

FAIR, FLOSS and Sustainable: Complementary Paradigms for Research Software

Christian E. Busse^[1,2,3,🐙]

[1] Div. of B Cell Immunology, German Cancer
Research Center, Heidelberg, DE

[2] ORCID: 0000-0001-7553-905X

[3] Github: bussec

[🐙] Mastodon: @bussec@scholar.social



My "hat" collection / disclosures



Project lead for
federated RDM of
immunological data



Co-speaker for
NFDI4RSE cross-
cutting initiative



Coordinator for
Open Science

Talk outline

- *Code* as Digital Object in Research
- Paradigms for *Code*
 - FAIR Software
 - Free/Libre/Open Source Software (FLOSS)
 - Sustainable Software
- Synthesis
- Outlook

What is "Code"

- *Code* is any set of machine-executable instructions that processes *data* in the course of a research project
- It therefore inherently documents the research process and allows (theoretically) for simple and direct replication of results
- Depending on its definition, "software" is either equal to or a subset of *Code*

Code as Digital Object

Code is distinct from other digital objects in several ways:

- (1) *Code* requires maintenance in order to ensure its continued usability in changing hard- and software environments
- (2) *Code* evolves over time, due to necessary adaptations and continuous development
- (3) *Code* receives external contributions after the initial publication, thus acknowledgement, accountability and IP/licensing issue can arise
- (4) *Code* has different demands to RDM infrastructure (less space, fine-grained versioning, no universally accepted PID system)
- (5) *Code* depends on libraries, which are also *code* objects (i.e., also change over time)
- (6) *Code* can be conveyed as source or as binaries, which is critical for re-use/licensing

Code is an object class of its own

*Treating code as "data"
is like
treating an animal as "sausage"*

(it is a transformation that you can perform, but you must be aware of the implications)

- Important to recognize that *code* is an own class of digital objects
- If a *code* object is treated as *data*, only the common properties remain

FAIR Software

- Can the original FAIR paradigm^[1] be applied to *code*?
- Should we:
 - change FAIR criteria,
 - extend them (more "letters") or
 - consider them as "base layer"?
- "Towards FAIR principles for research software"^[2]
 - ➔ Re-interpret FAIR criteria for *code* objects
- As usual, "FAIR" is not necessarily "Open"

[1] Wilkinson MD *et al.*, Sci Data 3:160018 (2016)
DOI:[10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18)

[2] Lamprecht AL *et al.*, Data Science 3:37 (2019)
DOI:[10.3233/DS-190026](https://doi.org/10.3233/DS-190026)

Free/Libre/Open Source Software (FLOSS)

- *Free Software*^[1] grants Four Freedoms to the user:
 - Use
 - Study
 - Share
 - Adapt
- *Open Source*^[2] uses different criteria, but functionally the definitions are nearly identical
- Well understood paradigm in software development, used for decades
- Focus is typically on licensing (i.e. "the R in FAIR")
- ➔ Clarity and compatibility: REUSE project^[3]

[1] FSF – Free Software Definition
[2] OSI – Open Source Definition
[3] REUSE Project

Sustainable Software

- *Code* objects have a dependency graph that must be maintained to ensure continued functionality
- Software Heritage^[1]
 - ➔ Conserve the current version of *code* and its dependencies
- Maintenance requires use of Good Development Practices
 - ➔ Development is a *process*, but FAIR and FLOSS only describe a *state*
- Maintenance requires limited resources (funds, work force)^[2]
 - ➔ Selection/decision processes on level of scientific domain or research institution
- Recognition of work, how to cite, what to cite^[3]
- Sustainable use of funding
 - ➔ Open Data Directive^[4]
 - ➔ Public Money, Public Code^[5]

[1] [Software Heritage](#)

[2] Anzt H *et al.* F1000 Res (2020)
DOI:[10.12688/f1000research.23224.1](https://doi.org/10.12688/f1000research.23224.1)

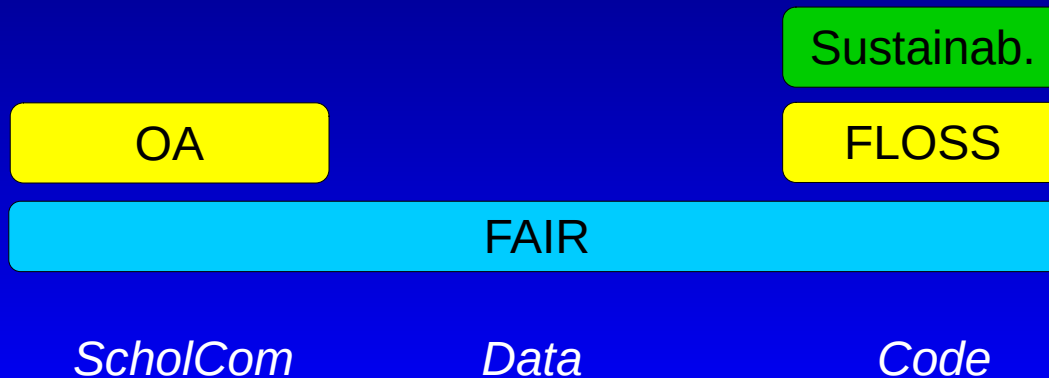
[3] DOI:[10.1109/MCSE.2019.2963148](https://doi.org/10.1109/MCSE.2019.2963148)

[4] [EU Directive 2019/1024](#)

[5] [PMPC](#)

Synthesis

- The paradigms are complementary and can be considered as layers
- FAIR is a necessary requirement for all digital objects in Open Science
- *Code* objects should fulfill additional criteria, i.e., FLOSS and Sustainable
- We are already using this principle for scientific publications (FAIR + OA)



Outlook / Recent developments

- Joint effort^[1] of RDA/ReSA/Force11 to address FAIR for Research Software
- "Software" is now increasingly recognized as first-class research output by policy makers^[2,3]
- "An environment for sustainable research software in Germany and beyond"^[4]
- NFDI4RSE Cross-cutting initiative^[5]

[1] [RDA - FAIR4RS WG](#)

[2] [EOSC Strategic Implementation Plan](#)
DOI:10.2777/202370

[3] [Draft Proposal EOSC Partnership](#)

[4] [Anzt H et al. F1000 Res \(2020\)](#)
DOI:10.12688/f1000research.23224.1

[5] [NFDI4RSE](#)