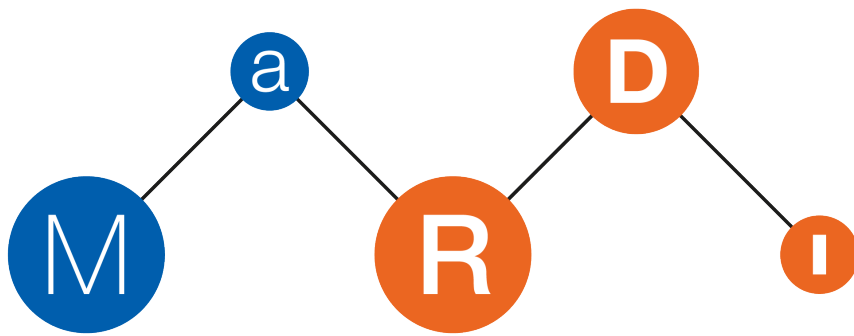# B-1 Progress Report Part 1

The Mathematical Research Data Initiative (MaRDI)

1 General Information

- Name of the Consortium: **The Mathematical Research Data Initiative, MaRDI**[1]

- Research domain of research methods addressed by the consortium is mathematics. In the initial funding phase, MaRDI has been focusing on computer algebra, scientific computing, statistics and machine learning and mathematics in interdisciplinary research. In these areas, exact, floating point and uncertain data, respectively, as well as data at the interfaces to other disciplines including mathematical models have been treated.

  Our DFG Review Board number is 3.31 in the subject area Mathematics.

- URL of the consortium website and repositories used for publishing output:

  MaRDI website: `https://www.mardi4nfdi.de/`

  MaRDI portal: `https://portal.mardi4nfdi.de/`

2 Summary

MaRDI, the only NFDI consortium dedicated to mathematical research data, has built services that make a broad range of data accessible via MaRDI's knowledge graphs and the MaRDI portal improving the reproducibility of results. Covering exact, floating point and uncertain data, respectively, as well as data at the interfaces to other disciplines including mathematical models, many of these services originate from developments in proto-typical areas such as computer algebra, scientific computing, statistics, and machine learning and are routinely advanced to the MaRDI portal via a community-oriented quality management. Thus, MaRDI is able to serve community needs in research data management (RDM) in many ways:

Contribution to existing standards and connection to international communities: MaRDI has enhanced the consistent description of software metadata in software citations and reusability of the *codemeta* standard in MathML. We have added improvements to the interoperability and discoverability of mathematical datasets with efficient metadata harvesting and exchange across repositories like DataCite. Our mathematical plugin for the general Research Data Management Organiser (MaRDMO), already available at several research institutions and universities, enables structured documentation of interdisciplinary workflows and aligns with Data Management Planning standards, as required nowadays, e.g., by the German Research Foundation (DFG) in project applications[2]. This tool not only supports the documentation and retrieval of workflows via the MaRDI portal but also enhances metadata standardisation, contributing significantly to the reproducibility and accessibility of scientific research.

---

[1] ROR `https://ror.org/04ncnzm65`
[2] Compare MaRDI's White Paper at `https://zenodo.org/records/10018246`

Internationally, MaRDI's goal is to combine EOSC and NFDI infrastructure. Correspondingly, we have successfully integrated experience and data from the mathematics case study from the FAIRCORE4EOSC project into the MaRDI portal. This involved the integration of metadata from both software and links from zbMATH Open and swMATH services with FAIRCORE4EOSC components to enhance the visibility and findability of mathematical data. In addition, a mathematical case study will be an integral part of the LUMEN project. Together with the French CNRS, LUMEN will develop a mechanism to support authors in Europe to better link software and research data in their publications. The MaRDI Layer structure for all developed services guarantees compatibility with EOSC by design.

Our interaction with the `OSCAR` community has been fruitful. The newly developed *mrdi* file format is being extended to many data types of `OSCAR` and proof of concept implementations are being developed for other computer algebra systems. The technical peer reviewing is well received at international conferences and is leading to an increase in the FAIRness of the corresponding publications. This showcases that MaRDI is significantly influencing the mathematical community outside of MaRDI and outside of Germany. Our development of the new *mrdi* file format, tailored for mathematical data, addresses challenges in data serialisation and establishes a benchmark for long-term storage and reproducibility in mathematical research.

Cultural shifts in approach to mathematical research data management. In response to the recent inclusion and gaining traction of Research Data Management plans in grant applications and accessibility of data, MaRDI has published a white paper on Research Data Management (RDM). It provides an introduction to RDM principles targeted specifically at the mathematical community and offers specific recommendations for standards to follow, emphasising the importance of metadata. The community has been reaching out to us via our Helpdesk for advice on developing data management plans and integrating metadata standards to research data. Our events are designed to engage with target communities, enhancing their understanding of RDM and the role MaRDI plays between RDM and researchers. We have organized 28 events since our launch in October 2021- for students, the scientific computing and computer algebra community, the social sciences community, and other NFDI consortia - to understand the needs of the communities. We also present our services at international conferences and annual meetings of mathematical societies and associations through mini-symposia, booths, or dedicated sessions. Responses are integrated into our strategic planning, ensuring that MaRDI remains open to changing requirements. Lastly, in May 2024, MaRDI participated in the German Science Year under the theme *Freedom* and inaugurated the interactive MaRDI station on the freedom of research data and the FAIR principles as a major exhibit on the science boat *MS Wissenschaft*. These events provide substantial benefits to the community by fostering collaboration, enhancing the understanding of RDM, creating multipliers across different institutions, and introducing new tools and workflows.

Developing a central MaRDI portal that users can query and retrieve relevant mathematical information. We have developed the first version of the MaRDI portal[3], a query platform compliant to the FAIR principles, for researchers to access our knowledge graph and contribute metadata. It has open-source accessibility, cost-free usage, user-enabled submission and retrieval of (meta)data. Simultaneously, we are developing our MaRDI knowledge graph pulling data from existing mathematical repositories for software and papers and from two newly developed MaRDI databases - MathModDB and MathAlgoDB. These knowledge graphs based on tailor-made ontologies interconnect mathematical models, algorithms, implementations, datasets, and relevant literature, making it easier for researchers to discover and access necessary information. Persistent identifiers are assigned to each dataset and tool, ensuring long-term accessibility and ease of citation. We are currently experimenting with the possibility of using large language models that enable users to interact with our knowledge graphs and connected databases via the portal.

One NFDI. We align with the vision of the NFDI and are committed to the development of basic services in the area of Data Management Planning and Knowledge Graphs. The development of central services across disciplines enhances data management practices over a broad user base and also allows the connection to other data sources built on the same Wikidata technology as our MaRDI portal. We continue to build relationships with other NFDI consortia through our regional networking event with institutes in the Berlin-Brandenburg area, the organisation of workshops on overlapping topics within physical sciences (NFDI4CAT, NFDI4ING, and MatWerk), and identifying synergies with social sciences (KonsortSWD, Text +, NFDI4Culture, BERD4NFDI) and participation in annual meetings.

---

[3]https://portal.mardi4nfdi.de/wiki/Portal

3 Composition of the consortium

Applicant institution	Location	Duration
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS) Leibniz-Institut im Forschungsverbund Berlin e. V.	Mohrenstraße 39 10117 Berlin	2020 -

Name of the consortium spokesperson	Institution, location	Duration
Michael Hintermüller email: michael.hintermueller@wias-berlin.de	WIAS Berlin Mohrenstraße 39, 10117 Berlin	2020 -

Co-applicant institutions	Location	Duration
Deutsche Mathematiker-Vereinigung e.V. (DMV) c/o WIAS; Current President: Joachim Escher	Berlin	2020 -
FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur GmbH (FIZ)	Eggenstein-Leopoldshafen	2020 -
Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e. V. für ihr Fraunhofer-Institut für Techno- und Wirtschaftsmathematik (ITWM)	Kaiserslautern	2020 -
Ludwig-Maximilians-Universität München (LMU)	München	2020 -
Mathematisches Forschungsinstitut Oberwolfach gGmbH (MFO)	Oberwolfach-Walke	2020 -
Max-Planck-Gesellschaft zur Förderung der Wissenschaften e. V. für ihr Max-Planck-Institut für Dynamik komplexer technischer Systeme (MPI DCTS)	Magdeburg	2020 -
Max-Planck-Gesellschaft zur Förderung der Wissenschaften e. V. für ihr Max-Planck-Institut für Mathematik in den Naturwissenschaften (MPI MIS)	Leipzig	2020 -
Rheinland-Pfälzische Technische Universität Kaiserslautern (RPTU)	Kaiserslautern	2020 -
Technische Universität Berlin (TUB)	Berlin	2020 -
Technische Universität München (TUM)	München	2020 -
Universität Leipzig (UL)	Leizig	2020 -
Universität Münster (Uni Münster)	Münster	2020 -
Universität Stuttgart (USTUTT)	Stuttgart	2020 -
Zuse Institut Berlin (ZIB)	Berlin	2020 -

Names of co-spokesperson	Institution, location	Task area(s)	Duration
Michael Hintermüller	WIAS, Berlin	T7	2020 -
President of DMV (2021-2022: Ilka Agricola)	DMV, Berlin	T6	2020 - 2022
President of DMV (Jan 2023: Joachim Escher)	DMV, Berlin	T6	2023 -
Peter Benner	MPI DCTS, Magdeburg	T2	2020 -
Bernd Bischl	LMU, München	T3	2020 -
Wolfram Decker	RPTU, Kaiserslautern	T1	2020 - 2023
Mathias Drton	TUM, Garching	T3	2020 -
Claus Fieker	RPTU, Kaiserslautern	T1	2020 -
Dominik Göddeke	USTUTT, Stuttgart	T4	2020 -
Michael Joswig	TUB, Berlin	T1	2020 -
Stephan Klaus	MFO, Oberwolfach-Walke	T6	2020 -
Mario Ohlberger	Uni Münster, Münster	T2	2020 -
Harald Sack	FIZ, Eggenstein-Leopoldshafen	T5	2020 -
Anita Schöbel	ITWM, Kaiserslautern	T4	2020 -
Christof Schütte	ZIB, Berlin	T5	2020 -
Rainer Sinn	UL, Leipzig	T6	2020 -
Bernd Sturmfels	MPI MIS, Leipzig	T6	2020 -

Participating institutions	Location	Duration
European Mathematical Society Department of Mathematics and Statistics	Helsinki, Finland	2020 -
Gesellschaft für Angewandte Mathematik und Mechanik e. V. (GAMM)	Dresden	2020 -
Gesellschaft für Operations Research e. V. (GOR)	Aachen	2020 -
IMAGINARY gGmbH	Berlin	2020 -
Hausdorff Center for Mathematics	Bonn	2023 -
Universität Bonn	Bonn	2023 -
Universität Marburg	Marburg	2023 -

Participating individuals	Location	Duration
Peter Bastian for the Exzellenzcluster "STRUCTURES: A unifying approach to emergent phenomena in the physical world, mathematics, and complex data" (EXC 2181)	Heidelberg	2020 -
Bettina Eick Technische Universität Braunschweig	Braunschweig	2020 -
Thomas Ertl for the Exzellenzcluster "Data-integrated Simulation Science" (EXC 2075)	Stuttgart	2020 -
Dieter Fellner for the Fraunhofer-Verbund IUK-Technologie	Berlin	2020 -
Michael Kohlhase Friedrich-Alexander Universität Erlangen-Nürnberg (FAU)	Erlangen	2020 -
Martin Skutella for the Exzellenzcluster MATH$^+$ (EXC 2046)	Berlin	2020 -