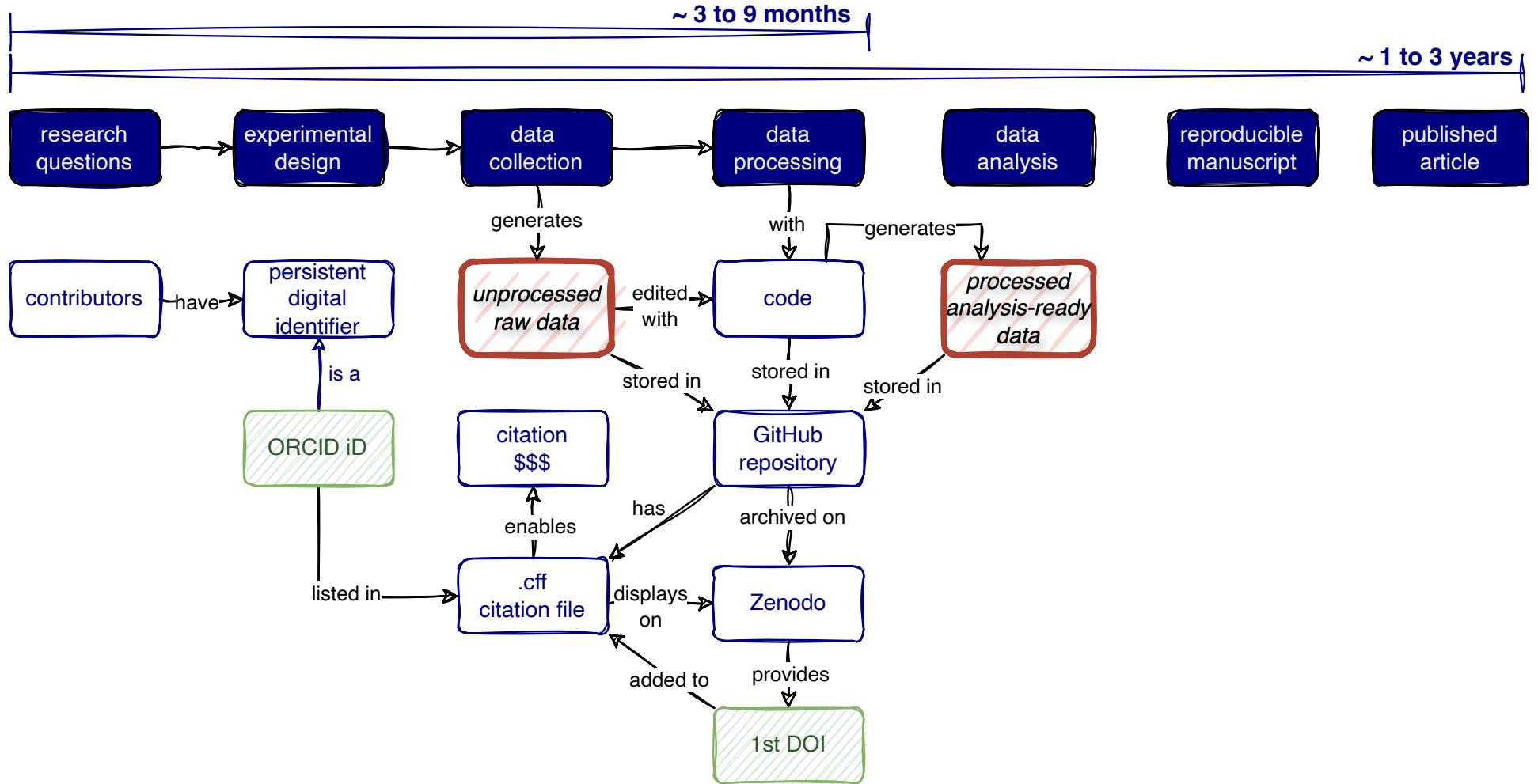


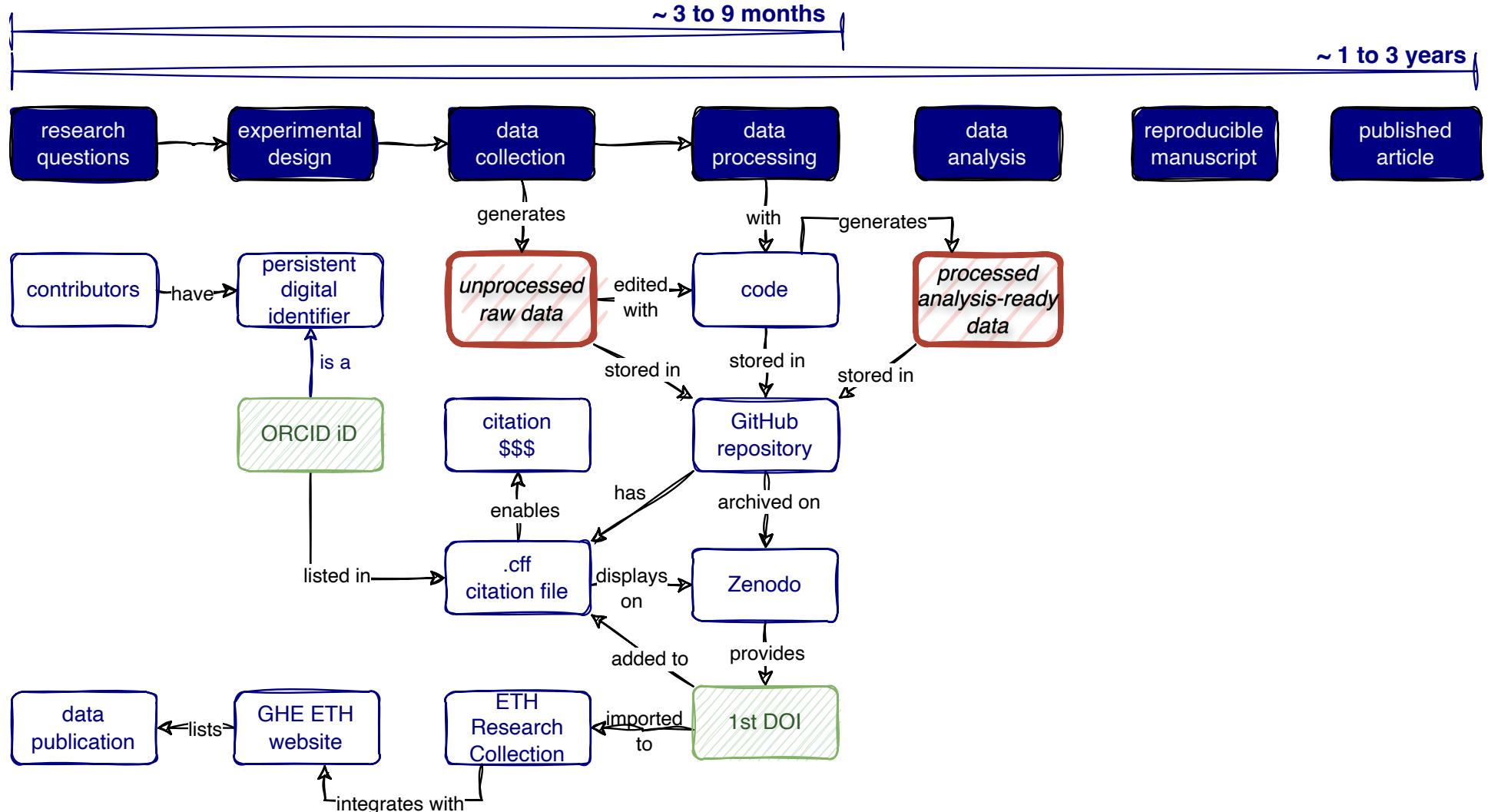


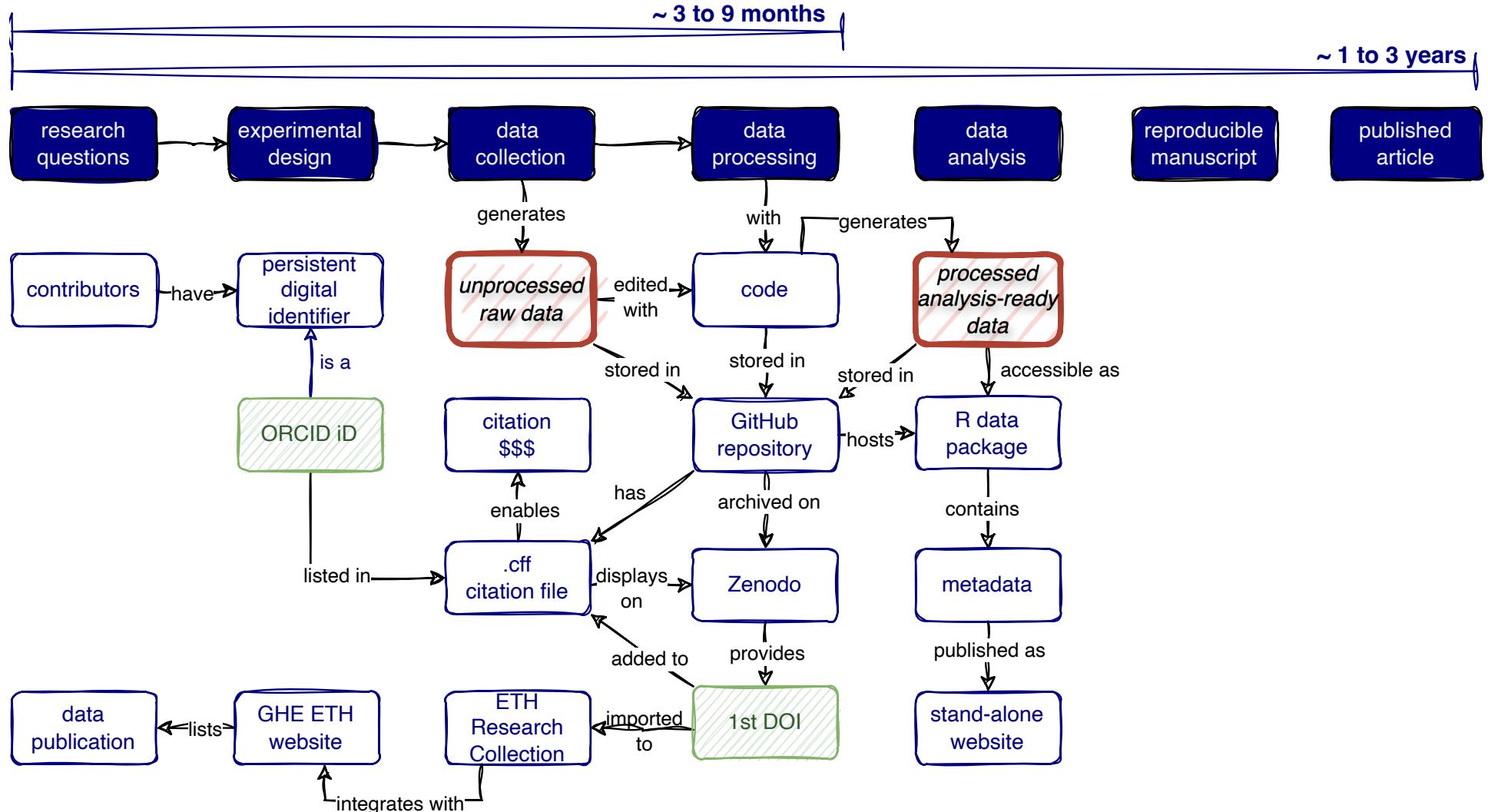


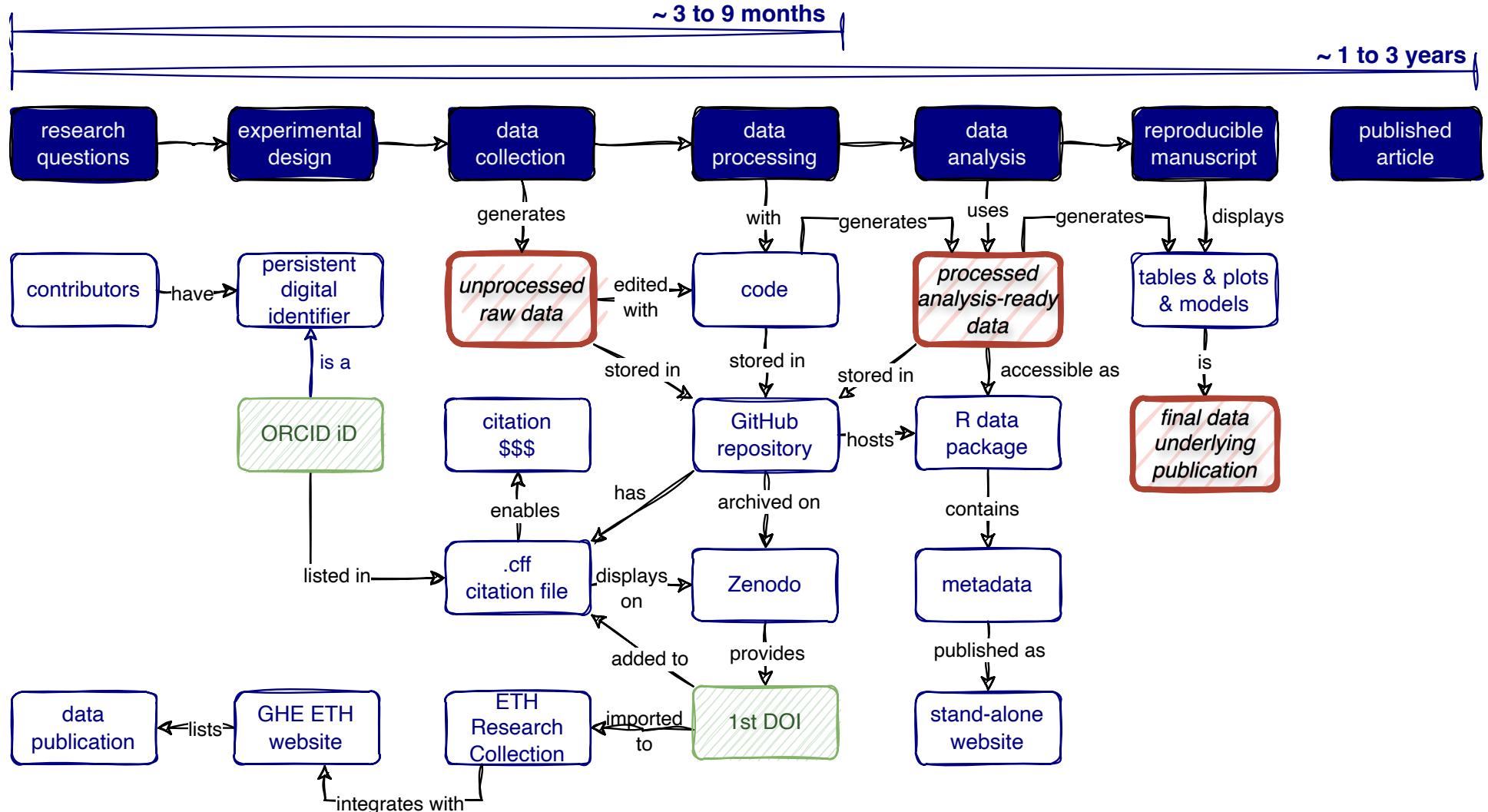


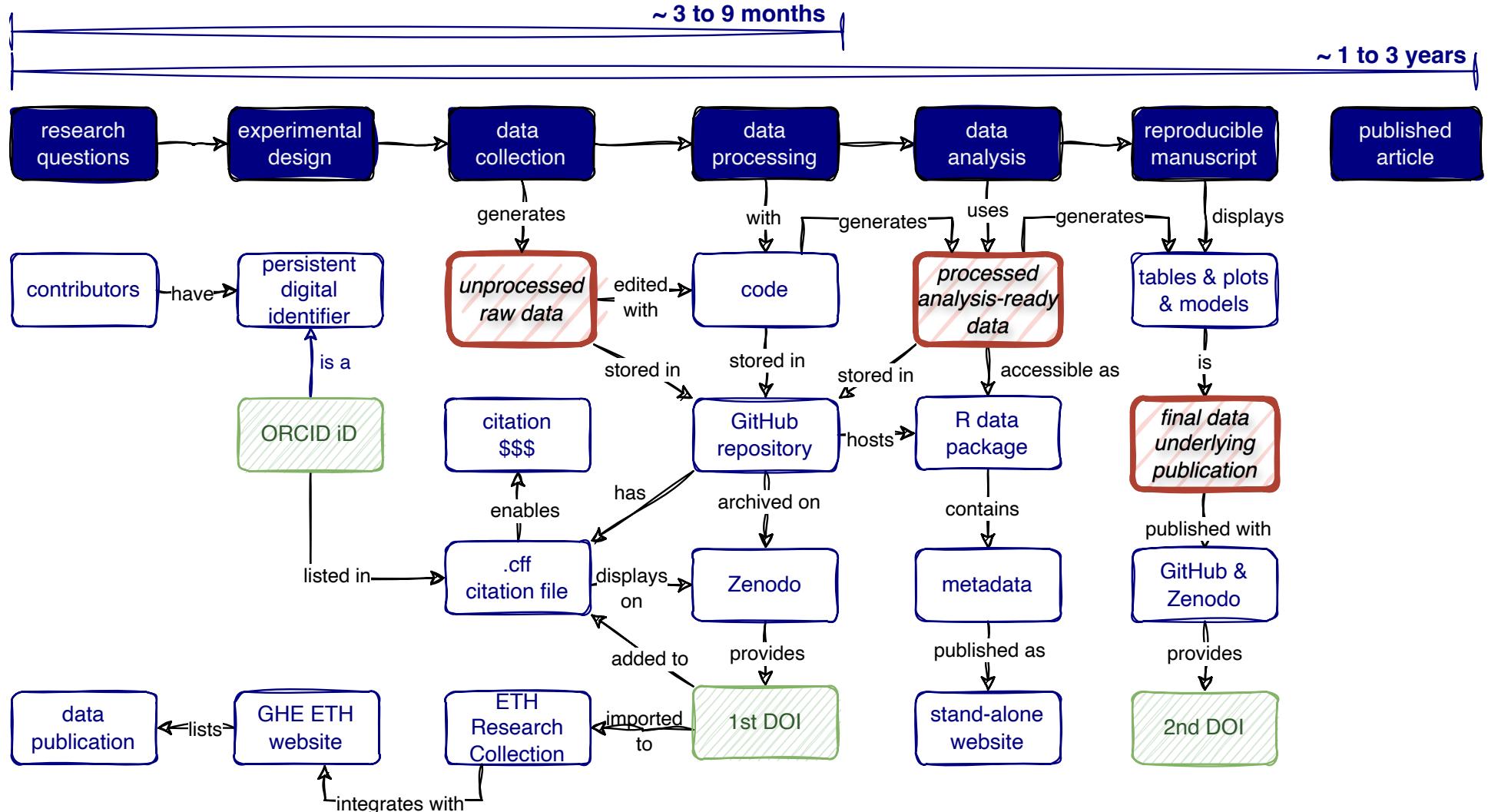


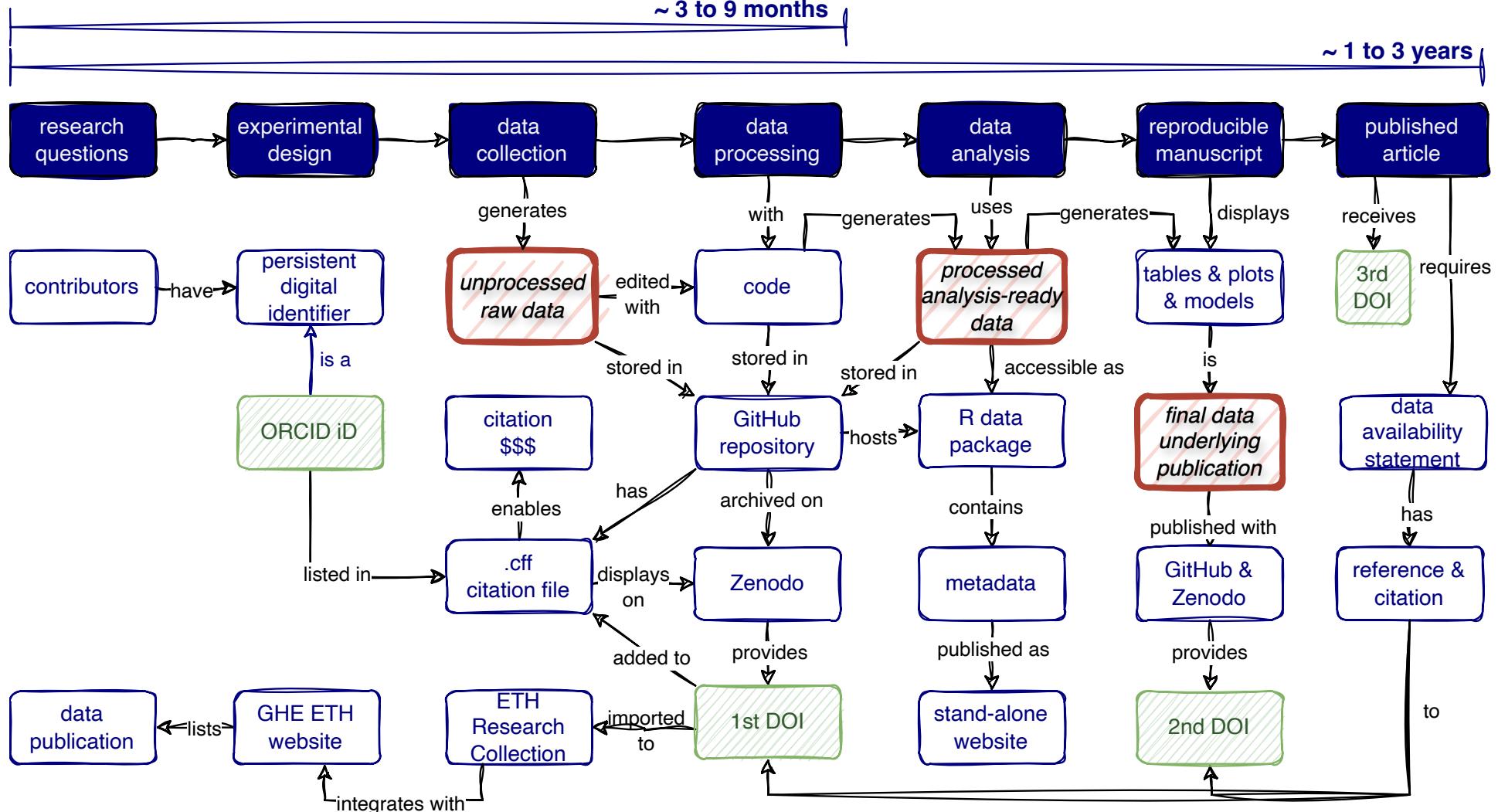


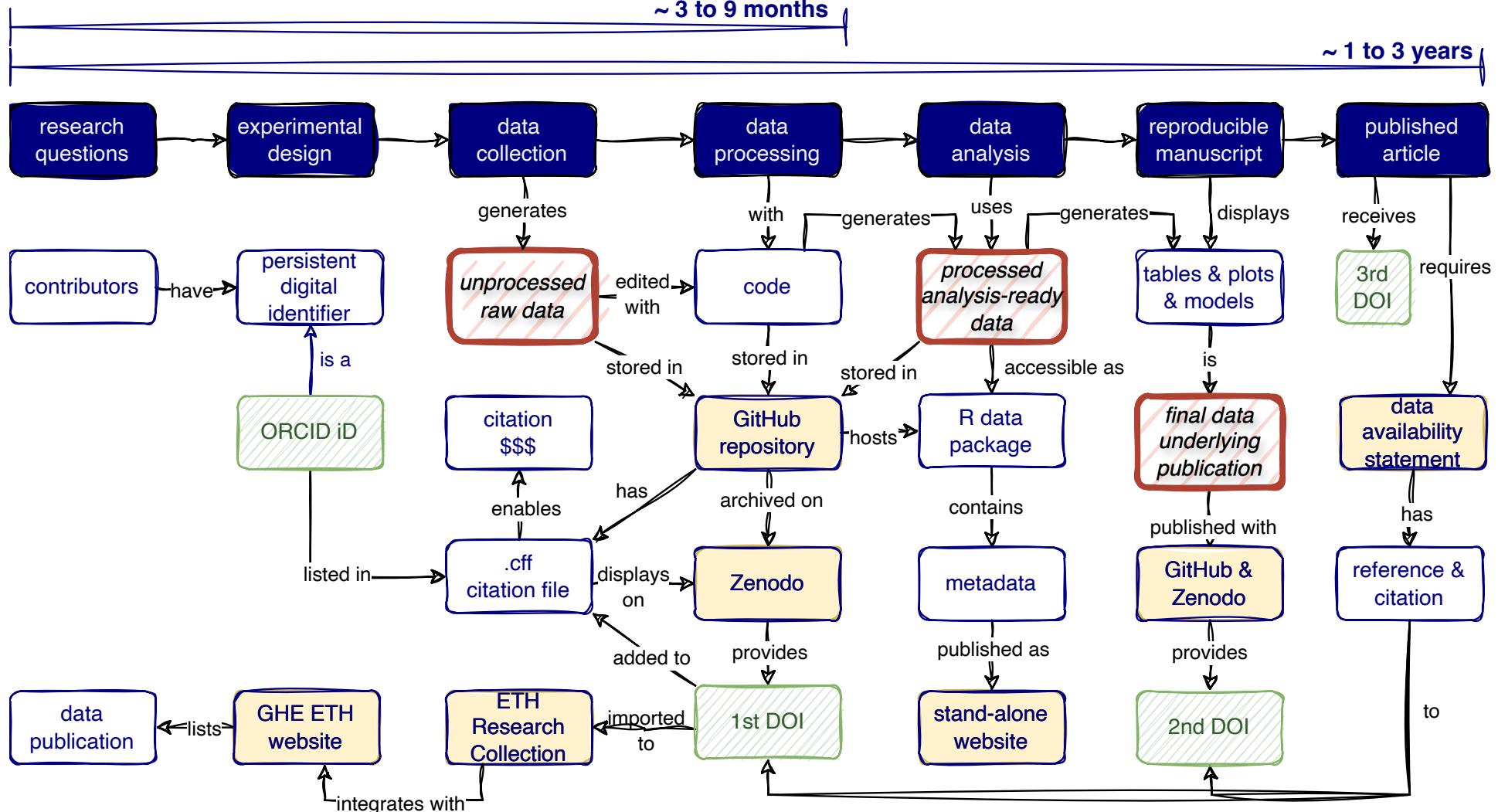












Invest now!

Funding schemes

swissuniversities

- 2021 - 2024: swissuniversities - Open Science I
- 2025 - 2028: swissuniversities - Open Science II (CHF 30 million requested)
- Watch: Measure B5.2 - Professionalisation of ORD specialists and related services
- Newsletter sign-up:
<https://sympa.ethz.ch/sympa/subscribe/isci>

Funding schemes

ETH Board

- 2021 - 2024: ~ 100 projects funded (~ CHF 10 million in total)
- ORD 8th Contribute (30k) guidelines, expected due date in August 2024
- motivates early career scientists to prepare proposals
- Newsletter sign-up: <https://open-research-data-portal.ch/> (bottom of page)

Thanks



slides at:
ghe-open.ch

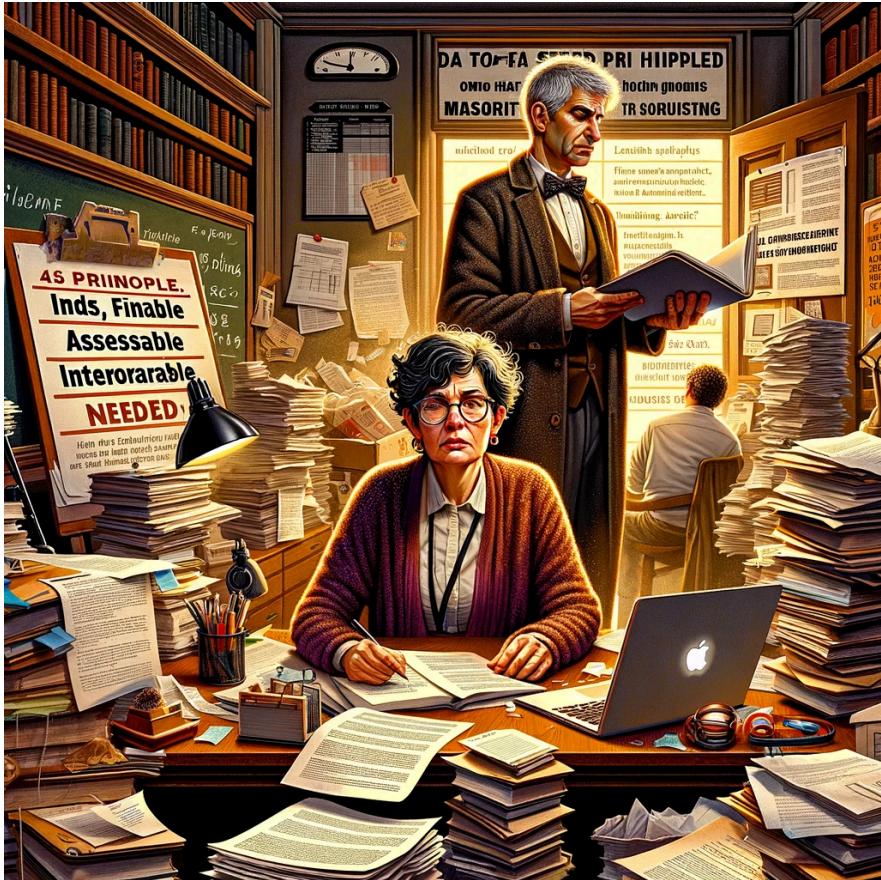
Slides created via revealjs and Quarto: <https://quarto.org/docs/presentations/revealjs/>

Source code of slides on GitHub: <https://github.com/Global-Health-Engineering/website/blob/main/slides/swissrn.qmd>

This material is licensed under [Creative Commons Attribution 4.0 International](#).

Additional slides

Meet a Prof



- working long hours
- sitting in committees
- stuck in meetings
- teaching, grading, supervising & mentoring

Meet a Prof



- working long hours
- sitting in committees
- stuck in meetings
- teaching, grading, supervising & mentoring

"We decided that you now need to share your data, following FAIR principles, and also publish your code for reproducibility." - the funding agency, the publisher

