

## 1 Binding letter of intent as advance notification of a full renewal proposal

- ✓ Binding letter of intent (required as advance notification for proposals in 2025)

## 2 Formal details

- Name of the consortium: **The Mathematical Research Data Initiative**
- Acronym of the consortium: **MaRDI**
- Applicant institution: Weierstrass Institute for Applied Analysis and Stochastics (WIAS)  
Mohrenstrasse 39, 10117, Berlin  
Prof. Dr. Michael Hintermüller (Director)
- Spokesperson: Prof. Dr. Michael Hintermüller  
E-mail: michael.hintermueller@wias-berlin.de  
Institutional affiliation: WIAS

Co-applicant institutions (new partners in bold)	Location
Deutsche Mathematiker-Vereinigung e.V. (DMV) c/o WIAS <b>President: Jürg Kramer; 2025 - 2026</b>	Mohrenstraße 39 10117 Berlin
FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur GmbH (FIZ) President & CEO: Wolfram Horstmann	Hermann-von-Helmholtz-Platz 1 76344 Eggenstein-Leopoldshafen
Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e. V. für ihr Fraunhofer-Institut für Techno- und Wirtschaftsmathematik (ITWM) Director: Anita Schöbel	Fraunhofer-Platz 1 67663 Kaiserslautern
<b>IMAGINARY gGmbH</b> Managing Director: Andreas Matt	Mittenwalder Str. 48 10961 Berlin
Ludwig-Maximilians-Universität München (LMU) President: Bernd Huber	Geschwister-Scholl-Platz 1 80539 München
Max-Planck-Gesellschaft zur Förderung der Wissenschaften e. V. für ihr Max-Planck-Institut für Dynamik komplexer technischer Systeme (MPI DCTS) Managing Director: Peter Benner	Sandtorstraße 1 39106 Magdeburg
Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau (RPTU) President: Malte Drescher	Gottlieb-Daimler-Straße 47 67653 Kaiserslautern
Technische Universität Berlin (TUB) President: Geraldine Rauch	Straße des 17. Juni 135 10623 Berlin
Technische Universität München (TUM) President: Thomas Hofmann	Arcisstraße 21 80333 München



Co-applicant institutions (new partners in bold)	Location
<b>Universität Bonn</b> (UBonn) Rector: Michael Hoch	Dechenstraße 3-11 53115 Bonn
Universität Stuttgart (USTUTT) Rector: Peter Middendorf	Keplerstraße 7 70174 Stuttgart
Universität Münster (UM) Rector: Johannes Wessels	Schloßplatz 2 48149 Münster
Zuse Institute Berlin (ZIB) President: Christof Schütte	Takustraße 7 14195 Berlin

Co-Spokesperson/E-mail (new partners in bold)	Institutional affiliation
<b>Jürg Kramer</b> juerg.kramer@hu-berlin.de	DMV
<b>Moritz Schubotz</b> moritz.schubotz@fiz-karlsruhe.de	FIZ
Anita Schöbel anita.schoebel@itwm.fraunhofer.de	ITWM
<b>Andreas Matt</b> andreas.matt@imaginary.org	IMAGINARY
Bernd Bischl bernd.bischl@stat.uni-muenchen.de	LMU
Peter Benner benner@mpi-magdeburg.mpg.de	MPI DCTS
Claus Fieker claus.fieker@rptu.de	RPTU
Michael Joswig joswig@math.tu-berlin.de	TUB
Mathias Drton mathias.drton@tum.de	TUM
<b>Martin Rumpf</b> martin.rumpf@uni-bonn.de	UBonn
Dominik Göddeke dominik.goeddeke@mathematik.uni-stuttgart.de	USTUTT
Mario Ohlberger ohlberger@uni-muenster.de	UM
Christof Schütte schuette@zib.de	ZIB



Participating institutions	Location
European Mathematical Society Department of Mathematics and Statistics	P.O.Box 68 00014 University of Helsinki, Finland
Mathematisches Forschungsinstitut Oberwolfach gGmbH (MFO) Director: Gerhard Huisken	Schwarzwaldstraße 9-11 77709 Oberwolfach-Walke
Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e. V. für ihren Fraunhofer-Verbund IUK-Technologie Spokesperson: Boris Otto	Anna-Louisa-Karsch-Straße 2 10178 Berlin
Gesellschaft für Angewandte Mathematik und Mechanik e. V. (GAMM) President: Karsten Urban	Geschäftsstelle der GAMM c/o Institut für Statik und Dynamik der Tragwerke Fakultät Bauingenieurwesen, 01062 Dresden
Gesellschaft für Operations Research e. V. (GOR) President: Jutta Geldermann	Kackertstraße 7, 52072 Aachen
Participating individuals (new partners in bold)	Location
<b>Prof. Dr. Thomas Nikolaus</b> for <b>Exzellenzcluster Mathematik Münster:</b> Dynamik–Geometrie–Struktur (EXC 2044)	Universität Münster Einsteinstraße 62 48149 Münster
<b>Prof. Dr. Sebastian Pokutta</b> for Exzellenzcluster MATH <sup>+</sup> (EXC 2046)	Technische Universität Berlin Schr. MA 2-2 Strasse des 17. Juni 136 10623 Berlin
<b>Prof. Dr. Stefan Müller</b> for Hausdorff Center for Mathematics at the University of Bonn (EXC 2047)	Mathematik Zentrum der Universität Bonn Endenicher Allee 60 53115 Bonn



### 3 Objectives, work programme and research environment in the second funding period

Research area of the proposed consortium (DFG classification system): 33.1 Mathematics

**Objectives.** In the current funding phase, *MaRDI* successfully laid the foundations for a national research data infrastructure in mathematics. We developed a broad range of interoperable services, curated data sources, and a robust technical backbone, including the *MaRDI* Portal and its extensive knowledge graph, which now contains over seven million items and more than 700 million semantic relations. These efforts have significantly advanced the FAIRness of mathematical research data and strengthened the culture of data sharing and reproducibility within the mathematical community, and neighbouring disciplines.

Building on these achievements, the next funding phase will focus on deepening, scaling, and extending our infrastructure and services. The overarching goal is to evolve *MaRDI* into a sustainable, community-driven ecosystem that not only ensures the reproducibility of research but also fosters true re-usability of mathematical data and knowledge across domains and disciplines. To this end, we formulate the following objectives:

**O1: From confirmability to re-usability:** Confirmability and reproducibility of results and workflows remains central to a mathematical approach to research data through the usage of stringent logic and derivation in mathematics. The next phase will focus on the re-use of existing mathematical facts and data to propel new research. Having established a number of data sources and a FAIR portal infrastructure, further development will focus on services that enable the R in FAIR and extend reproducibility to include reusability of mathematical research data.

**O2: Scaling of services along cross-cutting themes:** While the current funding phase of *MaRDI* led to the development of multiple - partly prototypical - services from within the disciplinary oriented task areas, we will focus on scaling and maturing them in the next phase, also addressing interdisciplinary needs. Services should include data sources from a broad interdisciplinary user community as well as from fields that were not originally targeted. E.g., the MathAlgoDB will therefore not only contain algorithms from external users within scientific computing but also from other fields like computer algebra, optimisation, or statistics to mention a few.

**O3: Extension into other mathematical domains:** One of the reasons to re-structure the work program for the next phase, see below, across cross-cutting topics is to allow for integration of new sub-communities beyond the original mathematical sub-disciplines. This includes consolidation and expansion of relevant structures. E.g., the development of metadata schemes and file formats within the sub-domains, like the *mrdi* file format, will be extended for use in other areas. Furthermore, *MaRDI* seeks to establish a process to include contributions from mathematical areas not yet covered in collaborations for specific goals.

**O4: Community involvement:** Scaling of many *MaRDI* services is only possible through community engagement as sources for data and knowledge including curation. *MaRDI* will there-



fore focus on training, user interfaces and suitable editorial processes to ensure growth of the data sources while maintaining high data quality and data curation.

**O5: Training and outreach:** The importance of research data management and knowledge curation in mathematics will remain as a cornerstone of scientific progress when it is embedded within a cultural shift that increasingly values data handling in science, generally, and in mathematics, specifically. *MaRDI* will implement targeted training programmes addressing researchers at various career levels and outreach to stakeholders within and outside of academia.

**O6: Integration with industry partners:** Mathematics is often the methodological foundation not only for scientific endeavours but also within applied projects in industry, with mathematical data and knowledge gaining increasing importance in collaborations such as simulations, statistical methods or artificial intelligence. Thus, the interface of mathematics and industry will be explicitly considered in the next funding phase.

**O7: Internationalisation:** Mathematical data and knowledge is ever acquired in international collaborations. However, national initiatives like the NFDI or *MaRDI* are not well-known internationally. Thus, the objective of the next funding phase is also to grow beyond national borders for community integration and user base establishment to increase the contributions to *MaRDI*'s infrastructure from a much larger pool of scientists.

**Task Areas.** During the next funding phase, integration of data via the cross-cutting themes will be crucial to fulfilling the objectives above. Specifically, we identified a number of task areas where the work program will be focussed on:

- Task Area A: Workflows: Dedicated to the comprehensive documentation of the mathematical research process and the use of data, methods, and software therein, in a comparable manner to electronic lab books in experimental sciences.
- Task Area B: Guidelines: Provides guidelines for reproducible research in mathematics.
- Task Area C: Data & Databases: Concerns itself with mathematical data sources, including the knowledge graph provided by *MaRDI*.
- Task Area D: Benchmarks: Covers important collections that are crucial for (applied) mathematical research problems as well as for other disciplines within the NFDI.

While *MaRDI* has already been implementing training and outreach during the current funding phase, work needs to be both targeted and broad in the next phase within the newly outlined

- Task Area E: Training & Outreach will provide a community bridge and extensive and scalable training material and "training the trainer" programs.

Like in the current funding phase, the

- Task Area F: *MaRDI* Portal & Infrastructure will integrate all services and data sources into a unified FAIR infrastructure. In the next phase, it will further expand the *MaRDI* knowledge graph and enhance the existing user interface with AI-based support for exploration, navigation, and improved tools for data contribution and curation.



**Usage of infrastructures, tools, services.** In the current funding phase, *MaRDI* looked at specific sub-disciplines in mathematics where an understanding of the community needs helped to build and shape the initial services. In particular, with the *MaRDI* Portal the consortium developed a FAIR infrastructure that serves (1) as an entry point to *MaRDI* services and (2) as a metadata hub for mathematical research data. It currently contains more than seven million items and over 700 million triplets in its *MaRDI* knowledge graph. Thus, it is essential that the future of the *MaRDI* portal, currently hosted at ZIB, is secured during the next funding phase. Technologically, it is based on MediaWiki and WikiBase technologies, thus relying on free and open software with a large developer community and fitting well into the storage strategies for many NFDI consortia. The technological compatibility with WikiData enables data exchange between various platforms, and in particular also between *MaRDI* and any WikiBase compatible platforms, such as WikiData. This makes *MaRDI* a valuable source for mathematical research data. As the *MaRDI* knowledge graph is a central source of connected information on mathematical data, objects and knowledge *MaRDI* actively contributes to the developments of the KGI4NFDI Base Service. Where *MaRDI* focuses on knowledge graphs, it relies on ontology developments within *MaRDI* (NFDIcore Ontology) or domain specific ontologies, e.g., QUDT for quantities and units in interdisciplinary applications.

While the platform is operating independently, the *MaRDI* Portal regularly harvests new entries for increased findability and integration into the *MaRDI* infrastructure. Similarly, *MaRDI* imports metadata from mathematical datasets stored at Zenodo making it easier for the user to find appropriate data for their research. *MaRDI* harvests the arXiv, one of the main preprint services for mathematics (and physics), data sources like the Digital Library of Mathematical Functions (DLMF), services like zbMATH Open for mathematical publications and swMATH and the Software Heritage for research software in mathematics. Further, metadata from the OpenML platform are included, which allows connections between mathematical research data and datasets, algorithms, and experiments from the machine learning community.

To assist research data management planning, *MaRDI* provides the MaRDMO plugin for the Research Data Management Organizer (RDMO). The plugin provides questionnaires specific to the mathematical data life cycles and will be extended to incorporate more mathematical use cases in the next funding phase. Through this, *MaRDI* actively contributes to the Base Service DMP4NFDI.

For computations, *MaRDI* considers Jupyter4NFDI and MultiCloud developments to provide mathematical notebooks and computations. Furthermore, central services for persistent identifiers (PID4NFDI), identity management (IAM4NFDI) for the *MaRDI* portal or terminology service (TS4NFDI) for interdisciplinary mathematical applications are important.

*MaRDI* integrates with the European Open Science Cloud (EOSC), e.g., through the FAIR-CORE4EOSC and LUMEN projects. From the beginning, the *MaRDI* layer structure X1: Core, X2: Data, X3: Exchange, and X4: Knowledge has facilitated the compatibility with EOSC and OpenAIRE.



*MaRDI* also relies on general international infrastructures like GitHub for development or the training of large-language models, e.g., for human-machine interaction, which generally is a very expensive endeavour.

**Regional Collaboration.** Our goals in engaging with institutes, either associated with or belonging to NFDI consortia in Berlin and Brandenburg, were to build networks and continually identify synergies at a regional level. We organised the inaugural regional NFDI meeting in October 2023, with participation from 24 out of 26 consortia. This meeting identified knowledge graphs and RDM in the curriculum as key cross-cutting topics of interest. *MaRDI* has continued to be active in organising subsequent chapter meetings in July 2024 and May 2025 (together with NFDI4DataScience) on specific topics such as the application of knowledge graphs and ontologies among different disciplines and AI in science. Activities will broaden with next thematic meetings at Wikimedia and in collaboration with NFDIxCS and strategic focus on cross-connecting consortia from all disciplines on a regional level.

All co-spokespersons have actively contributed to achieving *MaRDI*'s goals and propel mathematical research data management in their home institutions, and the other research consortia, i.e., excellence clusters, collaborative research centers etc, they are involved in.

**Cross-Consortia Collaborations.** *MaRDI*'s partnerships with other consortia within the physical and social sciences have helped to design interdisciplinary workflow templates. Our collaboration with NFDI4Cat focuses on consolidating existing ontologies for chemical reactors and integrating them into computational science and engineering workflows. We have also worked closely with NFDI4Ing and NFDI4Chem, leading to significant advancements in workflow templates and ontology integration. We invited representatives from the social sciences consortia for a digital humanities workshop to understand how we could incorporate their workflows into our templates and recently, we have entered a collaboration with MatWerk and NFDI4Culture on interdisciplinary workflows describing and including mathematical methods. The participation in panel discussions and in sessions during annual meetings of other consortia are further engagements with the NFDI community. Additionally, our partnership with the Munich Data Science Institute promotes synergistic relationships among consortia and facilitates knowledge sharing on research data related topics. With BERD4NFDI and NFDI4DataScience we collaborate on topics from machine learning and artificial intelligence.

*MaRDI* also actively collaborates with DAPHNE, FAIRmat, MatWerk, NFDI4Cat, NFDI4Chem, and PUNCH as part of the physical sciences in NFDI, organising talks given by experts on topics common to these consortia, showcasing good data practises.





## 4 International and national networking

Mathematics is a discipline in which data do not have geographical limitations, i.e., services that *MaRDI* develops are directly applicable, relevant for the international community.

*MaRDI* builds a research infrastructure for mathematical research data. Thereby, it collects from diverse international sources such as arXiv, LMFDB, the Software Heritage, or DMLF, to name a few. Data that have been harvested and curated within *MaRDI* is generally fed back into the original sources. Specific developments, e.g., for technical peer-review, software ecosystems or data sources, are co-developed within the respective sub-communities of mathematics.

*MaRDI* services are available free of charge to the international mathematical community with the *MaRDI* Portal as entry point to the services and the *MaRDI* knowledge graph. E.g., the zbMATH Open Service by FIZ Karlsruhe integrated into *MaRDI* provides reviews of current articles, conference papers, books, and other publications within mathematics and attracts a broad international community.

Our contributions to international infrastructures also include alignment to the EOSC. A prominent example is data from the Math case study of the FairCore4EOSC project that has been fully integrated into the *MaRDI* Portal. In the same vein, we have successfully included mathematics as a case study into the EOSC project LUMEN. Together with the CCSD/HAL within the National Center for Scientific Research (CNRS) from France, this cross-domain project will develop a mechanism to support authors in Europe to better link software and research data in their publications. We will make efforts to keep abreast of NFDI activities and Base4NFDI in alignment with the new EOSC EU Node services. *MaRDI* will contribute tools and services to the EOSC EU Node and participate in RDA Working Groups or Interest groups relevant to the field of mathematics.

*MaRDI* is part of the newly created R Consortium Infrastructure Steering Committee (ISC) working group "Marshalling and Serialization in R". The newly founded GAMM Fachausschuss "Research Software Engineering and Research Data Management in Mathematics and Mechanics" is co-headed by a *MaRDI* member and will serve as a structural connection to the large community organised within the GAMM.

*MaRDI* pitches its services and consortium at international conferences and to other national entities to establish contact and identify synergies and areas for future collaborations. For example, we presented *MaRDI* and our services to the Australian Research Data Commons during a visit to the Australian research institutes in February 2024 organized by the Leibniz association. A similar presentation will be given to the Digital Research Alliance of Canada, specifically the Digital Research Infrastructure, later in 2025.

*MaRDI*'s connection with the Society for Industrial and Applied Mathematics (SIAM) based in Philadelphia, USA, gives access to an international community of applied mathematicians, computer scientists, engineers, and researchers in quantitative life sciences and other scientific fields, in particular, the Computational Sciences and Engineering (CSE) community.

