

Advancing FAIR Digital Objects in RADAR: RO-Crate Integration, Machine-Actionable Access, and AI-Supported Metadata Annotation and FAIRness Assessment

Jan Schweikert^{1,*}  <https://orcid.org/0000-0003-4774-2717>, Stefan Hofmann¹  <https://orcid.org/0000-0003-0790-112X>,
Sergio Pollo Vázquez¹  <https://orcid.org/0009-0008-8702-0341>, Kerstin Soltau¹  <https://orcid.org/0000-0002-6368-1929>, and Felix Bach¹  <https://orcid.org/0000-0002-5035-7978>

¹ FIZ Karlsruhe, Germany

*Correspondence: Jan Schweikert, jan.schweikert@fiz-karlsruhe.de

Abstract. RADAR [1] is a cross-disciplinary, FAIR-aligned repository service operated by FIZ Karlsruhe – Leibniz Institute for Information Infrastructure [2]. It provides long-term, format-independent archiving and publication of digital research data across all scientific domains. While datasets published in RADAR have traditionally been accessible only via the WebUI and as complete TAR archives [3], this has limited machine-readability and automated reuse.

To address this, RADAR implemented FAIR signposting [4] in recent years and is now extending its metadata framework and access capabilities to make dataset contents both machine-readable and directly accessible, thereby strengthening compliance with the FAIR principles and enhancing interoperability within the research data ecosystem.

In the next development phase, RADAR introduces support for the RO-Crate [5] metadata standard as specified by researchobject.org [6]. For every published dataset, RADAR will automatically generate an accompanying `ro-crate-metadata.json` document available alongside the existing TAR-based dataset download. This structured file describes the internal dataset contents – including files, directories, and associated metadata such as size, format, and timestamps – while embedding all mandatory and optional dataset-level metadata (e.g., title, license, author information, and more).

In parallel, RADAR is implementing granular file-level access through streaming from TAR archives. This functionality enables users and machines to retrieve individual files without downloading the full dataset. To ensure precise and persistent referencing, RADAR extends its DOI-based [7] access scheme by introducing fragment identifiers that uniquely denote specific files within an already published TAR archive. When a DOI is resolved together with its fragment identifier, RADAR can directly stream the referenced file from the archived dataset. This mechanism realizes direct per-file access via DOI resolving, allowing every file in a dataset to be cited and accessed in a stable, machine-actionable way. Combined with RO-Crate, this enables the inclusion of these DOI-based fragment links within the `ro-crate-metadata.json`, thereby facilitating targeted, programmatic, and automated data retrieval.

Complementing these developments, RADAR is introducing AI-supported services that improve the metadata quality, FAIRness, and interoperability of research data. Using approaches such as KeyBERT-based [8] keyword extraction, RADAR automatically identifies and suggests

domain-relevant terms from both (a) descriptive metadata fields (e.g., title, abstract, description) and (b) text-based dataset files such as CSVs, README documents, and logs. Extracted terms are mapped to controlled vocabularies and ontologies via the TS4NFDI [9] service. These AI-driven enhancements support richer semantic descriptions, improve dataset discoverability across disciplines, and provide a foundation for intelligent linking between RADAR datasets and other infrastructures.

Additionally, a new AI-supported FAIRness assessment assists depositors in evaluating and improving their datasets' compliance with FAIR principles before publication. These features strengthen RADAR's integration within the German National Research Data Infrastructure (NFDI) [10], the European Science Cloud (EOSC) [11] and the emerging FAIR Digital Object (FDO) [12] ecosystem.

Together, these innovations position RADAR as a machine-actionable, interoperable, and FAIR-by-design repository service. By adopting RO-Crate, introducing streaming-based file access, fragment identifiers for referencing files within published datasets, and integrating AI-driven metadata and FAIRness services, RADAR continues to evolve toward a more connected and intelligent research data infrastructure.

Keywords: Research Repository, RO Crate, FAIR, Fragment Identifiers, Signposting, RADAR

Data availability statement

The poster associated with this abstract will be published in a publicly accessible repository or as part of the materials published by the conference. Our consent is hereby granted.

Author contributions

Conceptualization: all; Software: J.S., S.H. and S.P.V.; Validation: J.S., S.H., S.P.V., K.S.; Investigation: J.S., S.H. and S.P.V.; Resources: K.S. and F.B.; Writing-original draft preparation: J.S.; Writing-review and editing: K.S. and F.B.; Visualization: all; Supervision: K.S. and F.B.; Project Administration: K.S. All authors have read and agreed to the published version of the abstract.

Competing interests

The authors declare that they are the developers of the here presented research data repository RADAR.

References

- [1] "RADAR | Research Data." RADAR Research Data Repository. <https://radar-service.eu> (2025-11-01)
- [2] "About us | FIZ Karlsruhe." FIZ Karlsruhe. <https://www.fiz-karlsruhe.de/en/ueber-uns/ueber-uns> (2025-11-01)
- [3] "Tape Archive Format." PRONOM. <https://www.nationalarchives.gov.uk/PRONOM/x-fmt/265> (2025-11-01)
- [4] "FAIR Signposting Profile." Signposting the Scholarly Web. <https://signposting.org/FAIR/> (2025-11-01)
- [5] S. Soiland-Reyes, P. Sefton, et al. "Packaging research artefacts with RO-Crate," Data Science, vol.5, no.2, pp. 97-138, 2022, doi:[10.3233/DS-210053](https://doi.org/10.3233/DS-210053)
- [6] "Research Object Crate (RO-Crate)." reserachobject.org. <https://www.researchobject.org/ro-crate/> (2025-11-01)

- [7] "What is a DOI?." DOI Foundation. <https://www.doi.org/the-identifier/what-is-a-doi/> (2025-11-01)
- [8] "KeyBERT: Minimal keyword extraction with BERT." GitHub. <https://github.com/MaartenGr/keyBERT> (2025-11-01)
- [9] "TS4NFDI Service Portal." Base 4 NFDI. <https://terminology.services.base4nfdi.de/> (2025-11-01)
- [10] "Association | NFDI." NFDI. <https://www.nfdi.de/association/?lang=en> (2025-11-01)
- [11] "EOSC EU Node | European Open Science Cloud - EU Node." European Open Science Cloud. <https://open-science-cloud.ec.europa.eu/about/eosc-eu-node> (2025-11-01)
- [12] "FAIR Digital Objects." FAIR Digital Objects Forum. <https://fairdo.org/> (2025-11-01)