Fair Metrics for FAIR Software

Dr. Anna-Lena Lamprecht, Utrecht University
NWO Open Science in Practice Series
The FAIR Principles

Widely known as the “FAIR Data Principles”

Original paper demands that all scholarly digital research objects should be findable, accessible, interoperable and reusable

Increasingly recognized as essential for the transition towards Open Science

6700+ citations counted by Google Scholar

Software is (not) Data

Data (011011101…)

Information
Instructions (Software)
...

Digital Object
Data
Software
...

traditional computer science view
 computational science view
More FAIR Stuff

FAIR code
FAIR workflows
FAIR ML models
FAIR microscopy images
FAIR teaching material
FAIR physical samples
FAIR hardware
FAIR batteries

...
FAIR4RS (FAIR for Research Software)

**Workshop discussions**
- DTL Conference 2018 Utrecht, WOSSS 2019 The Hague, de-RSE 2019 Potsdam, Top 10 FAIR Things Global Sprint 2019, National eScience Symposium 2019 Amsterdam, ...

**2019**
- **Five Recommendations for FAIR Software** ([https://fair-software.eu/](https://fair-software.eu/))

**2020**
- International **FAIR4RS working grouped** formed (by RDA, ReSA & FORCE11), [https://www.rd-alliance.org/groups/fair-research-software-fair4rs-wg](https://www.rd-alliance.org/groups/fair-research-software-fair4rs-wg)

**2021**

**2022**
Key points of FAIR4RS

Software is not (just) data:
● Result of a creative process
● Executable
● Composite nature
● Frequent changes
● Decay

Warrant changes to several of the data-oriented principles.

Example principles:

F1. Software is assigned a globally unique and persistent identifier.
A2. Metadata are accessible, even when the software is no longer available.
I2. Software includes qualified references to other objects.
R1. Software is described with a plurality of accurate and relevant attributes.
R2. Software includes qualified references to other software.
Fair Metrics for FAIR Software

Project No. 203.001.124
NWO Open Science Fund 2020/2021

Team:
- Anna-Lena Lamprecht (PI), Utrecht University
- Michelle Barker, Research Software Alliance
- Jonathan de Bruin, Utrecht University
- Carlos Martinez-Ortiz, Netherlands eScience Center
- Jurriaan Spaaks, Netherlands eScience Center
- Extra RSE
- Student assistants
Problem

The application of the FAIR principles to software lags significantly behind their application to data.

One reason: Current lack of evaluated, community-endorsed metrics to assess software FAIRness and the tooling to support such assessment.

They would enable users to choose the FAIRest option, and in turn incentivize developers to make their software FAIRer.
Vision

We envision actionable metrics and workable tools for assessing the FAIRness of software, which maximize the utility for research software developers and users, and incentivize cultural change in line with the ideas and ideals of Open Science.
Aims

Towards this vision, this project aims at evaluating suggested FAIR software metrics with a focus on how to incentivize cultural change.

Four main objectives:
1. Identify candidate metrics based on recent discussions in the community.
2. Prototypically implement tooling for applying the candidate metrics.
3. Evaluate the candidate metrics on software projects from different research areas.
4. Aggregate the findings into recommendations for FAIR software metrics and tooling.
Candidate Metrics

- The Netherlands eScience Center’s “Five Recommendations for FAIR software” ([https://fair-software.nl/](https://fair-software.nl/)) and the howfairis tool ([https://github.com/fair-software/howfairis](https://github.com/fair-software/howfairis))
- The Barcelona Supercomputing Center’s “FAIRsoft approach”: [https://doi.org/10.1101/2022.05.04.490563](https://doi.org/10.1101/2022.05.04.490563)
- Australian Research Data Commons (ARDC) self assessment tool
- FAIR Software Assessment/Badging initiative (GO FAIR US)
- GO BUILD (FAIR assessment framework)

Anything missing? Please let us know!
 Five Recommendations for FAIR Software (NLeSC, 2019):

1. Use a publicly accessible repository with version control.
2. Add a licence.
3. Register your code in a community registry.
4. Enable citation of the software.
5. Use a software quality checklist.
howfairis

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

Tool to analyze a GitHub/GitLab repository's compliance with the fair-software.nl recommendations (NLeSC, 2020)

Binary assessment of the five recommendations

Gives score (0-5) and badge
howfairis in SWORDS@UU

Scan and Review of Open Research Data and Software (SWORDS) @ UU (https://github.com/UtrechtUniversity/SWORDS-UU)

A project to get stats and figures about how Utrecht University researchers develop and manage software

Summer 2021: first results for 143 users and 1356 repositories associated with Utrecht University, incl. the howfairis score
Eva Martín del Pico, Josep Lluís Gelpí, Salvador Capella-Gutiérrez: FAIRsoft - A practical implementation of FAIR principles for research software (bioRxiv preprint, https://doi.org/10.1101/2022.05.04.490563)

Key ideas:
● interpretation of the FAIR4RS principles as measurable indicators
● scoring system to weight the different indicators, to offer a quantitative assessment of software FAIRness
FAIRsoft: High-Level and Low-Level Indicators

From bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490563
FAIRsoft: High-Level and Low-Level Indicators

From bioRxiv preprint doi: https://doi.org/10.1101/2022.05.04.490563
FAIRsoft: Scoring System

Low-level indicators are assigned weights according to their relative relevance for the respective high-level indicators.

High-level indicators are assigned weights according to their relative relevance for the respective FAIR principle.

Results in a score between 0 and 1 for each FAIR principle.

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Name</th>
<th>LL weight</th>
<th>HL weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>Identity uniqueness</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F1.1 Uniqueness of name</td>
<td>0.8</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>F1.2 Identifiability of version</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>F2</td>
<td>Existence of Metadata</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F2.1 Structured Metadata</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F2.2 Standardized Metadata</td>
<td>0.4</td>
<td>0.2</td>
</tr>
<tr>
<td>F3</td>
<td>Searchability</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F3.1 Searchability in registries</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F3.2 Searchability in software repositories</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>F3.3 Searchability in literature</td>
<td>0.4</td>
<td></td>
</tr>
</tbody>
</table>

Example adapted from https://docs.google.com/document/d/1OcGxtX-LqloPSeEE-S1v3kgKXe6wiCgo_VMWF6idhM/edit#
FAIRsoft: Application to Life Science Tools

FAIRsoft indicators computed for 43,987 tool instances in the OpenEBench platform

Individual scores summarised to depict the FAIRness level of the evaluated collection
Next in the Project

Finish review of suggested FAIR software metrics

Implement tooling for automatically accessible metrics

Apply to 100+ open source research software projects

Evaluation of the effects on their developers

Formulate recommendations for FAIR software metrics

Please get in touch if you want to make sure yours is included!
Evaluation: Focus on Incentivization Potential

Goal are metrics that foster cultural change by incentivizing research software developers to make their software FAIR(er)

Points to consider:
- What are incentives for research software developers?
- Quantitative or qualitative FAIRness assessment?
- Score, badge or other form of summary?
- What is an appropriate level of detail?
Summary

FAIR essential for the transition towards Open Science

Software (and other kinds of research objects) need specialized FAIR principles

Various metrics for software FAIRness suggested

Project to evaluate metrics with focus on incentivization potential

Vision are metrics and tools that help fostering cultural change
Thank you!