3rd FAIR Digital Objects Forum Conference

TIBconferencesession

TIB-OP_will_set_DOI

Published: 1970-01-01

The FAIR Variable: Turning Observed Properties into FAIR Digital Objects

Barbara Magagna[12*], Arvin Rastegar[3], Sirko Schindler[4], Saurav Kumar[5], Anne Fouilloux[62], and Esteban Gonzalez[7]

[1]Semantics, Cybersecurity and Services, University of Twente, Enschede, Netherlands

[2]GO FAIR Foundation, Leiden, Netherlands

[3]Karlsruhe Institute of Technology, Garmisch-Partenkirchen, Germany

[4]German Aerospace Center (DLR), Institute of Data Science, Jena, Germany

[5]Federal Institute for Risk Assessment, Berlin, Germany

[6]Lifewatch, Seville, Spain

[7]Universidad Politécnica de Madrid, Madrid, Spain

*Correspondence: Barbara Magagna, barbara@gofair.foundation

Abstract.

Keywords: FAIR Digital Object, FAIR Variable, Interoperability

1. First section

Scientific variables are the building blocks of data from simulations, experiments, and other observations, which capture the property observed and in which context the measurement was obtained. Ideally, variables provide all the necessary information to understand the meaning of the corresponding result values. This meaning should be accessible and carried throughout all phases of the data lifecycle. Thus, it should be used at both the data and metadata levels. At metadata level, it is required when we want to identify specific types of data, e.g., in online repositories. At the data level, it is necessary to understand the actual meaning of the values and interpret them accordingly.

In parallel, FAIR Digital Objects (FDOs) are defined as self-contained digital artefacts comprising all essential information required for their interpretation and use by both humans and machines [1]. One of their important features is encapsulation, i.e. they can be composed of other FDOs, for example, to disclose their semantic nature. Properly represented variables can serve as FDOs within a dataset FDOs, conveying all information necessary for their accurate interpretation and demonstrating FDO's encapsulation principle. This work provides a framework for such variable FDOs relying on I-ADOPT to enable interoperability, and nanopublications to represent them for FDO-compliance.

Although variable descriptions should be findable and interoperable, it is still challenging to find coherent descriptions of variables in dataset metadata within repositories and even more difficult across different repositories. It is widely recognized that controlled variable lists based on Linked Data are part of the solution to this problem. However, the real challenge is to find a common way to describe variables consistently. The RDA I-ADOPT Working Group has made it its goal and mission to address this challenge. It developed the I-ADOPT Framework [2] which includes an ontology and recommendations subsequently endorsed by the RDA. Its purpose is to provide interoperable and machine-interpretable variable descriptions by decomposing them into atomic elements annotated with concepts from semantic artefacts. Its adoption is already growing in Earth and environmental science vocabularies such as the THEIA-OZCAR thesaurus [3] hosted in EarthPortal, the Environmental thesaurus [4] hosted in EcoPortal, or subsets of the BODC PUV and CF Standard Names in the NERC Vocabulary Server [5]. While the variables are represented as skos:Concept, their description components (iop:Property, iop:Entities playing description roles ObjectOfInterest, Matrix or ContextObject) are integrated in these vocabularies via the object properties of the I-ADOPT ontology, sometimes using terms both within and beyond the initial concept schema.

However, semantic concepts for variable constraints cannot be included directly in these representations because binding the target of the constraint to a constraint concept prevents reuse of that constraint in other variables. Constraints should therefore be modeled as separate nodes (named or blank) for each Constraint and Variable. These nodes can then be defined as instances of a semantic concept, allowing the latter to be reused in other variable descriptions. These nodes are very naturally embeddable in nanopublications which are based on the named graph concept. At the same time, they are suitable representations of FDOs as discussed in [6].

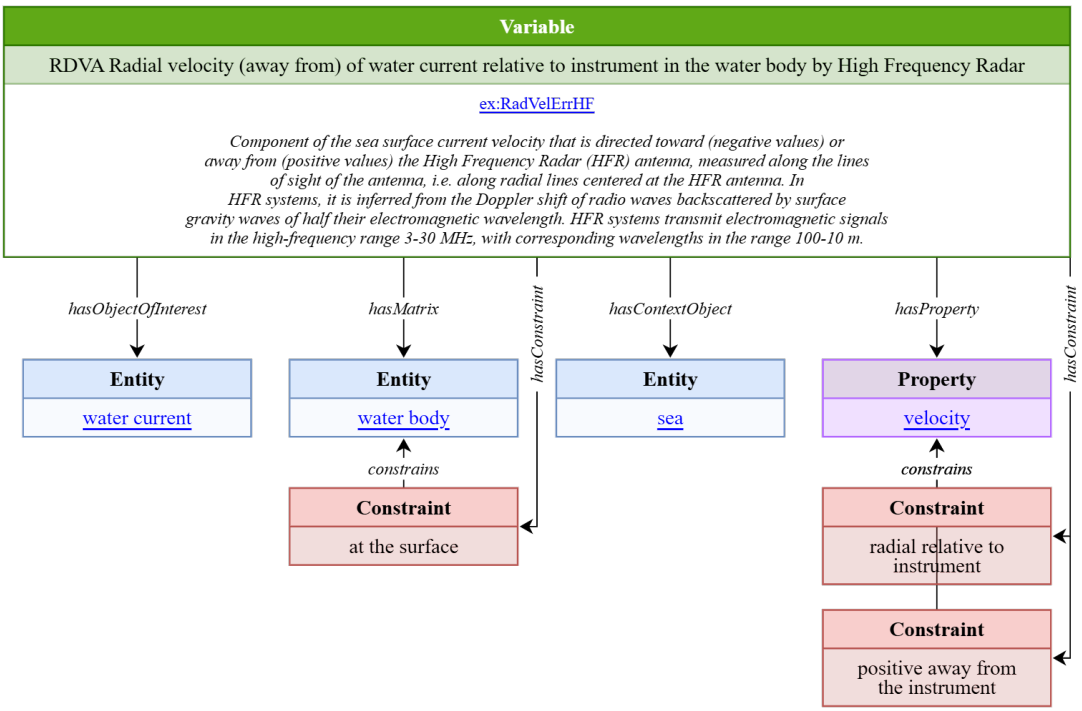

Figure 1. Typology of FAIR Supporting Resources (FSRs).

The paper describes how the I-ADOPT ontology is used as a nanopublication template for creating new I-ADOPT variables together with their underlying atomic compo-

nents. This approach offers two main key advantages that implement core FDO principles (compare with [1]: (a) Variables can be defined in alignment with the FDOs specification (FDO abstraction and encapsulation), (b) Persistent identifiers (Trusty URI) can be assigned for proposed variables and their components via the nanopub network infrastructure (FDO binding with PIDs). Furthermore, Including SHACL shape for the I-ADOPT variable definitions in the nanopublication template and having each variable instance declare “dcterms:conformsTo” the FDO profile, the template enables validation against the profile’s mandatory and optional attributes. These features will effectively complement the combined efforts of two projects that are currently developing an automated I-ADOPT service. The German NFDI4Earth Incubator Lab project aims to operationalize FDO creation by integrating large language models with the I-ADOPT ontology, enabling scalable and semantically consistent variable extraction, visualization, and publication workflows across environmental and Earth sciences. The EOSC FAIR2Adapt project [7] collaborates in this development, aiming to use this service for the EU climate adaptation community. The LLM-based service produces JSON-LD conforming to the I-ADOPT ontology, which is then automatically converted into RDF/Turtle (TTL) representations by a transformation script. Through a GitLab CI/CD pipeline, these TTL files are validated and published as nanopublications to the network. This automated workflow lowers technical barriers, allowing researchers from diverse backgrounds to engage with FAIR variable creation enabling them to validate the generated results. The I-ADOPT service could be embedded in any digital object registry enabling uploaded datasets to be described with FDO I-ADOPT variables to increase their findability and interoperability as part of their metadata, and overall also helping to make them more understandable and reusable by humans and machines. The enriched metadata can be utilized to enable semantic faceted search over the digital objects hosted in a repository. The FAIR2Adapt project will make use of this novel approach in ROHub [8], a repository of FDOs represented as RO-Crates, which is used to store research objects for climate adaptation solutions.

Author contributions

All authors:Writing - Review and Editing **Barbara Magagna**: Conceptualization, Writing - Original Draft

Competing interests

The authors declare that they have no competing interests.

Funding

This work is funded through the NFDI4Earth Program (NFDI4Earth Incubator Lab project LLM-enabled I-ADOPT Variable Extraction using Semantics) and the European Union’s Horizon Europe framework program under grant agreement No. 101188256 (FAIR2Adapt).

Acknowledgements

The authors acknowledge the RDA WG I-ADOPT core team for the I-ADOPT Framework development and Lorenzo Corgnati for providing a variable from the EuroGOOS

High Frequency Radar Task Team used as an I-ADOPT decomposition example in this paper.

References

[1] I. Anders et al., “FAIR digital object technical overview,” Apr. 13, 2023, Publisher: Zenodo Version Number: PEN 2.0. DOI: 10.5281/ZENODO.7824714. Accessed: Nov. 2, 2025. [Online]. Available: https://zenodo.org/record/7824714.

[2] B. Magagna et al., “InteroperAble descriptions of observable property terminologies (i-ADOPT) WG outputs and recommendations,” 2022, Publisher: Research Data Alliance Version Number: 1. DOI: 10.15497/RDA00071. Accessed: Nov. 2, 2025. [Online]. Available: https://zenodo.org/record/6520132#.YnOhMJNBybQ.

[3] C. Coussot, V. Chaffard, I. Braud, and S. Galle, *Theia/ozcar thesaurus*, 2025. Accessed: Oct. 31, 2025. [Online]. Available: https://w3id.org/ozcar-theia/.

[4] B. Magagna, J. Peterseil, M. Frenzl, M. Kertész, and U. Grandin, *Envthes – thesaurus for long term ecological research, monitoring and experiments*, 2025. Accessed: Oct. 31, 2025. [Online]. Available: https://vocabs.lter-europe.net/EnvThes.

[5] G. Moncoiffe and A. Kokkinaki, *The nerc vocabulary server (nvs)*, 2025. Accessed: Oct. 31, 2025. [Online]. Available: https://vocab.nerc.ac.uk/.

[6] B. Magagna, L. O. Bonino Da Silva Santos, T. Kuhn, L. Ferreira Pires, and E. Schultes, “Nanopublications as FAIR digital object implementations,” *Open Conference Proceedings*, vol. 5, Mar. 18, 2025, ISSN: 2749-5841. DOI: 10.52825/ocp.v5i.1417. Accessed: Nov. 2, 2025. [Online]. Available: https://www.tib-op.org/ojs/index.php/ocp/article/view/1417.

[7] A. Fouilloux and L. Bolands Eklund, “D1.1 - FAIR2adapt project handbook,” Apr. 3, 2025, Publisher: Zenodo Version Number: v1.0 - DRAFT NOT YET APPROVED BY THE EUROPEAN COMMISSION. DOI: 10.5281/ZENODO.15130762. Accessed: Nov. 2, 2025. [Online]. Available: https://zenodo.org/doi/10.5281/zenodo.15130762.

[8] R. Palma, J. Gòmez-Pèrez, and M. Wolniewicz, *Rohub*, 2025. Accessed: Oct. 31, 2025. [Online]. Available: https://www.rohub.org/.