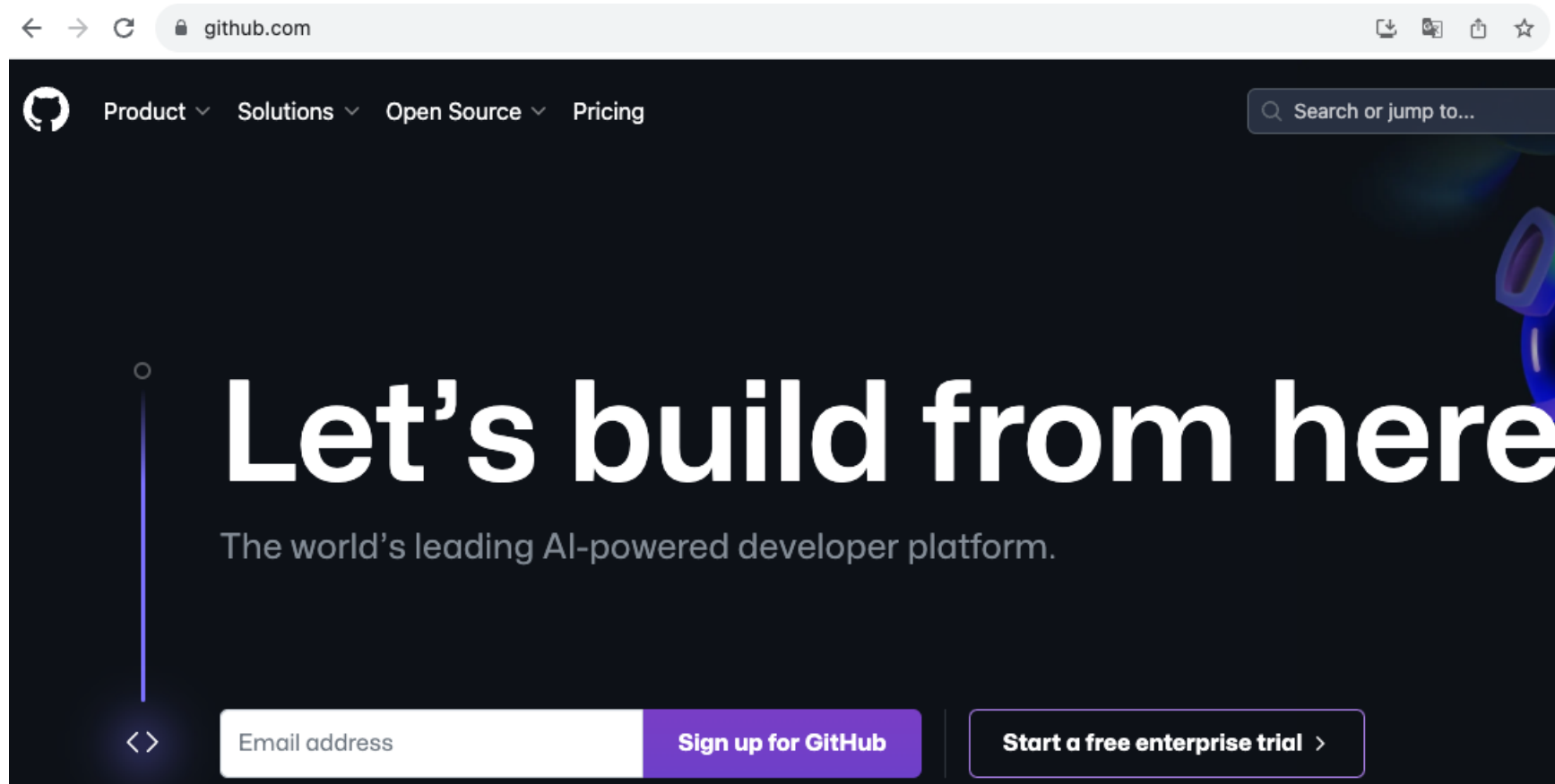


GitHub for Research (Data)



Agenda

- 1. Introduction**
2. Demo: Repository „ReproResearch“
3. Conclusions

Why „GitHub for research (data)“?

- Managing your research (data) and automating your research
 - Ensuring transparency and reproducibility of your research
 - Empowering collaboration across countries and disciplines

Research data

What is Data?

Data are objects that you use and produce during your research life cycle, encompassing datasets, software, code, workflow, models, figures, tables, images and videos, interviews, articles. Data are your research asset.

- **Data is not only a dataset**
- **Data is everything what makes your research (research articles) reproducible and replicable**

"The Turing Way: A handbook for reproducible, ethical and collaborative research"

<https://the-turing-way.netlify.app/reproducible-research/rdm/rdm-data.html>

Research and data in research

Two-step research:

1. Conducting research

- Data collection, creation or reuse (**dataset**)
- Data processing and analysis (**software, models, algorithms, methods, protocols, etc.**)



2. Publishing or perishing

- Publishing an article in a top-tier journal



Research data and code policies
of a journal => Data and code
availability statements
<=> **Replication package**



- Perishing without any research outputs



Research is not reproducible and
not replicable
<=> **No replication package**



Let's build from here

The complete developer platform to build, scale, and deliver secure software.

100+ million
Developers

4+ million
Organizations

420+ million
Repositories

90%
Fortune 100

Follow us on    

A version control and collaboration platform

- Git-based version control
- Pull requests and issue tracking
- GitHub Actions for automated workflows
- GitHub Pages for static website hosting
- Wiki for repositories
- Project management tools

The main concepts at GitHub

Repository (Repo): A folder where your project lives. It can be public or private.

Branch: A parallel version of a repository.

Commit: An individual change to a file or files.

Pull request: A method of submitting contributions to a project.

Merge: The process of taking the changes from one branch and integrating them into another.

Fork: A copy of a repository that you manage on your account.

Issue: A way to track tasks, enhancements, and bugs for your projects.

Clone: A copy of a repository that lives on your computer, and the act of making that copy.

Agenda

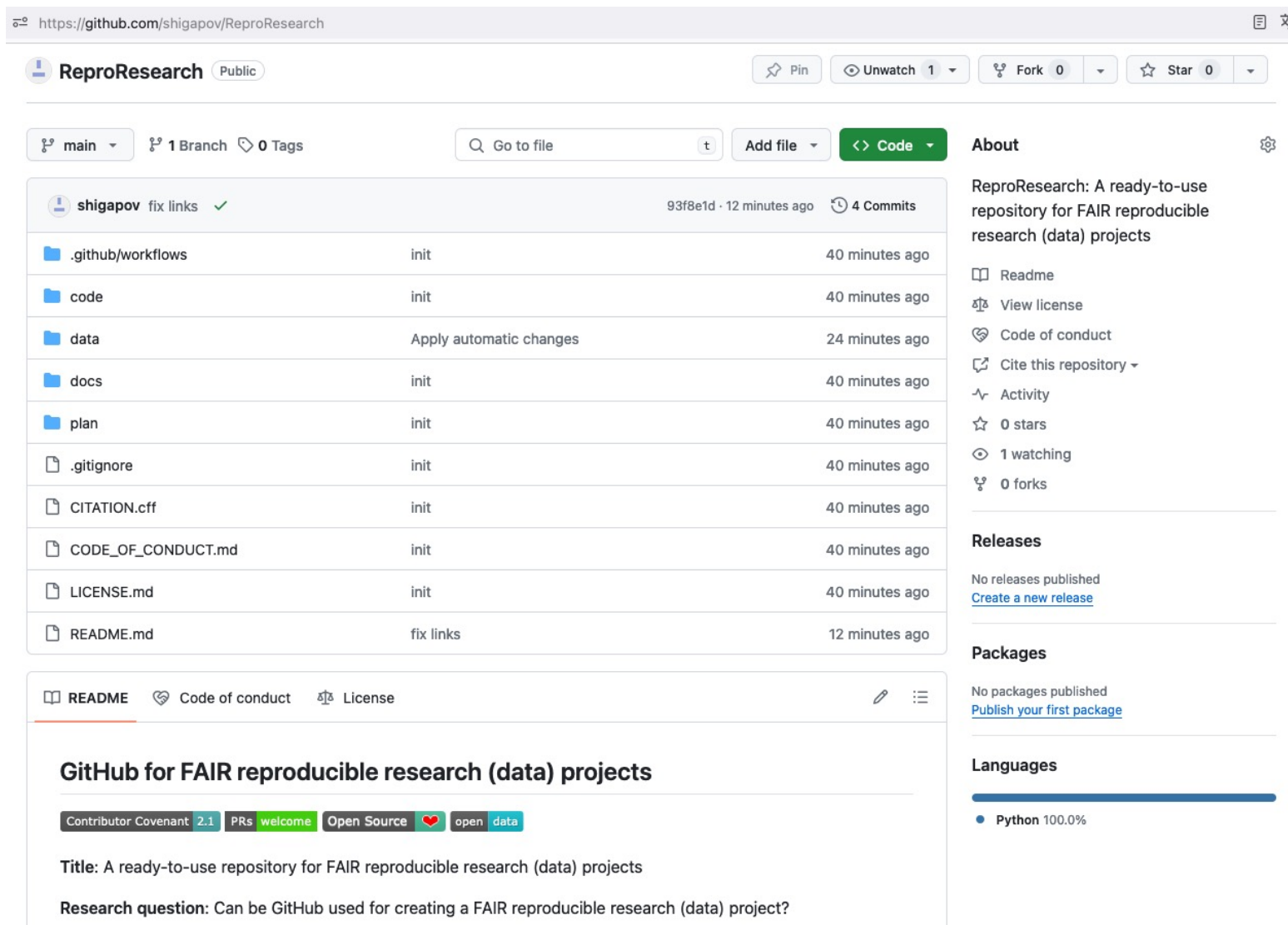
1. Introduction
- 2. Demo: Repository „ReproResearch“**
3. Conclusions

A ready-to-use repo for your research project

<https://github.com/shigapov/ReproResearch>

ReproResearch

- A ready-to-use or -fork or -clone
- A template for creating a replication package
- GitHub Actions for automated reproducible analysis
- GitHub Pages for static website hosting
- Wiki for repositories
- Project management board



https://github.com/shigapov/ReproResearch

ReproResearch Public

main 1 Branch 0 Tags

Go to file

Add file Code

About

ReproResearch: A ready-to-use repository for FAIR reproducible research (data) projects

Readme View license Code of conduct Cite this repository Activity 0 stars 1 watching 0 forks

Releases

No releases published [Create a new release](#)

Packages

No packages published [Publish your first package](#)

Languages

Python 100.0%

shigapov fix links ✓ 93f8e1d · 12 minutes ago 4 Commits

.github/workflows	init	40 minutes ago
code	init	40 minutes ago
data	Apply automatic changes	24 minutes ago
docs	init	40 minutes ago
plan	init	40 minutes ago
.gitignore	init	40 minutes ago
CITATION.cff	init	40 minutes ago
CODE_OF_CONDUCT.md	init	40 minutes ago
LICENSE.md	init	40 minutes ago
README.md	fix links	12 minutes ago

README Code of conduct License

GitHub for FAIR reproducible research (data) projects

Contributor Covenant 2.1 PRs welcome Open Source open data

Title: A ready-to-use repository for FAIR reproducible research (data) projects

Research question: Can be GitHub used for creating a FAIR reproducible research (data) project?

Agenda

1. Introduction
2. Demo: Repository „ReproResearch“
- 3. Conclusions**

What we have:

- A ready-to-use GitHub repository structured according to the best practices in research data management: <https://github.com/shigapov/ReproResearch>
- An automatic workflow using GitHub Actions for processing raw data, saving it, and making commit and push: <https://github.com/shigapov/ReproResearch/tree/main/.github/workflows>
- A project website using the Minimal Jekyll theme with GitHub Pages deployed via GitHub Actions: <https://shigapov.github.io/ReproResearch>
- A project wiki using GitHub Wiki (it needs separate cloning): <https://github.com/shigapov/ReproResearch/wiki>
- An open Kanban board for project management: <https://github.com/users/shigapov/projects/1>

Getting started with your research project at GitHub

1. You can use the ReproResearch repository as a template for your research project repository
2. Make your changes locally
3. Create your new repo at GitHub
4. Follow the instructions at the creation page.
5. Commit and push your local repo to your new repo at GitHub.