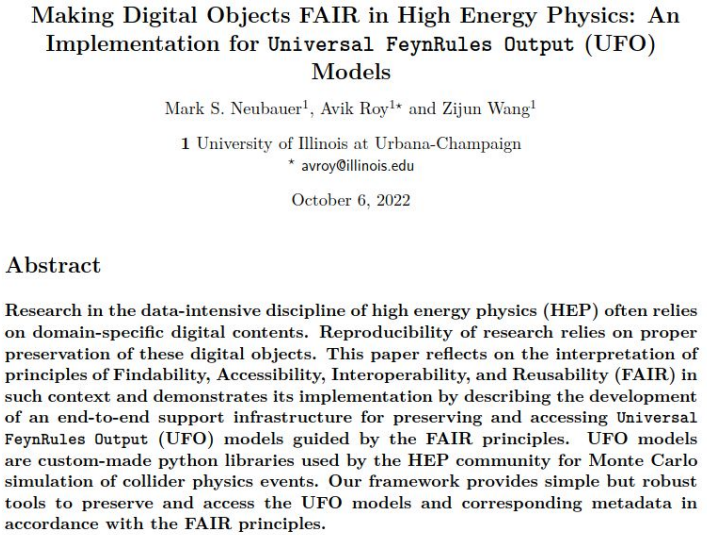
FAIR Principles for Better Citation: A Case Study with UFO Models

Avik Roy
University of Illinois at Urbana-Champaign

Nov 23, 2022



Making Digital Objects FAIR in High Energy Physics: An Implementation for `Universal FeynRules Output` (UFO) Models

Mark S. Neubauer[1], Avik Roy[1*] and Zijun Wang[1]

1 University of Illinois at Urbana-Champaign
* avroy@illinois.edu

October 6, 2022

Abstract

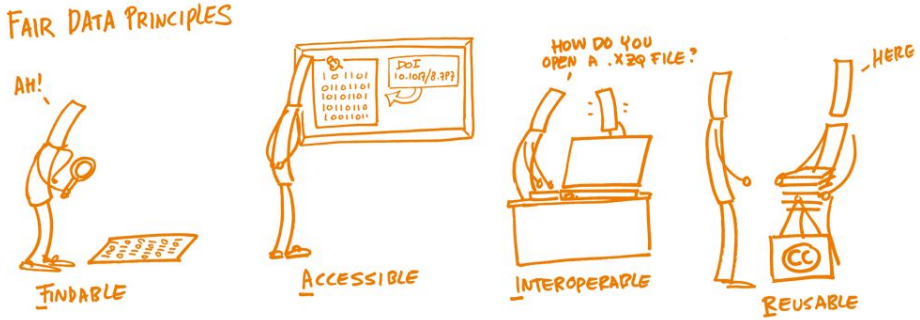
Research in the data-intensive discipline of high energy physics (HEP) often relies on domain-specific digital contents. Reproducibility of research relies on proper preservation of these digital objects. This paper reflects on the interpretation of principles of Findability, Accessibility, Interoperability, and Reusability (FAIR) in such context and demonstrates its implementation by describing the development of an end-to-end support infrastructure for preserving and accessing `Universal FeynRules Output` (UFO) models guided by the FAIR principles. UFO models are custom-made python libraries used by the HEP community for Monte Carlo simulation of collider physics events. Our framework provides simple but robust tools to preserve and access the UFO models and corresponding metadata in accordance with the FAIR principles.

Results from arxiv: 2209.09752
Git repo: https://github.com/Neubauer-Group/UFOManager
https://github.com/Neubauer-Group/UFOMetadata/

The FAIR Principles

- To inspire scientific data management for reproducibility and maximal reusability[1]
- Originally proposed for scientific data
- Can be interpreted as guidelines to manage and preserve other Digital Objects (DOs) e.g. research software[2] , tutorials and notebooks[3] , AI and ML models[4]
- Different working groups working on FAIR guidelines for different DOs (e.g FAIR4RS, FAIR workflows, FAIR VREs)

Findable:	locating DOs in a failsafe fashion
Accessible:	obtaining DOs along with their context, content, and format
Interoperable:	being usable across multiple computing platforms
Reusable:	specifying the context and extent of reusing DOs



Universal FeynRules Output (UFO) Models

- UFO models are used for simulating Beyond Standard Model Physics with Monte Carlo generators
- Custom Python libraries that pack necessary physics content as modules
- Designed to be **Interoperable** across multiple generator
- Heavily used in ATLAS and CMS analyses
- No uniform convention for management and citation practices for these models

Model-independent files	Model-dependent files
`__init__.py`	`particles.py`
`object_library.py`	`coupling_orders.py`
`function_library.py`	`parameters.py`
`write_param_card.py`	`vertices.py`
	`couplings.py`
	`lorentz.py`

Universal FeynRules Output (UFO) Models

- UFO models are used for simulating Beyond Standard Model Physics with Monte Carlo generators
- Custom Python libraries that pack necessary physics content as modules
- Designed to be **Interoperable** across multiple generator
- Heavily used in ATLAS and CMS analyses
- No uniform convention for management and citation practices for these models

Model-independent files	Model-dependent files
`__init__.py`	`particles.py`
`object_library.py`	`coupling_orders.py`
`function_library.py`	`parameters.py`
`write_param_card.py`	`vertices.py`
	`couplings.py`
	`lorentz.py`

Citing UFO models

- In many cases, only the theory paper describing the physics model is cited
- The actual digital object i.e. the implementation of the model itself is not always cited
- When cited, no uniform convention is followed
- Often digital repositories are cited, that may be lost/changed when the hosting service or the hosting account becomes inactive

[220] DM forum repository, *Higgs_scalar UFO model webpage*, https://svnweb.cern.ch/cern/wsvn/LHCDMF/trunk/models/Higgs_scalar_UFO/.

[221] DM forum repository, *Zp2HDM_UFO UFO model webpage*, https://svnweb.cern.ch/cern/wsvn/LHCDMF/trunk/models/EW_Higgs_2HDM/.

[222] DM forum repository, *DMS_tloop UFO model webpage*, https://svnweb.cern.ch/cern/wsvn/LHCDMF/trunk/models/Monojet_DMS_tloop/.

[223] DM forum repository, *DMScalarMed_loop UFO model webpage*, https://svnweb.cern.ch/cern/wsvn/LHCDMF/trunk/models/HF_S+PS/.

[224] DM forum repository, *dmS_T UFO model webpage*, https://svnweb.cern.ch/cern/wsvn/LHCDMF/trunk/models/Monojet_tChannel/contributed_by_Amelia_Brennan/.

From the references in JHEP 05 (2019) 142

Making UFOs FAIR

- Making UFOs FAIR requires consistent standards for
  - preserving the models with persistent identifiers like DOIs (F, A)
  - storing domain-specific enriched metadata (I, R)
  - allowing search and download models (F, R)
- FAIRification should be independent of developing the UFO, compatible with existing models

UFOManager

Dedicated software developed to
- Validate models
- Create enriched metadata
- Store said metadata in dedicated repo
- Publish model in Zenodo with DOI
- Facilitate search and download of models
- Allows version controlling

UFOMetadata

Dedicated repository developed to
- Store metadata of published UFO models
- Validate metadata format with continuous integration

FAIR UFOs: Standardized Citation Practices

- Making UFOs FAIR can allow better citation practices
- Using DOIs allow persistent citations
- Allows the developers get the right credit for developing the software (along with the associated theory work)
- These DOIs can be used to look for these models and download their specific versions with UFOManager
- Associated metadata can be easily retrieved from the UFOMetadata

[46] L. Panizzi and B. Fuks, VLQ UFO model at NLO QCD with 5FS, doi:10.5281/zenodo.6991118, Available at: https://doi.org/10.5281/zenodo.6991118 (2022).

[47] L. Panizzi, Vector-like quark UFO model at NLO QCD with five flavour scheme, doi:10.5281/zenodo.7038823, Available at: https://doi.org/10.5281/zenodo.7038823 (2022).

[48] L. Panizzi, Vector-like quark UFO model at NLO QCD with five flavour scheme (third generation), doi:10.5281/zenodo.7041528, Available at: https://doi.org/10.5281/zenodo.7041528 (2022).

From the references in [arxiv: 2209.09752]

References

1. Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18
2. Chue Hong, Neil P., Katz, Daniel S., Barker, Michelle et al. RDA FAIR4RS WG. (2022). FAIR Principles for Research Software (FAIR4RS Principles) (1.0). https://doi.org/10.15497/RDA00068
3. Richardson, R. A., et al. "User-friendly Composition of FAIR Workflows in a Notebook Environment." Proceedings of the 11th on Knowledge Capture Conference. 2021. https://doi.org/10.1145/3460210.3493546
4. Katz, D. S., Psomopoulos, F. E., and Castro, L. J. "Working towards understanding the role of FAIR for machine learning." DaMaLOS@ ISWC (2021): 1-7. https://doi.org/10.4126/FRL01-006429415