Placeholder Name of the conference

Placeholder Session title

https://doi.org/10.52825/xxxx..... DOI placeholder (WILL BE FILLED IN BY TIB Open Publishing)

© Authors. This work is licensed under a Creative Commons Attribution 4.0 International License

Published: (WILL BE FILLED IN BY TIB Open Publishing)

Aligning and globally indexing diverse FDOs with nanopublications

Virginia Balseiro[1,*] https://orcid.org/0009-0009-0118-9195, Ziroli Plutschow[1] https://orcid.org/0009-0008-3635-347X, Ashley Caselli[1] https://orcid.org/0000-0001-8492-0354, and Tobias Kuhn[1] https://orcid.org/0000-0002-1267-0234

[1]Knowledge Pixels, Switzerland

*virginia@knowledgepixels.com

FAIR Digital Objects (FDOs) [1] and nanopublications [2] are emerging paradigms for structured, interoperable data. While conceptually overlapping, they differ in their focus, approach, and implementation [9]. We show how they can be effectively combined to align FDOs of different kinds and to provide global indexing.

FDOs aim to align digital content, such as images, datasets, software, and tables, with the FAIR principles by making them Findable, Accessible, Interoperable, and Reusable. They do that by specifying [3] how a persistent identifier maps to a digital artifact and its structured metadata for semantic interpretation and reuse. Each metadata record refers to an FDO profile, which defines its structure and semantics through attribute definitions, and both profiles and attributes are themselves FDOs.

Different FDO implementations have emerged, but their lack of full compatibility currently hampers adherence to the FAIR principles. FDOs have been implemented based on the Handle system [4], via the JSON-based RO-Crates approach [5], as well as grounded in RDF/OWL technology with the FAIR Digital Object Framework (FDOF) [6]. Apart from their lack of alignment, these current approaches also do not provide reliable methods to represent version histories, and they do not come with an infrastructure to globally search FDOs.

Nanopublications can be seen as a fourth strand of related work. They are small independent packages of semantic data, encoded in the RDF model. Nanopublications can contain any kind of structured knowledge, including FDO metadata, as their content (“assertion” part), together with relevant metadata about this assertion (“provenance” and “publication info” parts). Nanopublications can be used to implement FDOs [9], and we show here how they can unify existing approaches, improving compatibility and interoperability. They also provide built-in versioning and a decentralized network for global semantic querying, addressing key limitations of current FDOs.

Handle-based FDOs [4] implement FDOs with the metadata records of the Handle System. In our approach, we represent such FDOs as nanopublications by mapping each Handle record’s key-value pairs to RDF assertion triples. In this model, every valid Handle record corresponds to a nanopublication in which the FDO’s persistent identifier (PID) becomes the subject, the record key becomes the predicate, and the record value becomes the object. The nanopublication preserves the original semantics, mapping Handle identifiers deterministically to URIs. Additionally, we map the attribute that links an FDO to its profile onto the Dublin Core dct:conformsTo predicate that links the FDO to its profile definition, and express the

corresponding FDO Profile definitions as SHACL [7] shape constraints, which can be used to validate that an FDO instance conforms to its declared profile. This approach allows existing Handle-based repositories to publish their digital objects additionally as nanopublications without modifying the underlying Handle infrastructure.

RO-Crates [5,8] provide a standardized way to structure and package research data, including tables, publications, and workflows, all with well-structured metadata. Their metadata, stored as a single JSON-LD file, can be interpreted as RDF triples, which we place directly into nanopublication assertions. RO-Crates use the same dct:conformsTo predicate to link to RO-Crate Profiles, naturally mapping to FDO Profiles and providing a key point of alignment across FDO types.

The FAIR Digital Object Framework (FDOF) [6] provides a general semantic model for representing digital resources as FAIR Digital Objects, defining the relations between the identifier, metadata records, and the digital object itself. With our approach, we use nanopublications as a vessel to instantiate this model with specific FDO definitions by placing the respective triples in the assertion part. We further augment these definitions with FDO Profile references using the same dct:conformsTo predicate.

Our initial implementation efforts demonstrate the feasibility of generating and publishing nanopublications from existing Handle-based FDOs as well as RO-Crates. Crucially, these mappings preserve all existing structures and identifiers, providing an additional publishing channel for existing FDO metadata rather than creating new ones. In addition to this pathway, our tools can also mint new FDOs from scratch by publishing their metadata records as nanopublications, assigning persistent identifiers at creation time and ensuring full alignment with FDO profiles.

Specifically, we can publish all these FDOs in the existing decentralized nanopublication network to a Nanopub Registry [10] instance, and query them via Nanopub Query [11]. Overall, nanopublications thereby enable decentralized indexing, explicit provenance and versioning, and uniform representation across diverse FDO types for automated reasoning and integration. This approach supports a robust, interoperable ecosystem where updates, citations, and metadata can be tracked transparently, promoting reproducibility and reuse, and fostering a more open, verifiable, and interconnected digital object landscape.

Author contributions

All authors have contributed to the conceptualization, methodology, and writing.

Competing interests

The authors have no competing interests.

References

[1] FAIR Digital Object Forum, Basic FDO Definition, https://fairdo.org/basic-fdo-definition/

[2] Tobias Kuhn et al., Nanopublications: A Growing Resource of Provenance-Centric Scientific Linked Data, Semantic Web, 2018, https://doi.org/10.1109/eScience.2018.00024

[3] FAIR Digital Object Forum, FAIR Digital Object (FDO) Specification, https://fairdo.org/specification/

[4] Sam Sun, Larry Lannom, and Brian Boesch. Handle system overview. Network Working Group, No. rfc3650. 2003. https://www.rfc-editor.org/rfc/rfc3650

[5] RO-Crate Community, Research Object Crate (RO-Crate), https://www.researchobject.org/ro-crate/

[6] Luiz Olavo Bonino da Silva Santos et al. Towards a conceptual model for the FAIR Digital Object Framework. Formal Ontology in Information Systems. IOS Press, 2023. https://doi.org/10.3233/FAIA231131

[7] W3C, Shapes Constraint Language (SHACL), https://www.w3.org/TR/shacl/

[8] Stian Soiland-Reyes et al. Creating lightweight FAIR digital objects with RO-Crate. Research Ideas and Outcomes 8 (2022): e93937. https://doi.org/10.3897/rio.8.e93937

[9] Barbara Magagna et al. Nanopublications as FAIR Digital Object Implementations. Open Conference Proceedings. Vol. 5. 2024. https://doi.org/10.52825/ocp.v5i.1417

[10] Nanopub Registry. https://github.com/knowledgepixels/nanopub-registry

[11] Nanopub Query. https://github.com/knowledgepixels/nanopub-query