

what's next?

data stewardship for WASH Center

Lars Schöbitz

Global Health Engineering, ETH Zurich

February 6, 2025

WASH Center: Computational Competency

Where do we stand?

- Computational competency is a key skill for solid research data management
- Survey: <https://forms.gle/XSyUJ5GrmzvXtTFN8>
- 24 responses

FAIR data publishing workshop

FAIR data publishing workshop

- 19 participants
- starting point to understand overall goals of FAIR data publishing
- hands-on experience with data publishing tools
- next step is building a foundation in R programming for team members

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

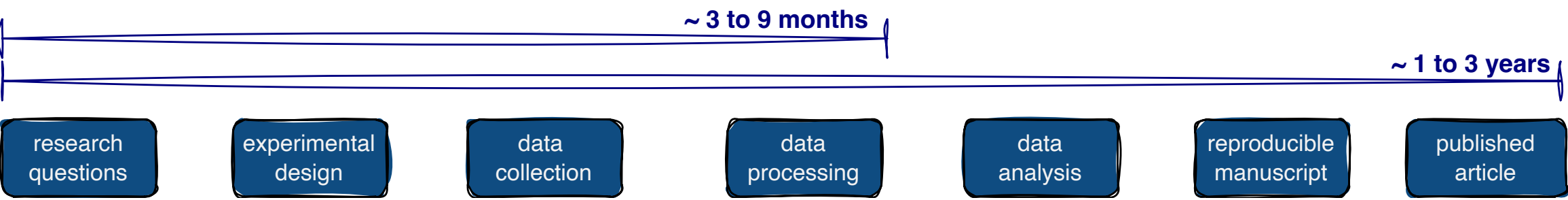
term	explanation	file format
unprocessed raw data	data that is not processed and remains in its original form and file type	often XLSX, also CSV and others
processed analysis-ready data	data that is processed to prepare for an analysis and is exported in its new form as a new file	CSV, R data package

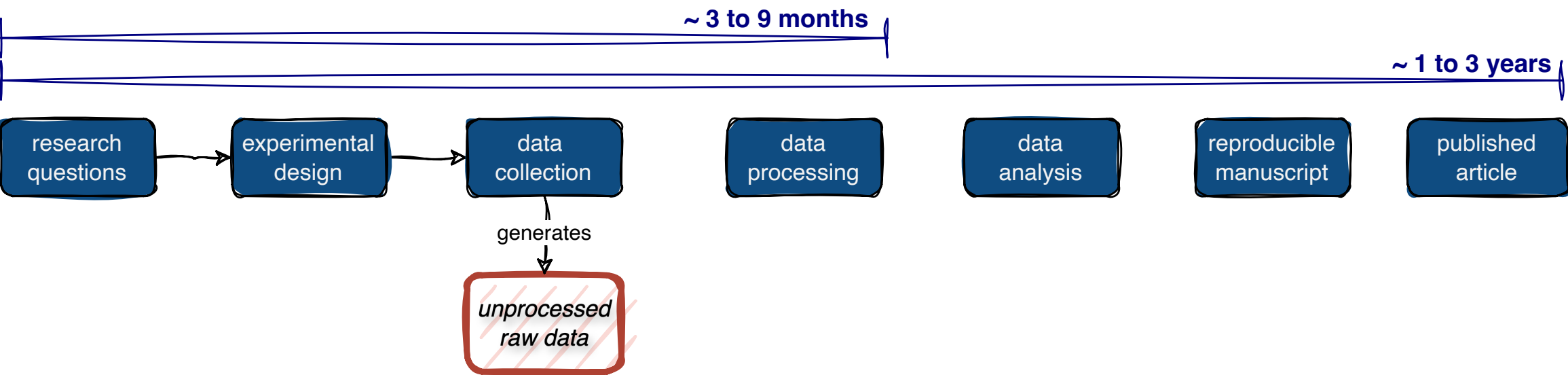
Three terms for three stages

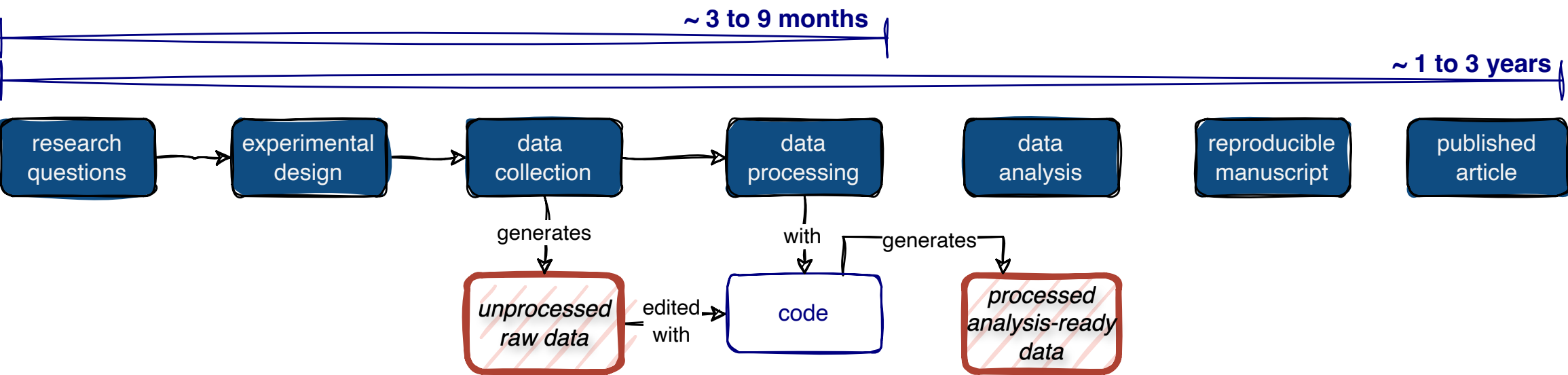
term	explanation	file format
unprocessed raw data	data that is not processed and remains in its original form and file type	often XLSX, also CSV and others
processed analysis-ready data	data that is processed to prepare for an analysis and is exported in its new form as a new file	CSV, R data package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

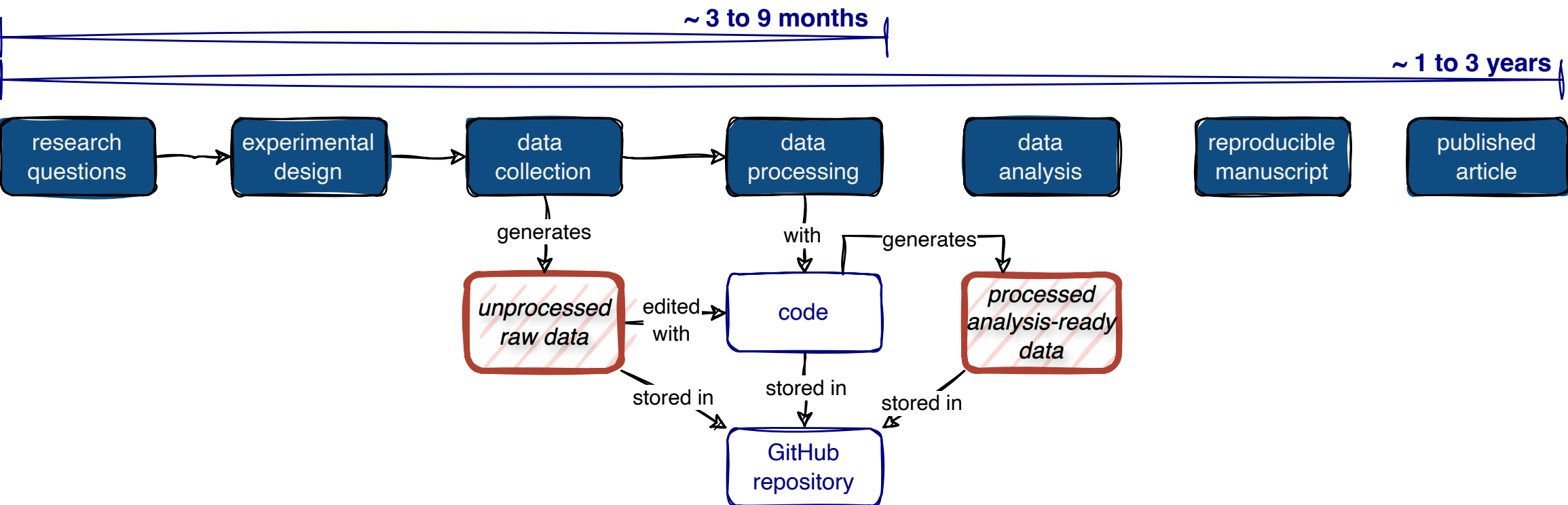
Data Management Strategy

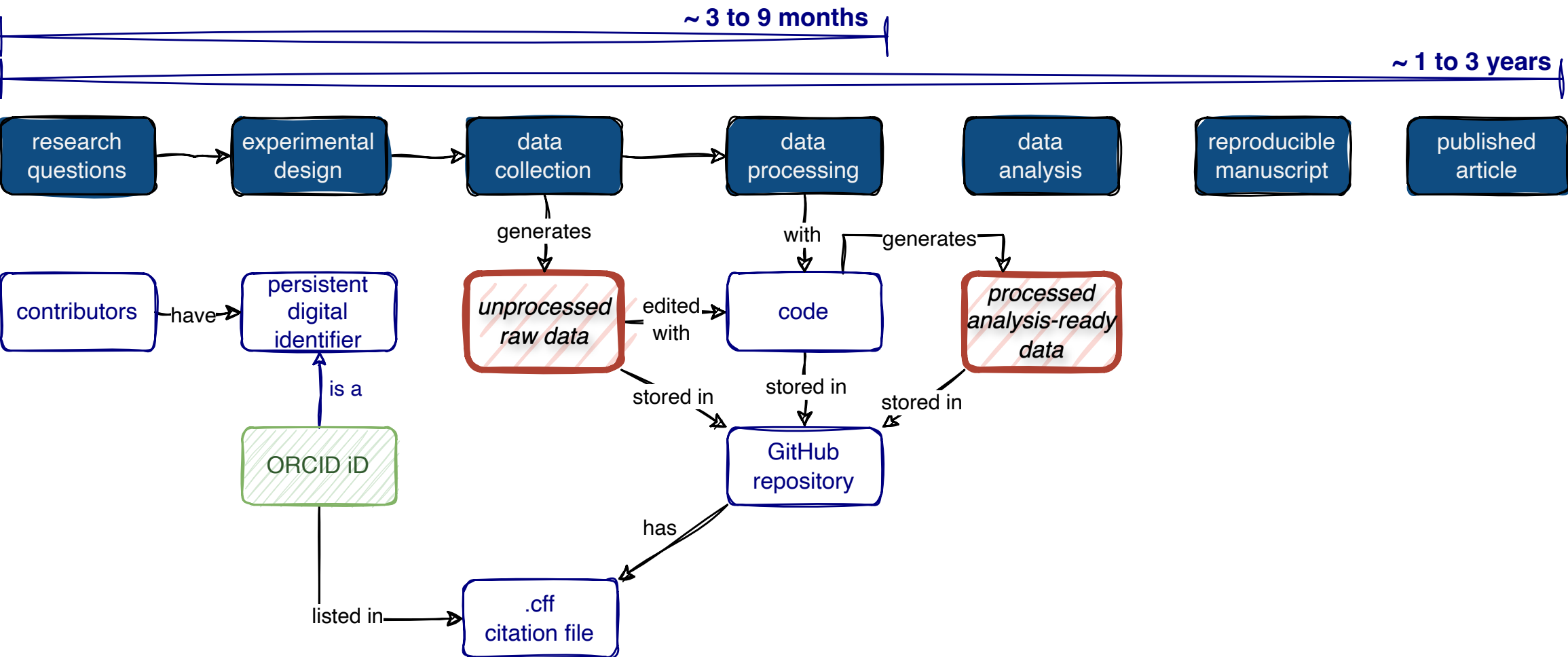


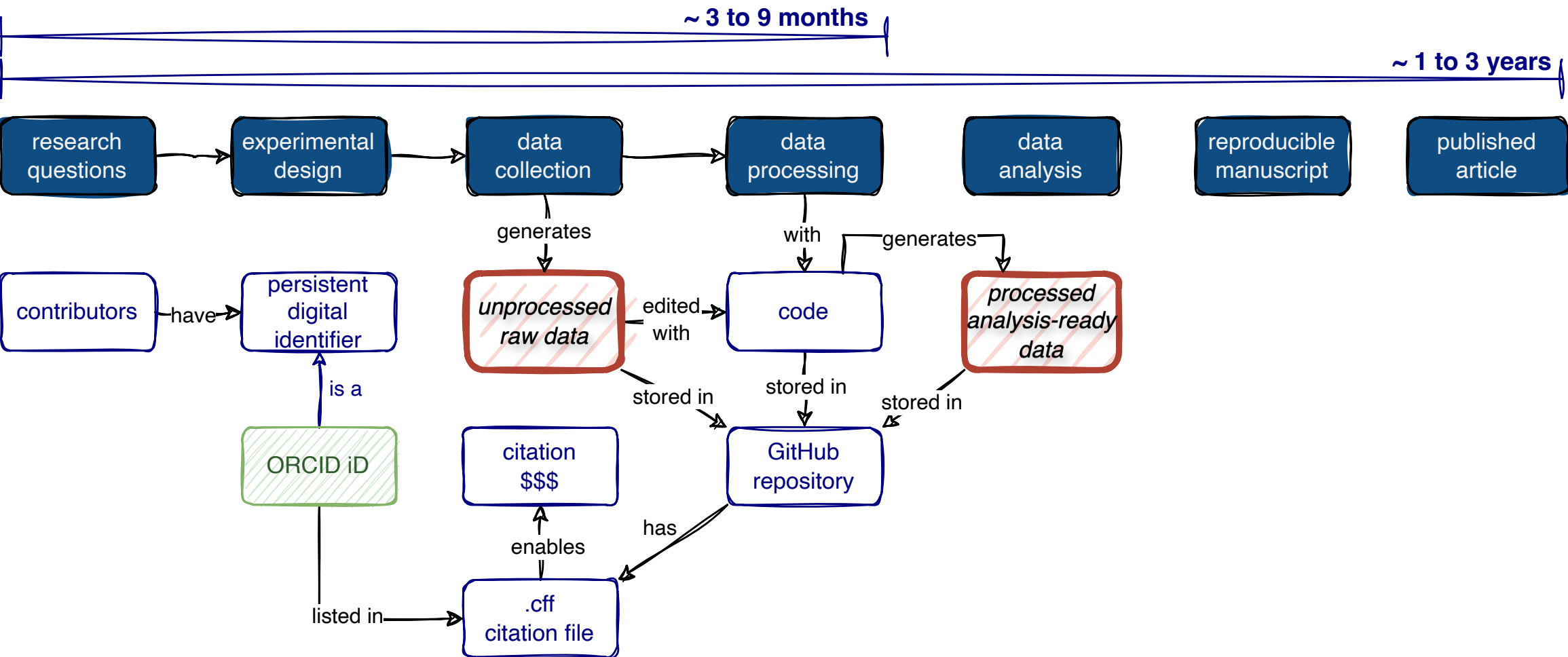


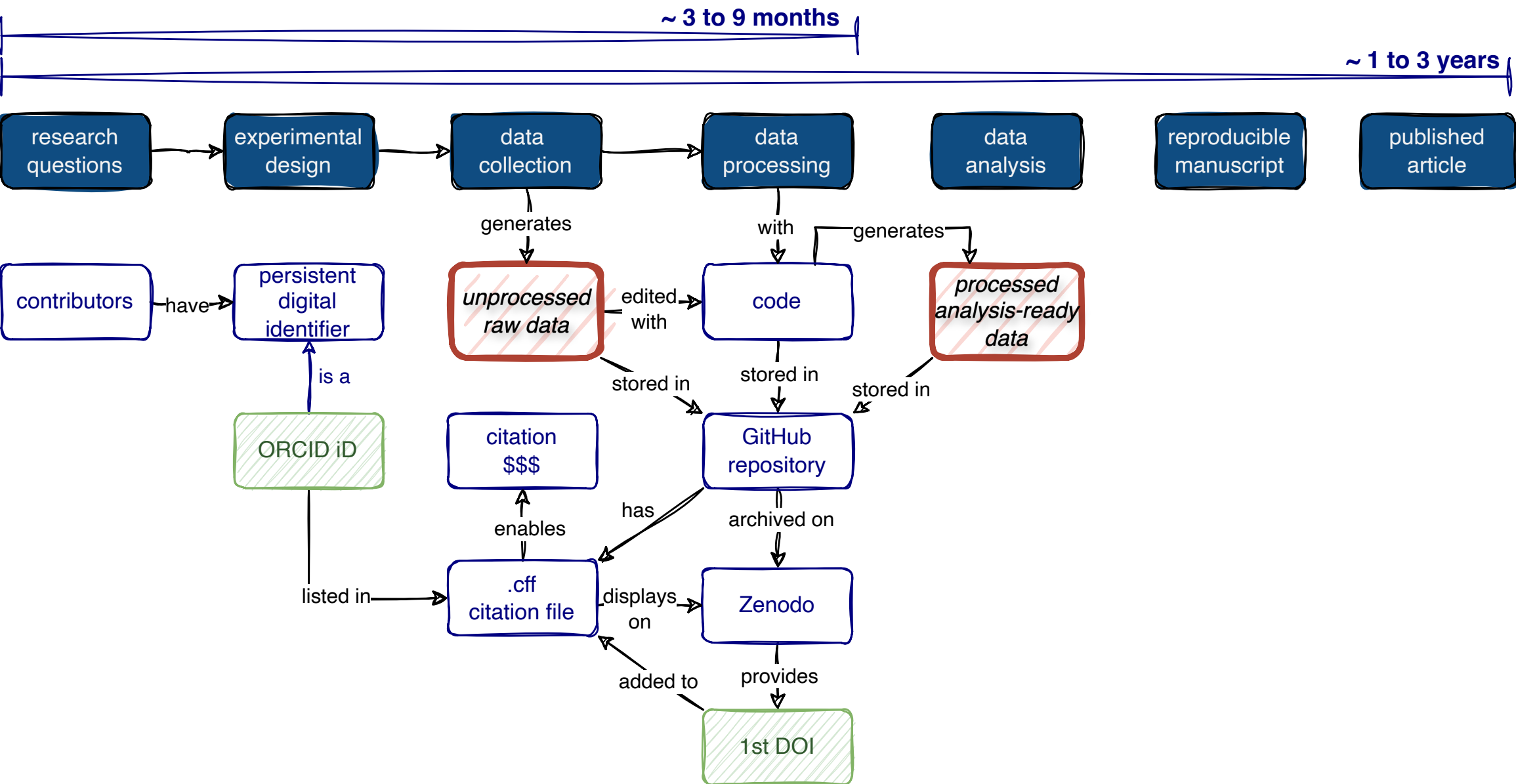


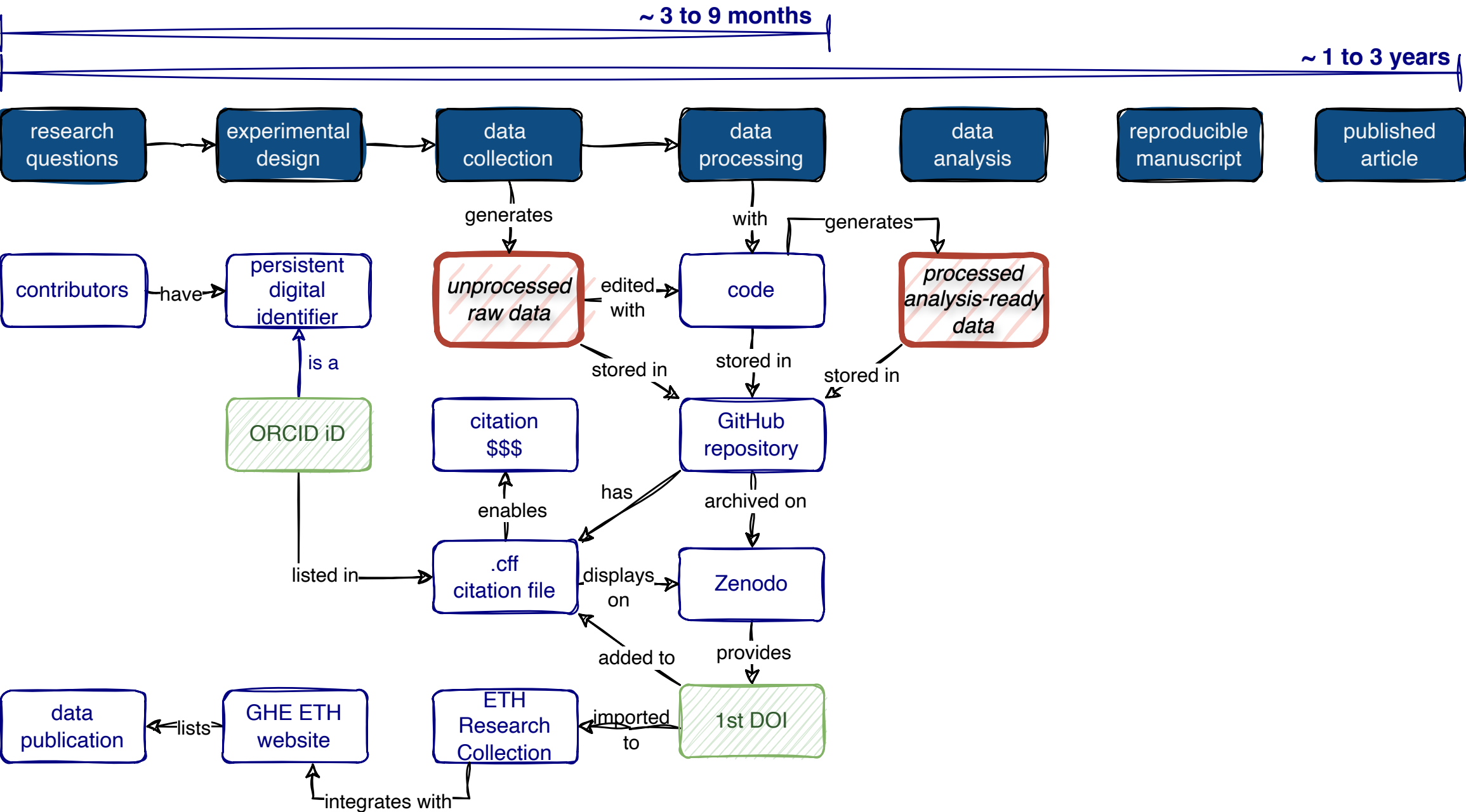


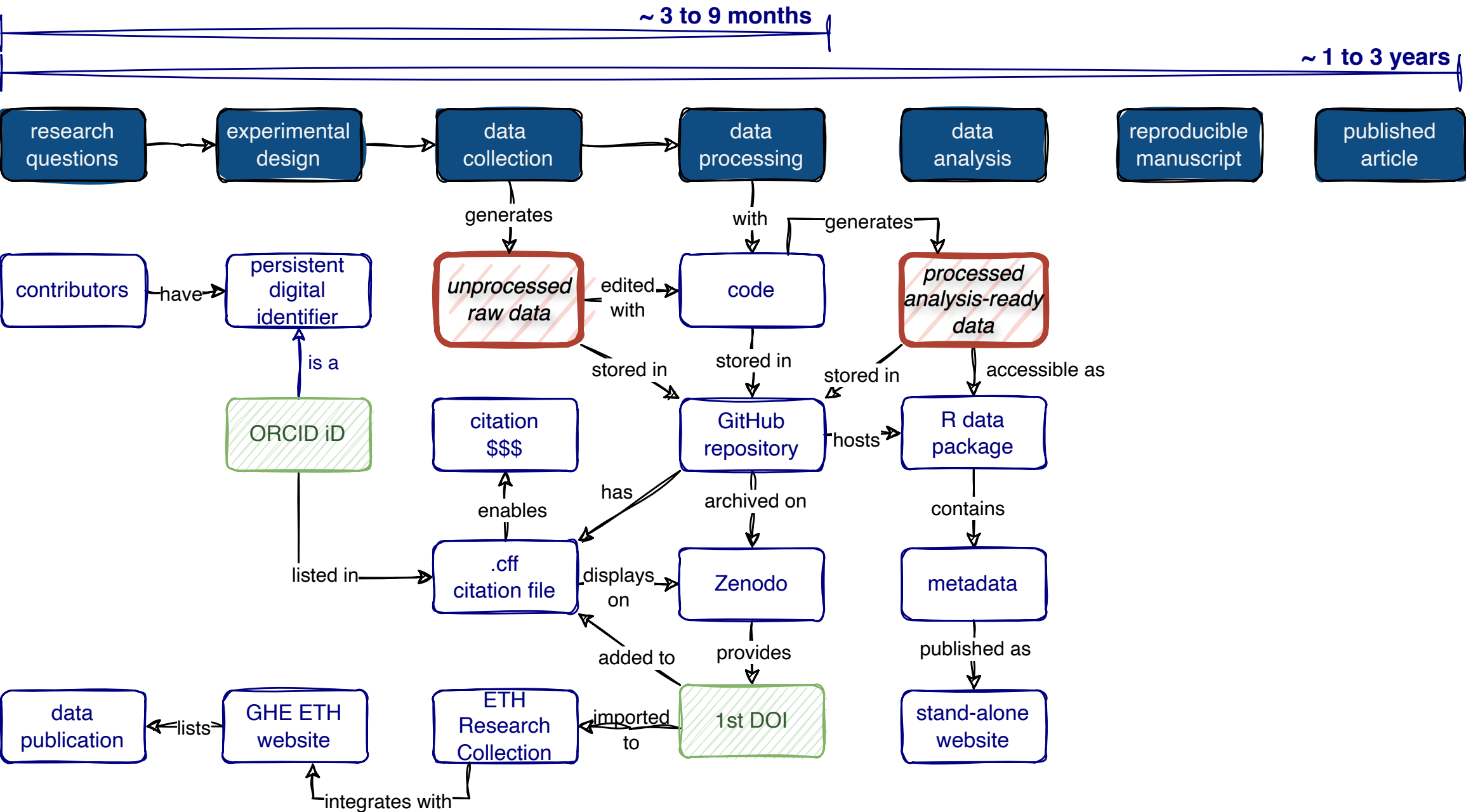


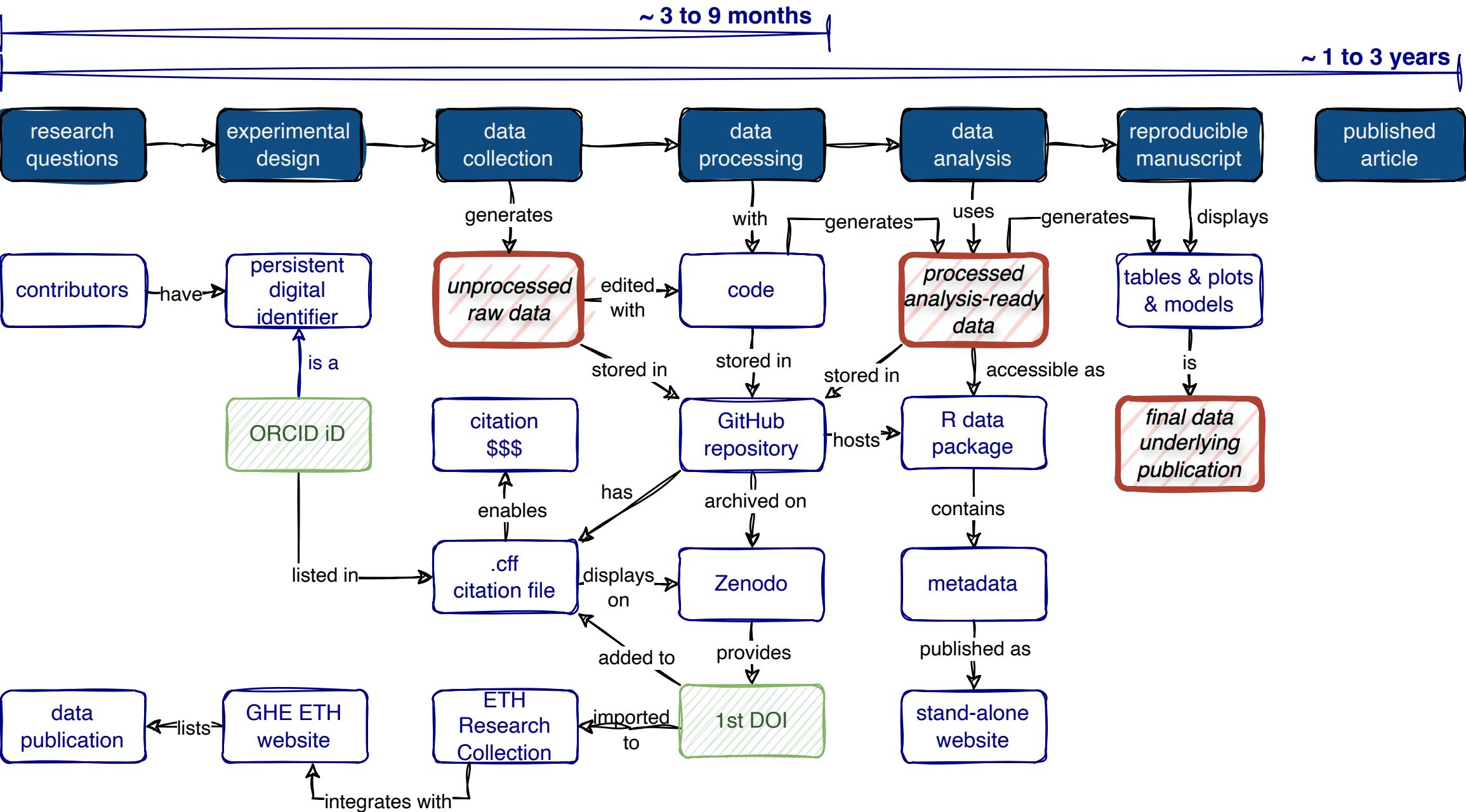


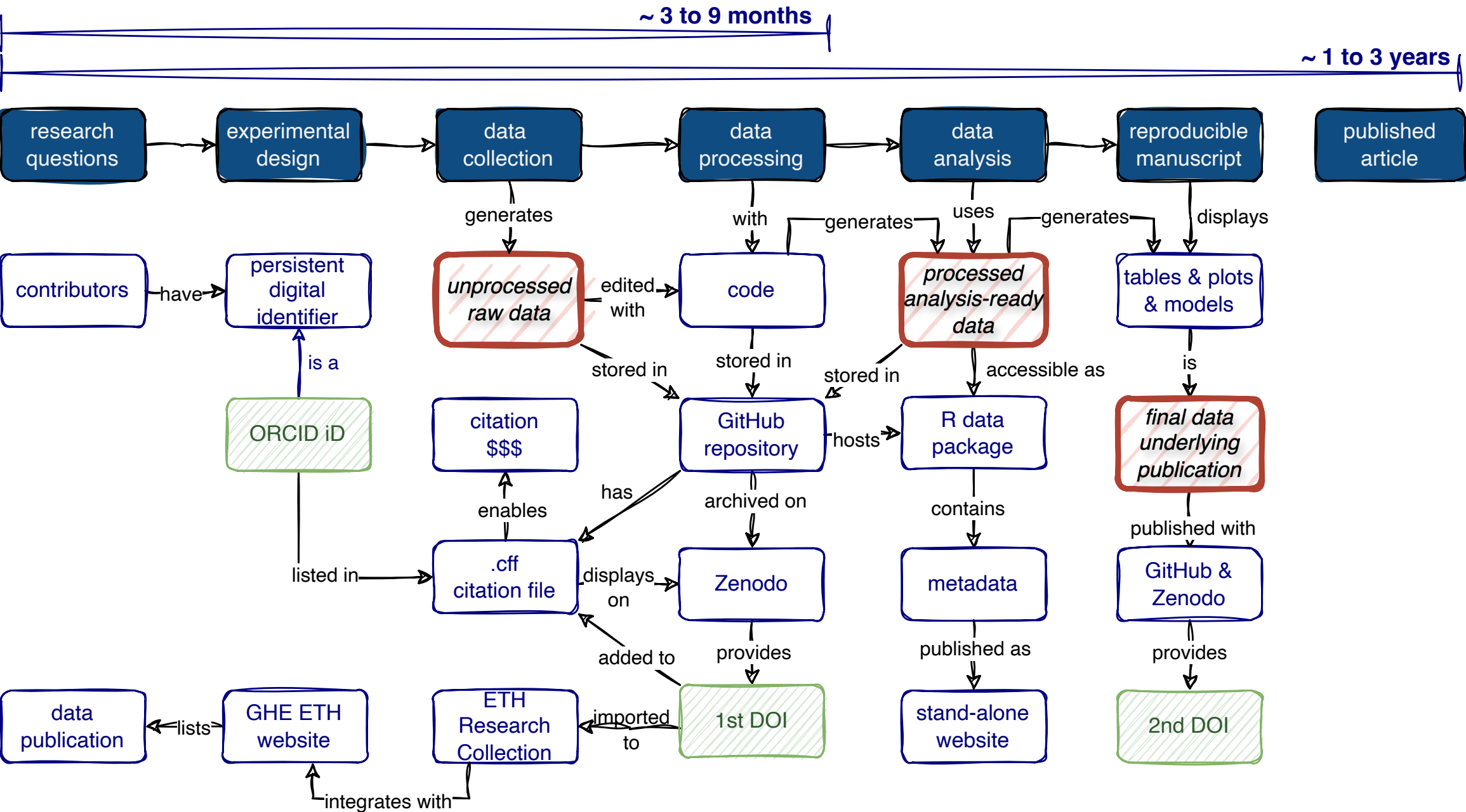


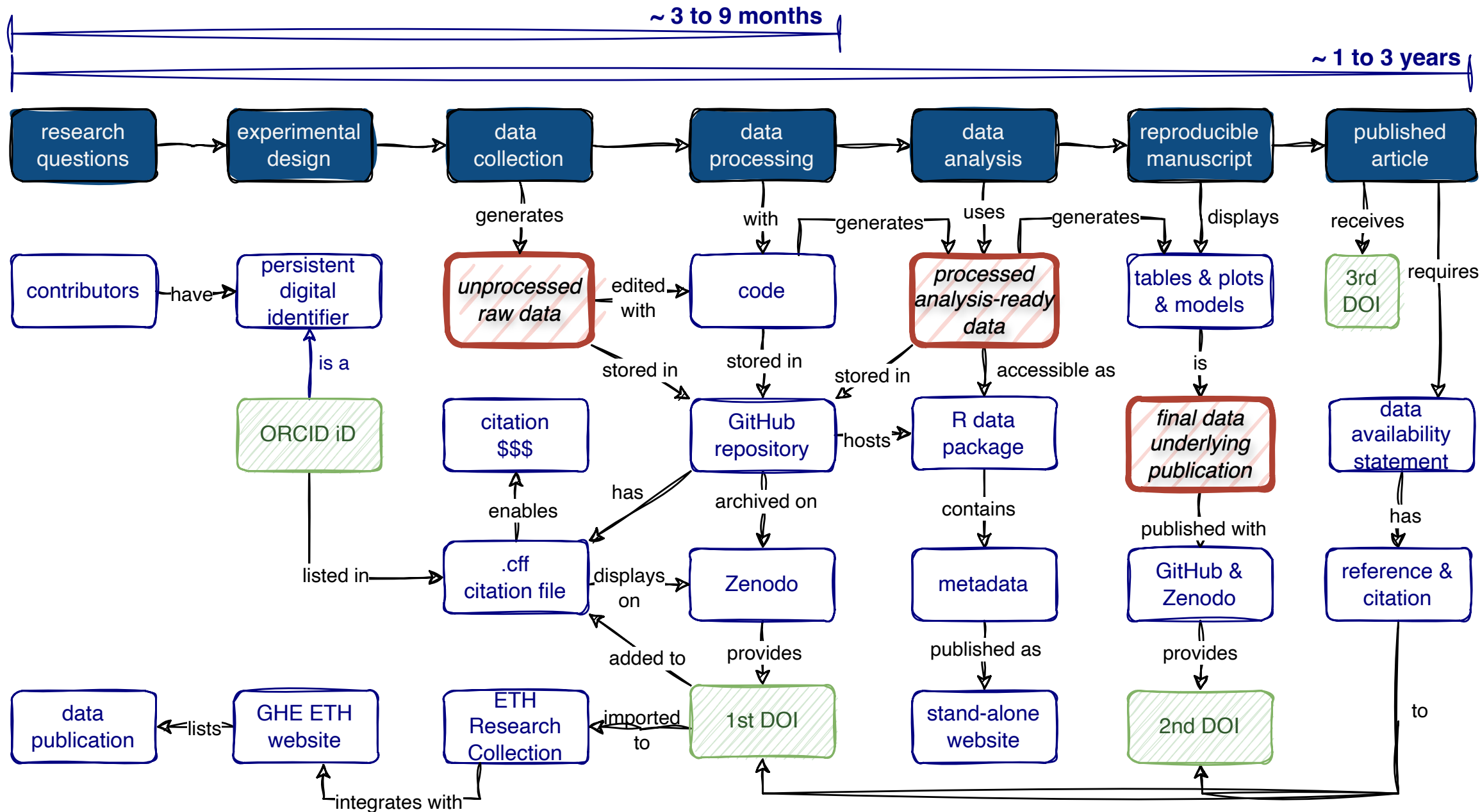


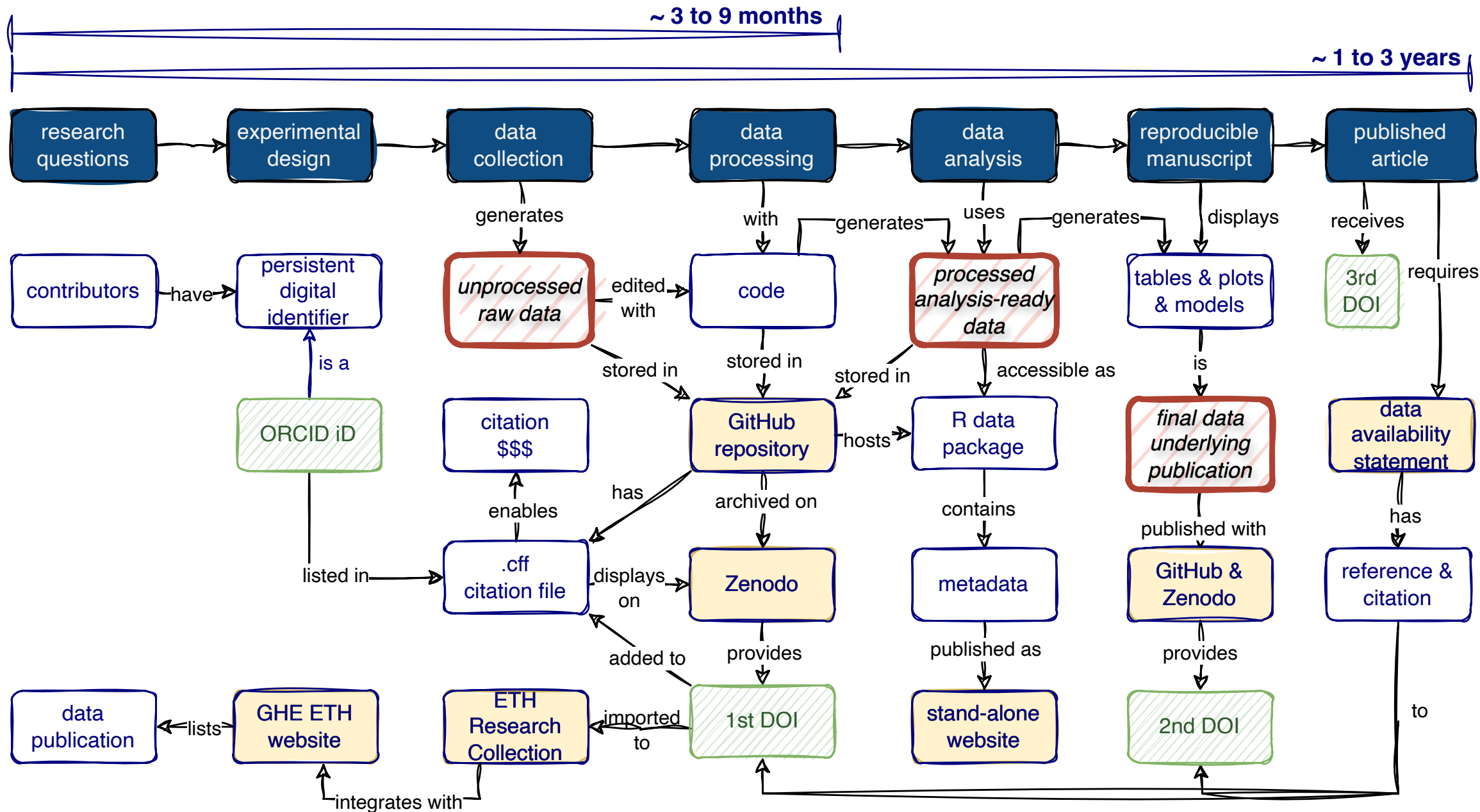












Data steward support

Data steward for WASH R&D Center

- A fully funded 2-year position, hopefully extended to 5-years with 3rd party funding
- 600 applications, screened to 6, invitations for interviews going out today
- Going through a 12-month programme together with data steward at NGO BASEflow in Malawi

Data steward activities (WP1)

- Activity 1.3: Identify how ethical approval for data collection differs for types of organizations (university, NGO) and types of data (quantitative, qualitative).
- Activity 1.4: Identify current data management practices and develop a draft data management strategy for organization.
- Activity 1.5: Publish at least 10 datasets of two different types that are available to the organization, following openwashdata data publishing workflow.

Data steward activities (WP1)

- Activity 1.3: Identify how ethical approval for data collection differs for types of organizations (university, NGO) and types of data (quantitative, qualitative).
- Activity 1.4: Identify current data management practices and develop a draft data management strategy for organization.

Future

- Activity 1.5: Publish at least 10 datasets of two different types that are available to the organization, following openwashdata data publishing workflow. **Past**

Data steward activities (WP3)

Community expansion

- Activity 3.1: Offer advanced data science training and workshops to community members.
- Activity 3.2: Develop a mentorship program to support new members in adopting ORD practices.
- Activity 3.3: Organize community events to foster networking and collaboration.

1st openwashdata forum/retreat/gathering

- Switzerland
- 23rd to 27th June 2025
- Data stewards (Switzerland, Malawi, South Africa)
- Programme to be defined

data science for openwashdata 002

At the end of the workshop, participants will be able to:

1. Be able to use a common set of data science tools (R, RStudio IDE, Git, GitHub, tidyverse, Quarto) to illustrate and communicate the results of data analysis projects.
2. Learn to use the Quarto file format and the RStudio IDE visual editing mode to produce scholarly documents with citations, footnotes, cross-references, figures, and tables.

data science for openwashdata 002

ds4owd-001.github.io/website/

Course Overview

Thank you for your interest in this course. Your course instructors: [Lars Schöbitz](#) & [Mian Zhong](#) & [Sophia Skorik](#) are looking forward to meet you.

We will meet on Zoom for 10 modules over 17 weeks (see Course Calendar below) at the following times:

- **Start:** 31st October 2023 - 2 pm to 4:30 pm CET
- **End:** 20th February 2024 - 2 pm to 4:30 pm CET

data science for openwashdata 002

- analytics dashboard: <https://openwashdata.github.io/dashboard/dashboard.html>
- analytics data: <https://github.com/openwashdata/analytics>
- graduates 2024: <https://openwashdata.org/pages/academy/graduates/>

data science for openwashdata 002

- free, live, online, 10-week programme
- next iteration from Thursday, 11th September 2025 at 13:30 to 16:00 (South Africa timezone)
- Sign-up: <https://forms.gle/MP5rNYZagBdfG2ZRA>

ds4owd-001.github.io/website/

Course Overview

Thank you for your interest in this course. Your course instructors: [Lars Schöbitz](#) & [Mian Zhong](#) & [Sophia Skorik](#) are looking forward to meet you.

We will meet on Zoom for 10 modules over 17 weeks (see Course Calendar below) at the following times:

- **Start:** 31st October 2023 - 2 pm to 4:30 pm CET
- **End:** 20th February 2024 - 2 pm to 4:30 pm CET

News

Support us: Sign up to our newsletter



<https://buttondown.email/openwashdata>

Your location

Your first name*

Your email (you@example.com)

Subscribe

Thanks

This project was supported by the [Open Research Data Program of the ETH Board](#).

The slides were created via revealjs and Quarto:
<https://quarto.org/docs/presentations/revealjs/>

You can [view source code of slides on GitHub](#)

Or you can [download slides in PDF format](#)

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