Codefair: Easily Make Biomedical Research Software Findable, Accessible, Interoperable, Reusable (FAIR)

Dorian Portillo, BS[1], Sanjay Soundarajan, MS[1], Bhavesh Patel, PhD[1]

[1]FAIR Data Innovations Hub, California Medical Innovations Institute, San Diego, CA, USA

Background

Research today is powered by software, whether it's a simple script or a complex machine learning model.

Making research software Findable, Accessible, Interoperable, and Reusable (FAIR) is critical to enable reproducible, transparent research and prevent duplicate effort.

This requires several steps, as specified by the FAIR-BioRS guidelines, such as selecting a clear usage license, following best coding practices, documenting software, Including essential metadata files, and archiving software.

Problem

Researchers **lack awareness, training, and time** for making research software FAIR. Most researchers are not formally trained in software development so this could also become **challenging** for them.

Solution: Codefair

To tackle these challenges, we are developing Codefair, **a free and open source GitHub App** that can be installed on any GitHub repository and acts as a personal assistant for making biomedical research software FAIR. Codefair is designed to be easy to use:

Install codefair-app from the GitHub Marketplace

Code as usual

Address FAIR compliance issues through the Codefair GitHub issue dashboard and web platform

Installed on **400+** repositories

35+ Individual and org users

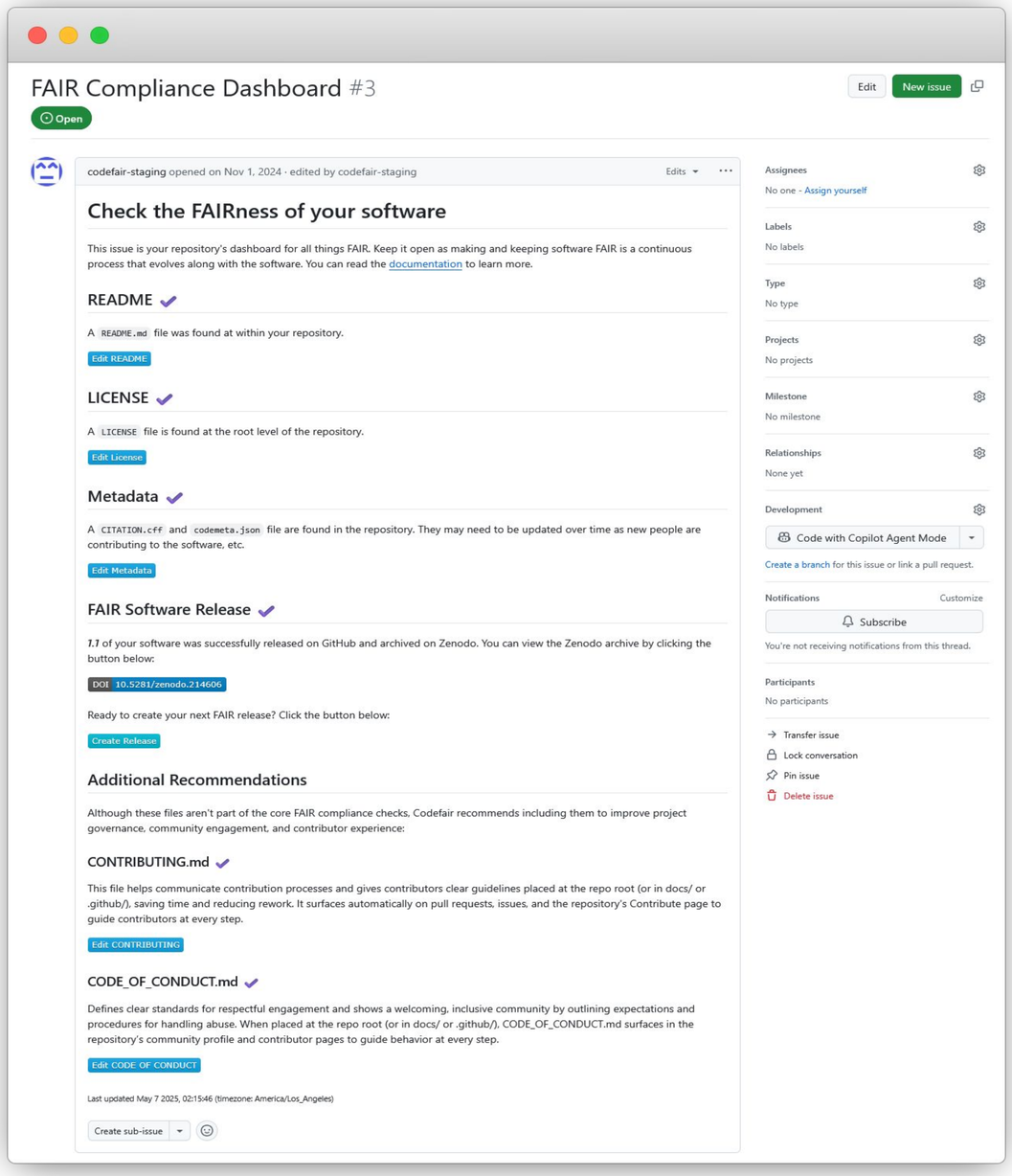

Figure 1. Screenshot of Codefair's GitHub issue dashboard

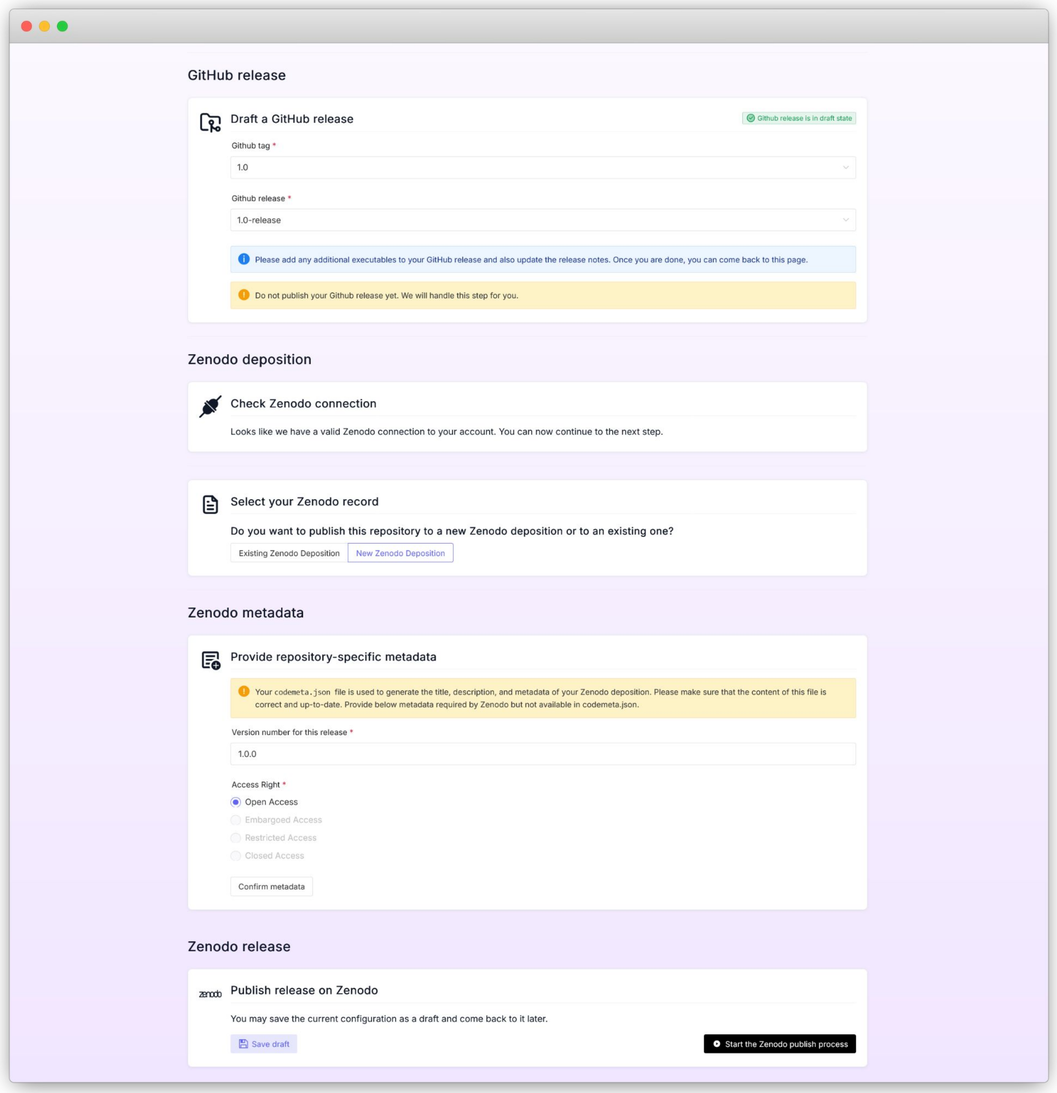

Figure 2. The codefair user interface for easily archiving software on Zenodo

Features

By leveraging GitHub tools such as Probot and GitHub's API, Codefair actively monitors repositories, continuously gathers essential metadata, and flags FAIR compliance issues via a GitHub issue dashboard (**Figure 1**).

For each compliance check, a link is provided in the GitHub issue to allow users to address the issue through an intuitive interface and submit a pull request (**Figure 2**).

Codefair provides several features for making software FAIR, including some newly implemented since the presentation at BOSC 2024:

Add a standard license

Include essential metadata files: README, CITATION.cff, codemeta.json, CODE_OF_CONDUCT.md, CONTRIBUTING.md

Archive software on Zenodo with your GitHub releases

Get validation reports for Common Workflow Language (CWL) files

Future work

Additional features are being added to provide support for complying with best practices, using LLM assistance, registering on bio.tools, and much more to cover all the requirements for making software FAIR.

The impact of Codefair can be very broad as it can be extended to other fields outside of biomedical and also provide support for aspects outside of FAIR, such as software quality and security.

We believe Codefair is an essential tool for enabling software curation at scale and turning FAIR software into reality.

Try Codefair today at codefair.io!

This work is supported by Azure sponsorship credits granted by Microsoft's AI for Good Research Lab

Dorian Portillo
dportillo@calmi2.org
fairdataihub.org

Bhavesh Patel
bpatel@calmi2.org
fairdataihub.org

Find this poster and all related references here

bit.ly/codefairBOSC25