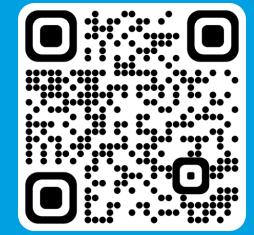
NLeSC Software Quality tools and some personal opinions.

Faruk Diblen

10 October 2022

netherlands
Science center



https://doi.org/10.5281/zenodo.7193991









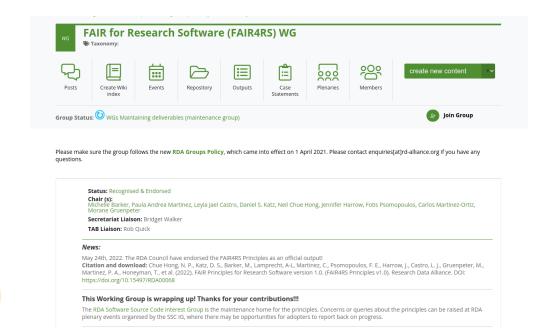


FAIR Principles

https://www.go-fair.org/fair-principles/

The FAIR Guiding Principles for scientific data management and stewardship

https://www.nature.com/articles/sdata201618



FAIR data is different than FAIR software

FAIR for Research Software (FAIR4RS) WG

https://www.rd-alliance.org/groups/fair-4-research-software-fair4rs-wg

FAIR Principles for Research Software (FAIR4RS Principles)

https://zenodo.org/record/6623556





FIVE RECOMMENDATIONS FOR FAIR SOFTWARE

LET'S GO! →

F A I R

WHAT'S FAIR ABOUT US https://fair-software.nl

F A I R

#1 USE A PUBLICLY ACCESSIBLE REPOSITORY WITH VERSION CONTROL

#2 ADD A LICENSE





- @NLeSC we prefer Apache 2.0 for software
- No default for data but usually CC BY (https://creativecommons.org/about/ cclicenses/)
- We deviate only when it is necessary





https://choosealicense.com/



License compliance is important!

The \$100 Million Court Case for Open Source License Compliance

- CoKinetic Systems Corporation, one of the major global players in the in-flight entertainment (IFE) market, has recently filed suit against Panasonic Avionics Corporation in a New York federal court, seeking damages of over \$100 million.
- CoKinetic claim that Panasonic, which holds a roughly 70% share of the embedded IFE hardware market, has willfully violated GPL v2 open source licensing requirements, in addition to a long list of other unlawful actions aimed at monopolizing the market for in-flight entertainment software and media services.

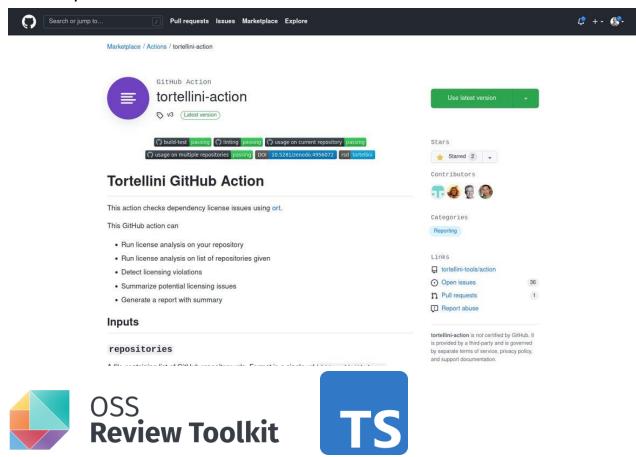
Vizio sued for breach of copyleft opensource software license

 On October 19, 2021, the Software Freedom Conservancy (SFC) sued Vizio, Inc. for alleged violations of the GNU General Public License covering software incorporated into certain Vizio smart TVs.

More examples:

https://en.wikipedia.org/wiki/Open source license litigation https://www.slashdata.co/blog/the-open-source-trialshanging-in-the-legal-balance-of-copyright-and-copyleft

Example: Akurat Font



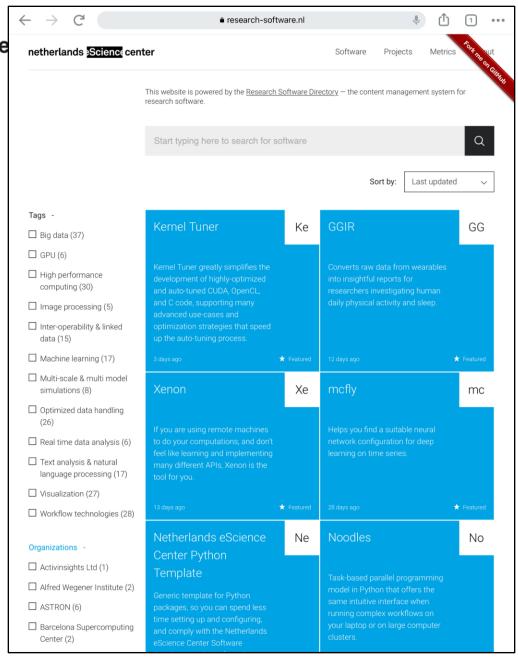
- https://github.com/oss-review-toolkit/ort
- https://github.com/tortellini-tools/action
- https://github.com/marketplace/actions/tortellini-action

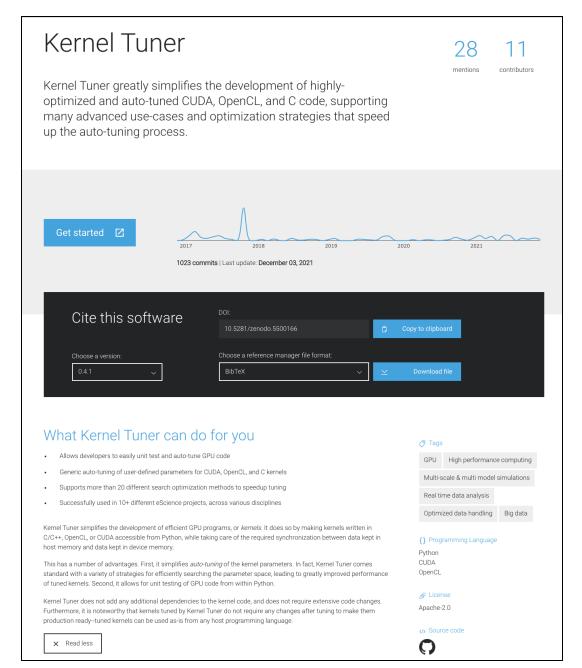
F A I R

REGISTER YOUR CODE IN A COMMUNITY REGISTRY

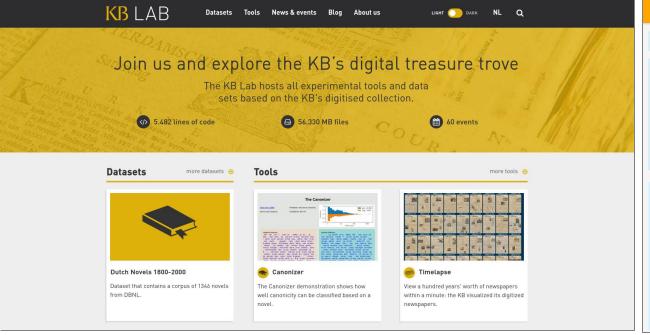






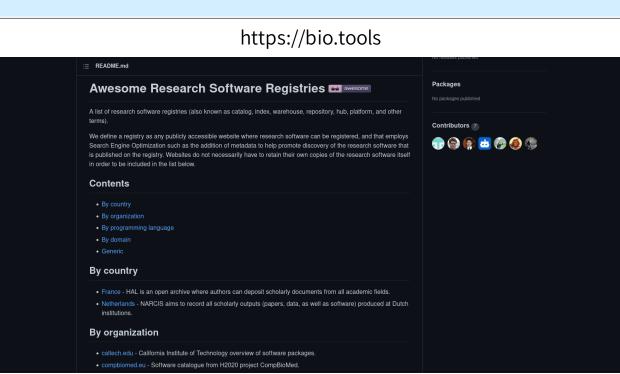






https://lab.kb.nl





Experts from scientific communities are key in identifying what are the methods and software resources used in their research activities. We rely upon scientific communities to improve the terminology and description

This site uses cookies. By continuing to browse the site you are agreeing to our use of cookies. Find out more here.

Explore communities in bio.tools and see what resources are used by different scientific communities.

We strongly encourage everyone to take part in the community effort and contribute. Email us for more info.

Explore and search bio.tools domains and manage your own domains. Read more about bio.tools domains here.

Essential scientific and technical information about software tools, databases and services for bioinformatics and the life sciences.

of resources in different domains of the Life Sciences. This includes the communities being fostered by ELIXIR, and any other group of scientists with an interest to help out.

In bio.tools we try, with the help of the community, to create content domains or "slices" relevant to a specific community, research area or task (e.g. Proteomics, Rare diseases, COVID-19).

bio.tools

Communities in bio.tools

Communities and domains

Search bio.tools

https://github.com/NLeSC/awesome-research-software-registries



Awesome Research Software Registries

- Collection of research software registries organized by
 - Country
 - Organization
 - Programming language
 - Scientific domain
- Collective effort
- Please contribute!

https://github.com/NLeSC/awesome-research-software-registries





A list of research software registries (also known as catalog, index, warehouse, repository, hub, platform, and other terms).

We define a registry as any publicly accessible website where research software can be registered, and that employs Search Engine Optimization such as the addition of metadata to help promote discovery of the research software that is published on the registry. Websites do not necessarily have to retain their own copies of the research software itself in order to be included in the list below.

Contents

- Bv country
- By organization
- By programming language
- By domain
- Generic

By country

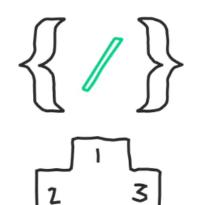
- France HAL is an open archive where authors can deposit scholarly documents from all academic fields.
- Netherlands NARCIS aims to record all scholarly outputs (papers, data, as well as software) produced at Dutch institutions

By organization

- caltech.edu California Institute of Technology overview of software packages.
- compbiomed.eu Software catalogue from H2020 project CompBioMed.
- darpa.mil US Defense Advanced Research Projects Agency (DARPA).
- b2find.eudat.eu Data and software discovery service based on metadata harvested from research data collections from EUDAT data centres and other repositories.
- kb.nl Dutch Royal Library.
- Ibl.gov Lawrence Berkeley National Lab.
 - Applied Mathematics Software
 - Computational Science Software
 - Computer Science Software
 - Data Science & Technology Software
- nasa.gov NASA software catalog.
- ontosoft.org Software from a variety of disciplines, annotated and searchable using the OntoSoft vocabulary for describing software.



#4 ENABLE CITATION OF THE SOFTWARE



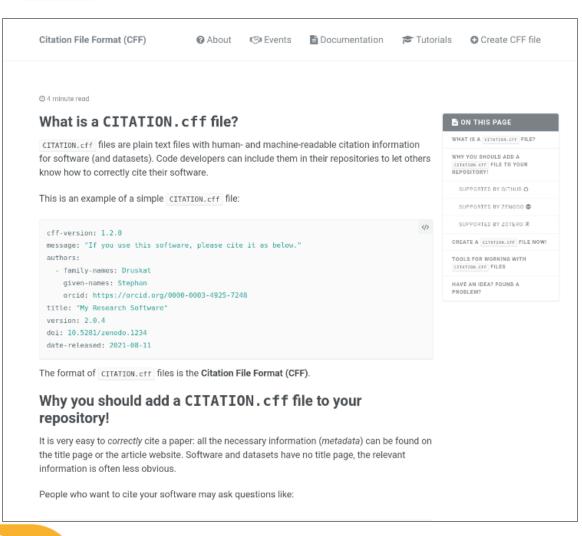


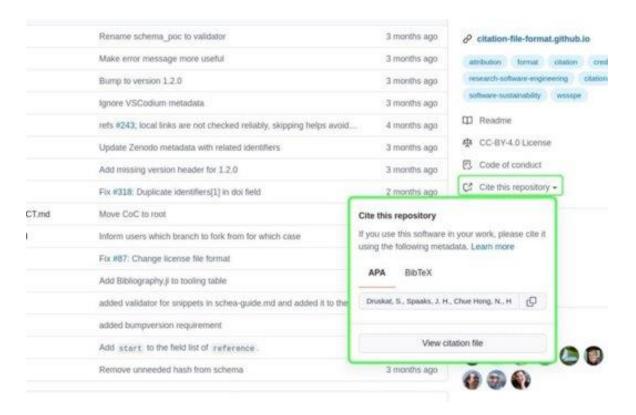


 $\underline{https://blog.esciencecenter.nl/cffinit-now-is-the-time-to-get-more-recognition-for-your-software-e2e6ef617f8e}$



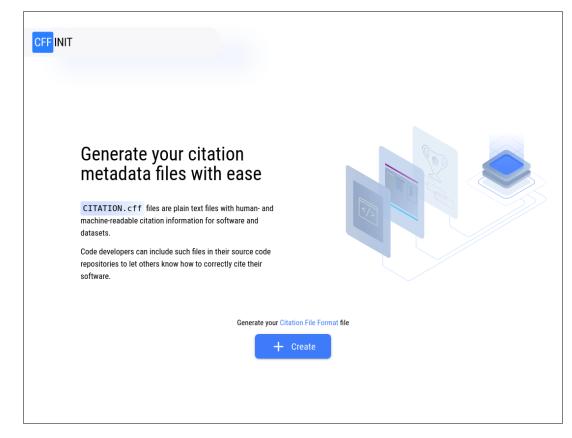




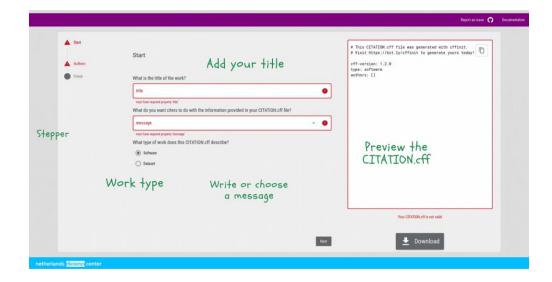


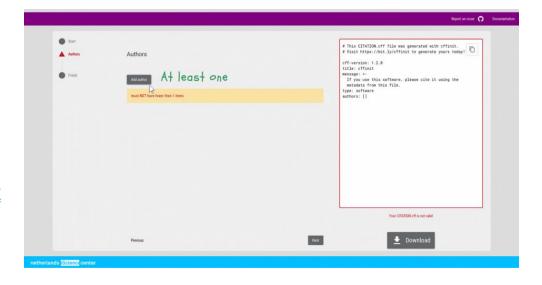
https://citation-file-format.github.io/





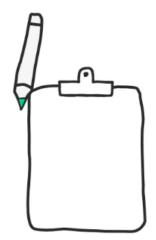
http://bit.ly/cffinit
https://citation-file-format.github.io/cff-initializer-javascript





#5
USE A SOFTWARE
QUALITY
CHECKLIST

F A I R











OpenSSF Best Practices Badge Program

Get Your Badge Now

The Open Source Security Foundation (OpenSSF) Best Practices badge is a way for Free/Libre and Open Source Software (FLOSS) projects to show that they follow best practices. Projects can voluntarily self-certify, at no cost, by using this web application to explain how they follow each best practice. The OpenSSF Best Practices Badge is inspired by the many badges available to projects on GitHub. Consumers of the badge can quickly assess which FLOSS projects are following best practices and as a result are more likely to produce higher-quality secure software.

You can easily see the criteria for the passing badge. More information on the OpenSSF Best Practices Badging program is available on GitHub. Project statistics and criteria statistics are available. The projects page shows participating projects and supports queries (e.g., you can see projects that have a passing badge). You can also see an example (where we try to earn our own badge). This project was formerly known as the Core Infrastructure Initiative (CII) Best Practices badge. and was originally developed under the CII. It is now part of the OpenSSF Best Practices Working Group (WG). The OpenSSF is a foundation of the Linux Foundation (LF).

Privacy and legal issues: Please see our privacy policy, about cookies, and terms of use. The code for the badging application itself is released under the MIT license (projects pursuing a badge are under their respective licenses). All publicly-available non-code content managed by the badging application is released under at least the Creative Commons Attribution License version 3.0 (CC-BY-3.0); newer non-code content is released under CC-BY version 3.0 or later (CC-BY-3.0+). If referencing collectively or not otherwise noted, please credit the OpenSSF Best Practices badge contributors.

Some badge earners:

















Please share this:









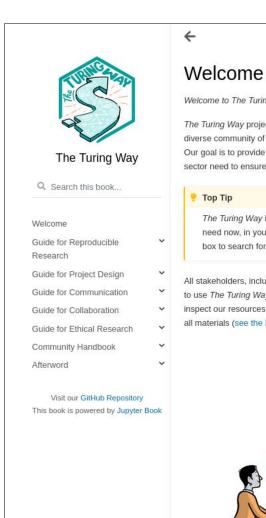


















!≡ Contents

Our Community History

Citing The Turing Way

Welcome to The Turing Way handbook to reproducible, ethical and collaborative data science.

The Turing Way project is open source, open collaboration, and community-driven. We involve and support a diverse community of contributors to make data science accessible, comprehensible and effective for everyone. Our goal is to provide all the information that researchers and data scientists in academia, industry and the public sector need to ensure that the projects they work on are easy to reproduce and reuse.

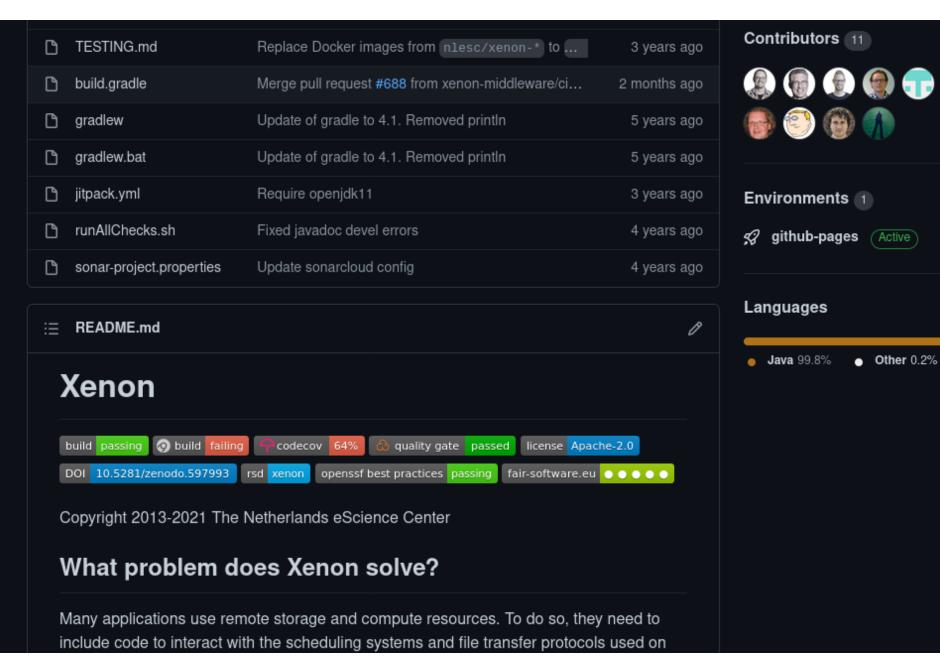
P Top Tip

The Turing Way is not meant to be read from start to finish. Start with a concept, tool or method that you need now, in your current work. Browse the different guides that make up the book, or use the search box to search for whatever you would like to learn about first.

All stakeholders, including researchers, software engineers, project leaders and funding teams, are encouraged to use The Turing Way to understand their roles and responsibility of reproducibility in data science. You can inspect our resources on GitHub, contribute to the project as described in our contribution guidelines and re-use all materials (see the License).

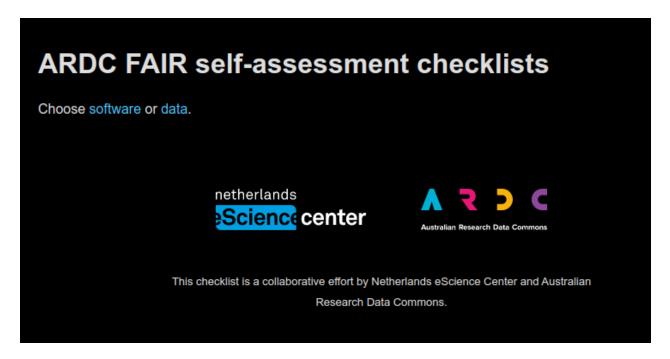


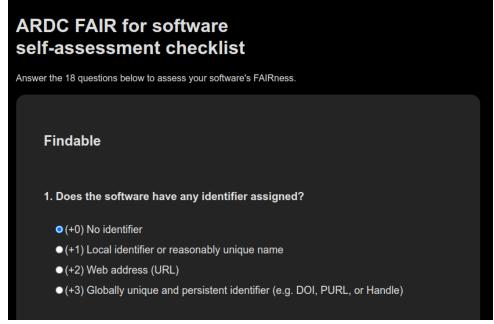






ARDC FAIR-software checklist



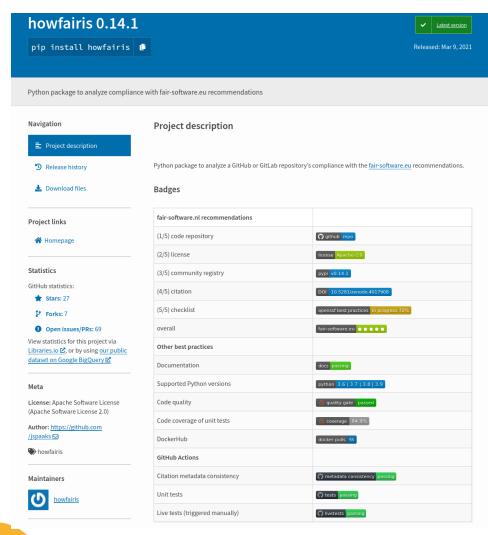




Compliance with 5 recommendations?







https://pypi.org/project/howfairis https://github.com/fair-software/howfairis





Badges

fair-software.nl recommendations	
(1/5) code repository	github repo
(2/5) license	license Apache-2.0
(3/5) community registry	pypi v0.14.2
(4/5) citation	DOI 10.5281/zenodo.4017908
(5/5) checklist	openssf best practices in progress 72%
overall	fair-software.eu • • • •

Install

pip3 install --user howfairis

Verify that the install directory is on the PATH environment variable. If so, you should be able to call the executable, like so:

howfairis https://github.com/<owner>/<repo>

howfairis supports URLs from the following code repository platforms:

- 1. https://github.com
- https://gitlab.com (not including self-hosted instances)

https://doi.org/10.5281/zenodo.7193991



Assess compliance with fair-software.eu

To enable this checker, add the following snippet as <code>.github/workflows/fair-software.yml</code> in your GitHub repository.

```
name: fair-software

on: push

jobs:
    verify:
        name: "fair-software"
        runs-on: ubuntu-latest
        steps:
        - uses: fair-software/howfairis-github-action@0.2.1
        name: Measure compliance with fair-software.eu recommendations
        env:
            PYCHARM_HOSTED: "Trick colorama into displaying colored output"
        with:
            MY_REPO_URL: "https://github.com/${{ github.repository }}"
```

https://github.com/fair-software/howfairis-github-action





Compliance of multiple repositories

fairtally

Python application to analyze multiple GitHub and GitLab repositories compliance recommendations.

fair-software.nl recommendations	Badges
1. Code repository	github repo
2. License	license Apache-2.0
3. Community Registry	pypi v0.1.0 rsd fairtally
4. Enable Citation	DOI 10.5281/zenodo.4590882
5. Checklist	openssf best practices in progress 31%
Other best practices	
Continuous integration	Build passing Linting passing
DockerHub	docker pulls 15

https://pypi.org/project/fairtally/
https://github.com/fair-software/fairtally

- Checks compliance of multiple repositories using howfairis
- Generates a report that includes compliance of each repository
- Can be automated to keep track of the changes



Questions?

netherlands
Science center