Making FAIR Fair to the Researchers

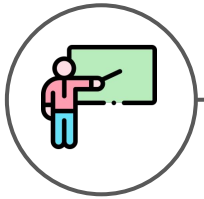
Bhavesh Patel, Ph.D.

Research Professor

CALIFORNIA MEDICAL INNOVATIONS INSTITUTE

About This Presentation

20 min + Q&A

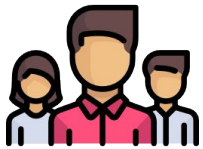


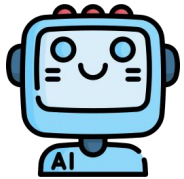
Share some lessons I have learned for making FAIR more fair

FAIR Principles

The FAIR Principles
Overview

The FAIR Principles (2016) provide instructions for making research outcomes reusable by humans and machines





https://doi.org/10.1038/sdata.2016.18

The FAIR Principles
Adoption

The FAIR Principles have been supported and adopted worldwide by all stakeholders including researchers, publishers, and funders

Example: The new Data Sharing Policy of the National Institutes of Health (NIH) effective January 2023 requires all grant proposals to include a plan for making data FAIR

https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html

My Encounter with FAIR

My Encounter with FAIR

The NIH SPARC Program

SPARC is an NIH Common Fund Program aimed at accelerating the development of neuromodulation therapies (2015-2025)

$350M + awarded

97 projects

(2017)

https://commonfund.nih.gov/sparc

My Encounter with FAIR

The NIH SPARC Program

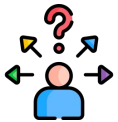
SPARC imposes strict guidelines to awardees for making data FAIR

- Organize data into a standard folder structure
- Follow file and folder naming conventions
- Include metadata using standard file formats and vocabularies
- Upload dataset on the SPARC data platform Pennsieve
- And more!

I was put in charge of doing this for the SPARC data at our Institute

https://doi.org/10.1101/2021.02.10.430563

My Encounter with FAIR
Challenges



Difficult - I was not trained for this

Time-consuming - The guidelines are very extensive

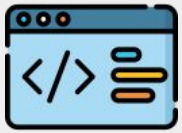
Boring - I wanted to do research, publish papers!

My Encounter with FAIR

Solution: Automation

Idea: Automate the process of preparing and sharing SPARC data

2017:
Started developing Python scripts for our group's data



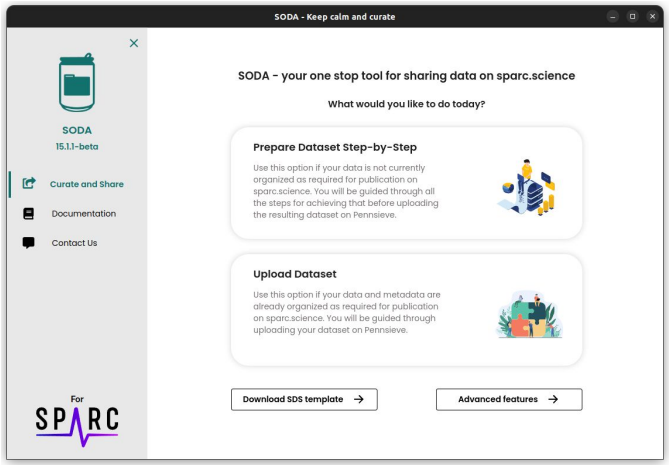
2018:
Presented the idea at the SPARC Hackathon

2019:
Received funding from SPARC for further development

My Encounter with FAIR

Software to Organize Data Automatically (SODA)

SODA is an open source and free desktop software that simplifies the preparation and sharing of datasets according to the SPARC guidelines



Like a tax filing tool but for filing FAIR SPARC datasets

My Encounter with FAIR

SODA Performance

Reduces time, errors, and difficulty

Has helped process 27+ TB of data (400k+ files) from 1000+ users (since 2021)

My Encounter with FAIR

SODA-inspired Tools

SODA has inspired many tools aimed at supporting researchers in making they data FAIR

FAIRshare

Making immunology related data FAIR (NIAID)

FAIRhub

Making clinical research data FAIR (NIH Bridge2AI)

GUIDE

Another team has forked SODA to help make neurophysiology data FAIR (NIH Brain Initiative)

https://fairdataihub.org

My Encounter with FAIR

Lesson Learned #1

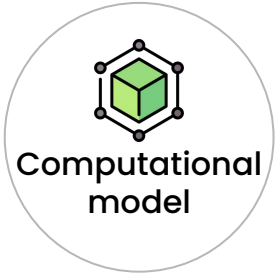
Tools that support researchers can help reduce researchers time and effort, and are needed to make FAIR more fair to them

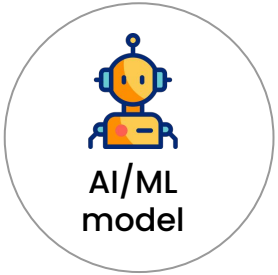
FAIR Research Software

FAIR Research Software

Definition of Research Software

Any software created during the research process or for a research purpose





https://doi.org/10.5281/zenodo.5504016

FAIR Research Software

FAIR4RS Principles

FAIR Principles for Research Software (FAIR4RS Principles, 2022):
17 principles tailored for research software

F1. Software is assigned a globally unique and persistent identifier.
F1.1. Components of the software representing levels of granularity are assigned distinct identifiers.
F1.2. Different versions of the software are assigned distinct identifiers.
F2. Software is described with rich metadata.
F3. Metadata clearly and explicitly include the identifier of the software they describe.
F4. Metadata are FAIR, searchable and indexable.

(and more)

https://doi.org/10.1038/s41597-022-01710-x

FAIR Research Software
Problem

The FAIR4RS Principles by design
do not provide actionable instructions

How do I assign a unique identifier?

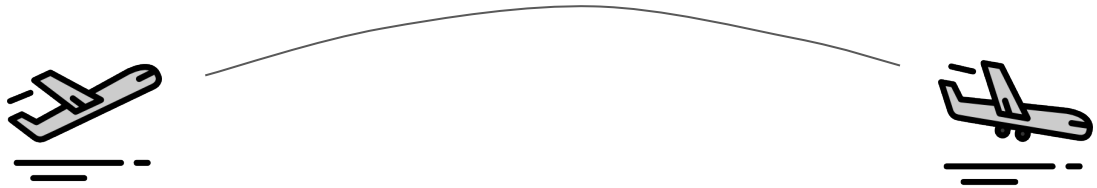
How do I provide rich metadata?

FAIR Research Software

FAIR-BioRS Guidelines

FAIR Biomedical Research Software (FAIR-BioRS) Guidelines

Minimal, actionable, step-by-step guidelines for complying with each of the FAIR4RS principles



December 2021
Beginning of this effort

August 2023
Manuscript published

https://doi.org/10.1038/s41597-023-02463-x

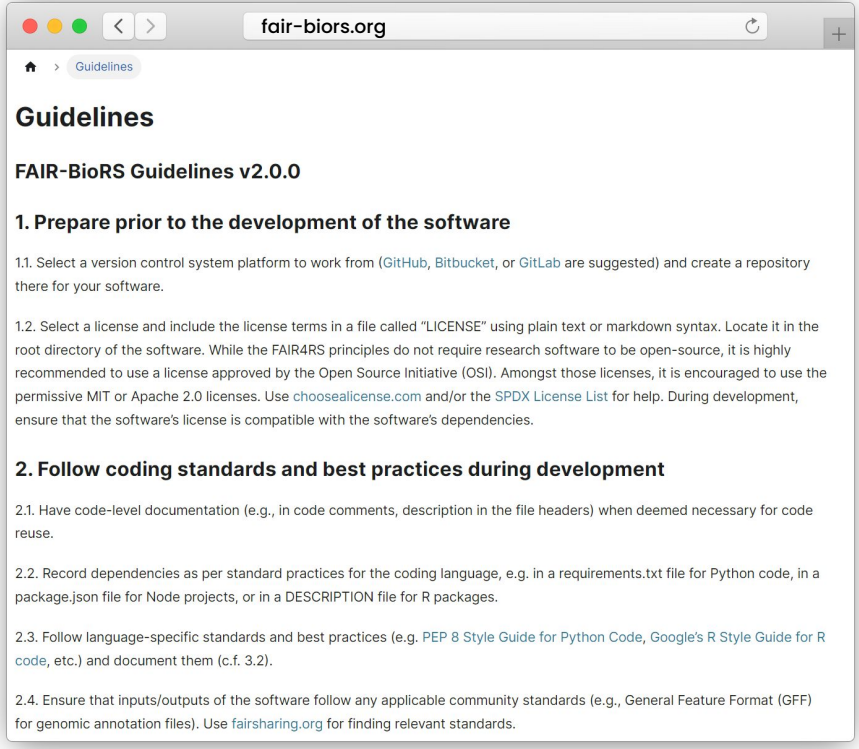
FAIR Research Software

FAIR-BioRS Guidelines Overview

Full guidelines: fair-biors.org

Step-by-step process

Clear & actionable instructions



FAIR Research Software

Lesson Learned #2

Actionable guidelines for making research outcomes FAIR can reduce burden on researchers and are needed to make FAIR more fair to them

FAIR Research Software

Codefair

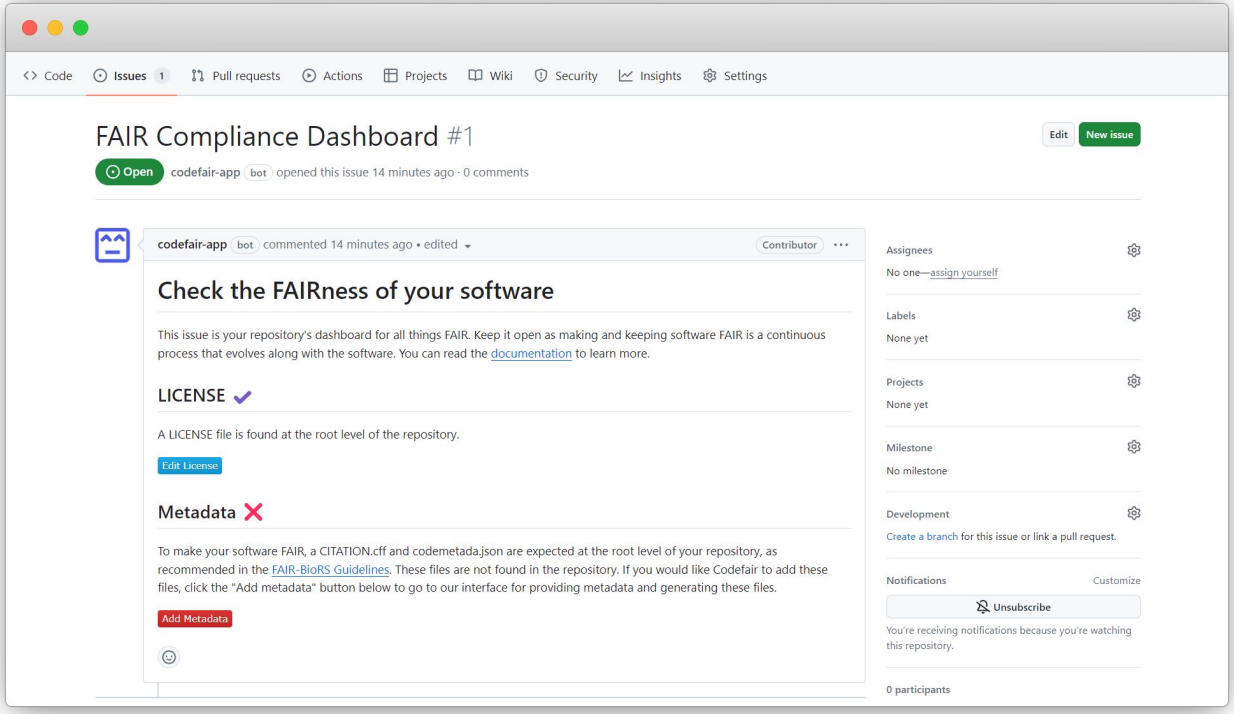
Codefair is a free and open source GitHub app that acts as your personal assistant for making software FAIR

1. **Install** codefair (codefair.io) on the GitHub repository of your software
2. **Develop** your software as usual from GitHub
3. **Track** and address FAIR compliance issues through the Codefair GitHub issue

https://codefair.io

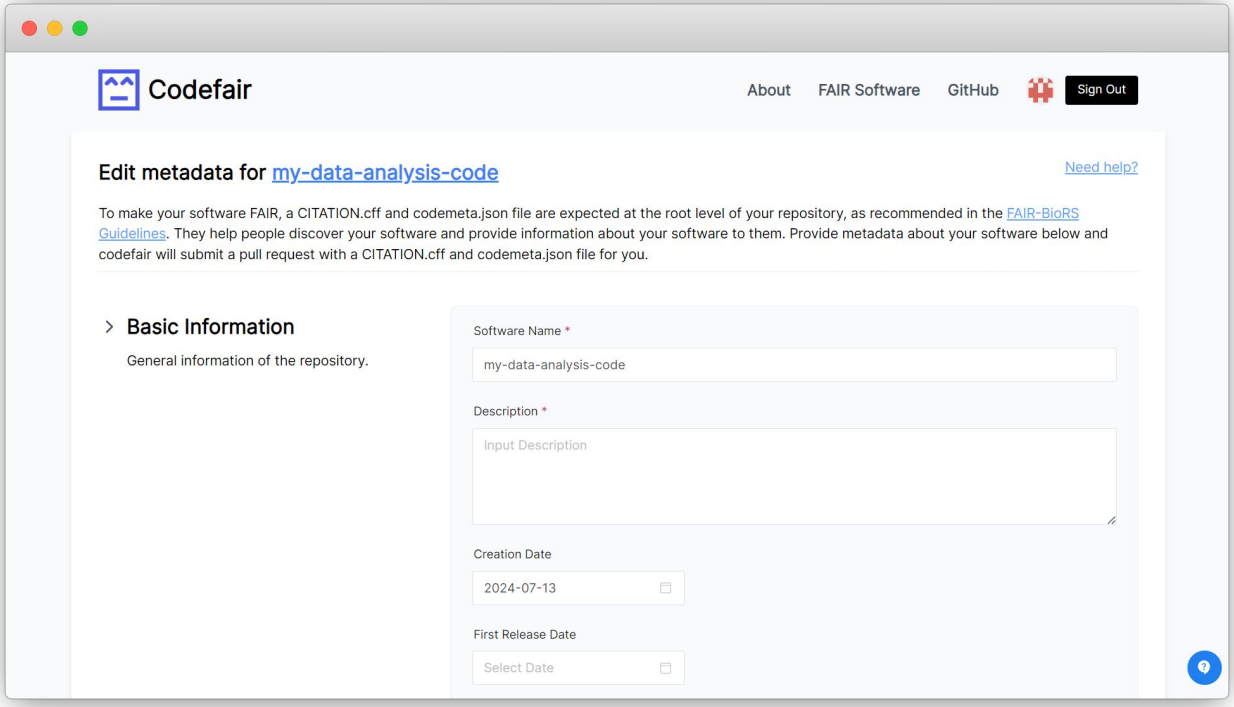
FAIR Research Software

Codefair Issue Dashboard



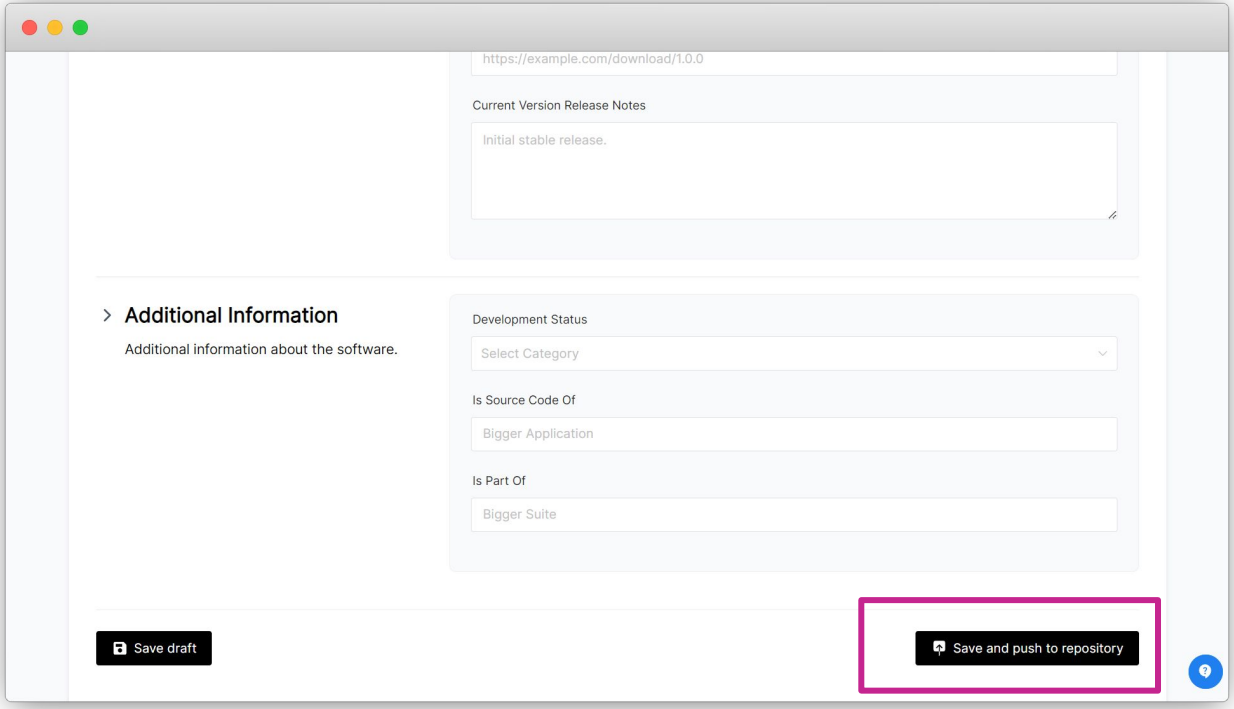
FAIR Research Software

Codefair UI



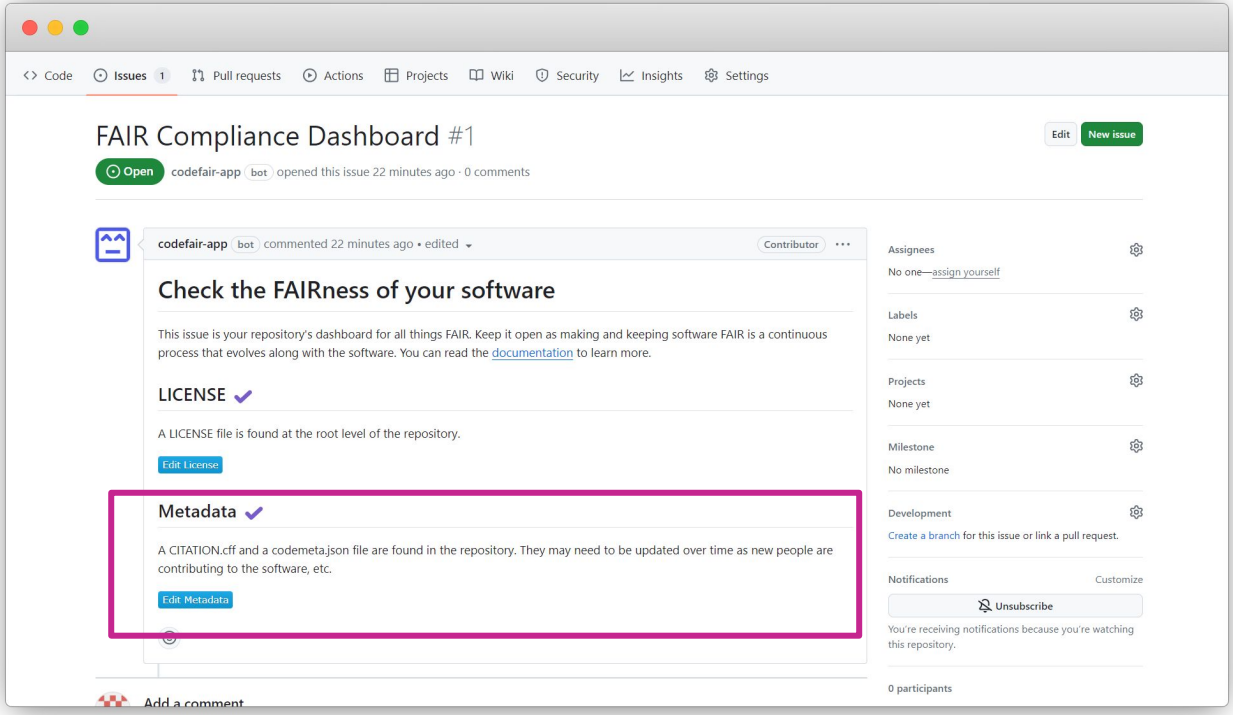
FAIR Research Software

Codefair UI



FAIR Research Software

Codefair Issue Dashboard



Closing Comments

Closing Comments

Requiring researchers to make their research outcomes FAIR can rapidly become unfair to them

FAIR can be made more fair to the researchers through actionable guidelines and tools for implementing them

Thank You!

bpatel@calmi2.org

fairdataihub.org

Find these slides and all resources here

bit.ly/FAIRfair