Research Data Management (RDM) in Collaborative Research Centres must bridge disciplinary boundaries, heterogeneous data types, and evolving standards. In CRC/TRR 196 MARIE ("Mobile Material Characterization and Localization by Electromagnetic Sensing"), an integrated, modular RDM framework (Figure 1) has been developed and implemented to support the complete data lifecycle—from experiment planning to the long-term provision of FAIR (Findable, Accessible, Interoperable, Reusable) data [1]. This framework combines a suite of interconnected tools, a collaboratively developed metadata schema tailored to Terahertz (THz) research, and targeted change-management measures. Together, these components enable sustainable, reproducible, and open science across institutional boundaries and disciplinary cultures [2].

**Modular Infrastructure**

Organisation and documentation of experiments is key for understanding, reproducing and reusing the resulting data sets, especially in projects that work hand in hand as is the case in MARIE. Therefore, the **electronic lab notebook eLabFTW** was made available for all MARIE associated laboratories, ensuring provenance and reproducibility across the entire workflow [3]. Further, a central **Nextcloud collaboration hub** was established to provide an intuitive, cross-institutional file exchange which synchronizes with the large storage capacities of the University Duisburg-Essen. Acting as the project's (meta)data hub and the core of MARIE's infrastructure, an **internal Dataverse-based repository** was implemented allowing the MARIE community to collect and share versioned sets of metadata enriched data. To allow precise data description, a metadata schema for THz research was developed in close cooperation of MARIE's researchers and the NFDI4Ing-associated group Metadata4Ing to ensure connectivity to overarching standards [4, 5]. Finally, referenced but restricted data may be bit-stream preserved within the **Coscine archive** using the same metadata schema as Dataverse providing discover- and interoperability without compromising confidentiality [6]. In order to ensure the availability of the tools even beyond project end, the MARIE infrastructure was integrated into the general RDM services and the IT infrastructure of the University of Duisburg-Essen [7].

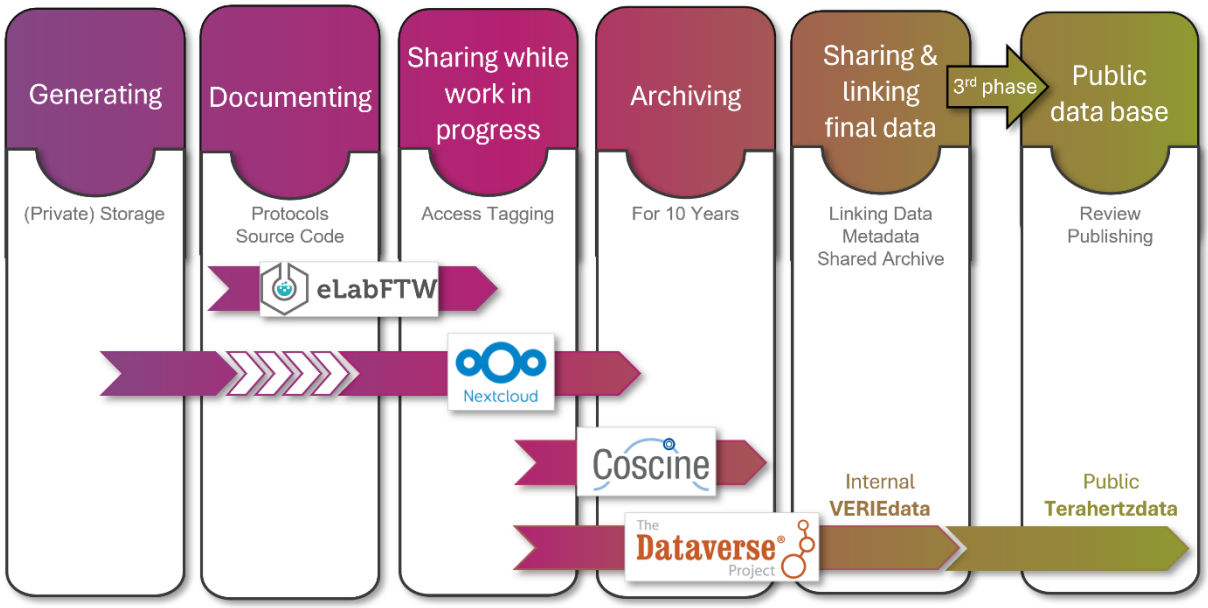


Figure 1: Illustration of the four-module core RDM infrastructure for MARIE.

However, technology alone does not change practice. The integration of MARIE's RDM tools into the work routine of THz researchers is accompanied by standardised onboarding, ongoing training and workshop programmes with hands-on practice. Furthermore, efforts in pioneering RDM tools and best practices in FAIR data handling are recognized and disseminated with a Data Champion Award.

**Public Database for Terahertz Research**

Since MARIE's approaches are operating in a previously little explored frequency regime, MARIE's generated data collections concerning material properties and device design are unique and highly relevant for further research in the terahertz community. Therefore, in the third phase, an externally available public THz database shall emerge from MARIE's internal data collection, which will serve as a constantly growing knowledge base for THz research. However, this will also lead to new challenges that need to be addressed. To further optimise the mapping of the heterogeneous and challenging data structure of THz research, data models on competence issues are continuously elaborated with the MARIE researchers. Furthermore, these are used to ensure an accurate description of the data and at the same time ensure maximum user-friendly handling. Moreover, procedures for conducting data quality assessments and internal evaluations will be designed and put in place to ensure that the data made available in the public repository meets high standards. To support this, scientifically grounded characterization criteria and quality benchmarks for channels, materials, and scenarios will be defined and applied.

The shift toward an open public repository also introduces significantly higher demands on the metadata schema: it must comply not only with MARIE-internal requirements but also align with established and emerging standards in the wider RDM ecosystem. To meet these demands, the close cooperation with NFDI4Ing and Metadata4Ing has proven essential [4].
In addition, a key challenge arises with regard to ensuring the quality and trustworthiness of stand-alone datasets that are not linked to peer-reviewed publications. For such datasets, standard peer review processes fall short, and satisfactory alternative solutions are still lacking. MARIE addresses this challenge by actively exploring structured quality assessment workflows and leveraging open access journals such as ing.grid, which aims to provide quality-assured publication mechanisms for independent research data [8].

## References

[1] Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). Comment: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*. https://doi.org/10.1038/sdata.2016.18

[2] Rehwald, Stephanie; Moussavi Alashloo, Seyed Yaser; Rolfes, Ilona; Kaiser, Thomas
Research Data Management for Terahertz Research
In: 2025 International Conference on Mobile and Miniaturized Terahertz Systems (ICMMTS) / International Conference on Mobile and Miniaturized Terahertz Systems (ICMMTS), 23-26 February 2025, Dubai, United Arab Emirates 2025

[3] Timm, H., Wittkamp, A., & Beyer, S. (2021). *Automatisiertes Deployment von Elektronischen Laborbüchern mit Ansible (CC BY-SA 4.0)*. https://doi.org/10.11588/heibooks.979.c13744

[4] Alashloo, Y., Leimer, S., Hendriks, S., & Rehwald, S. (2022). *Competency Questions in Practice: Metadata Schema for Terahertz Research*.

[5] S. Arndt, B. Farnbacher, M. Fuhrmans, S. Hachinger, J. Hickmann, N. Hoppe, M. Horsch, T. Martin, D. Iglezakis, A. Karmacharya, G. Lanza, S. Leimer, J. Munke, D. Terzijska, J. Theissen-Lipp, and C. Wiljes. "Metadata4Ing: An ontology for describing the generation of research data within a scientific activity." (1.0.0). Zenodo, 2022, https://doi.org/10.5281/zenodo.5957104

[6] Allan, D., Becker, C., Lang, I., Nellesen, M., & Politze, M. (n.d.). *RDM Platform Coscine-FAIR play integrated right from the start*. https://doi.org/10.48694/inggrid.3952

[7] Rehwald, S., & Stegemann, J. (2021). Roadmap zur Servicestelle für Forschungsdatenmanagement am Beispiel der Universitätsbibliothek Duisburg-Essen Implementation roadmap to service points for Research Data Management using the example of the University Library of Duisburg-Essen. *Information-Wissenschaft Und Praxis*, *72*(4), 194–203. https://doi.org/10.1515/iwp-2021-2161

[8] Logan, K. T., Leštáková, M., Kleinhans, A., & Pelz, P. F. (2025). *ing.grid : Bridging the Troubled Waters of Peer Review : Towards Openness, Fairness and F.A.I.R.ness*. https://doi.org/https://doi.org/10.11588/heidok.00036267