

MaRDMO Plugin

Document and Retrieve Interdisciplinary Workflows Using the MaRDI Portal.

Marco Reidelbach¹[\[https://orcid.org/0000-0002-1919-1834\]](https://orcid.org/0000-0002-1919-1834), Eloi Ferrer¹[\[https://orcid.org/0009-0009-2327-2619\]](https://orcid.org/0009-0009-2327-2619),
and Marcus Weber¹[\[https://orcid.org/0000-0003-3939410X\]](https://orcid.org/0000-0003-3939410X)

¹Mathematics of Complex Systems, Zuse Institute Berlin, Berlin, Germany

Abstract: MaRDMO, a plugin for the Research Data Management Organiser, was developed in the Mathematical Research Data Initiative to document interdisciplinary workflows using a standardised scheme. Interdisciplinary workflows recorded this way are published directly on the MaRDI portal. In addition, central information is integrated into the MaRDI knowledge graph. Next to the documentation, MaRDMO offers the possibility to retrieve existing interdisciplinary workflows from the MaRDI Knowledge Graph to allow the reproduction of the initial work and to provide scientists with new research impulses. Thus, MaRDMO creates a community-driven knowledge loop that could help to overcome the replication crisis.

Keywords: Interdisciplinary Workflows, RDMO, MaRDI Portal

1 Introduction

Reproducibility is a key element of science. Increasing reports about reproduction problems [1], [2] impair the scientific credibility and limit synergies arising from reusing and advancing scientific work. The reasons for reproducibility issues are manifold, including intentional misbehaviour [3], but also incomplete documentation of applied methodologies and software, missing parameters or conditions, and inconsistent naming schemes [2].

To overcome these issues the Mathematical Research Data Initiative (MaRDI [4]) developed a standardised documentation scheme to summarise the relevant aspects of interdisciplinary workflows independent of the research area and its design (theoretical/experimental) [5]. Thereby, scientists from various disciplines are able to document their interdisciplinary workflows and publish them on the MaRDI portal, a unique portal that will contain all sorts of mathematical research data. Likewise, scientists are able to retrieve documentations through the corresponding Knowledge Graph (KG) to reproduce, develop, and re-document the original work to retain the knowledge in a reproducible form.

To enable scientists from the entire research community to document and retrieve interdisciplinary workflow documentations using the MaRDI portal, a low-threshold access is necessary. This means that the access should ideally be integrated into the

normal working day, without the need to call up the MaRDI portal directly and edit/query it manually.

The Research Data Management Organiser (RDMO) is an open source web application for research data management. It provides discipline-specific question catalogues and allows the output of data management plans. Currently, RDMO is used in many research institutions in Germany and tested in other European countries [6]. It is very likely that use will continue to increase in the coming years, especially since most of the National Research Data Infrastructure consortia have endorsed the use of RDMO [7]. Thus, it is an integral part of the work of many scientists from the entire research community.

Since RDMO encourages the design of individual question catalogues and additional functionalities, a question catalogue for interdisciplinary workflow documentation and a functionality connecting RDMO with the MaRDI portal could establish the desired low-threshold portal access.

Here, we present MaRDMO, a plugin for RDMO, to document and retrieve interdisciplinary workflows using the MaRDI Portal.

2 MaRDMO Plugin

The MaRDMO Plugin contains a questionnaire and an export/query functionality. It adds a "MaRDI Export/Query" button to the project view by which the entire plugin is controlled. The remaining settings, cf. 1, are made via the question catalogue.

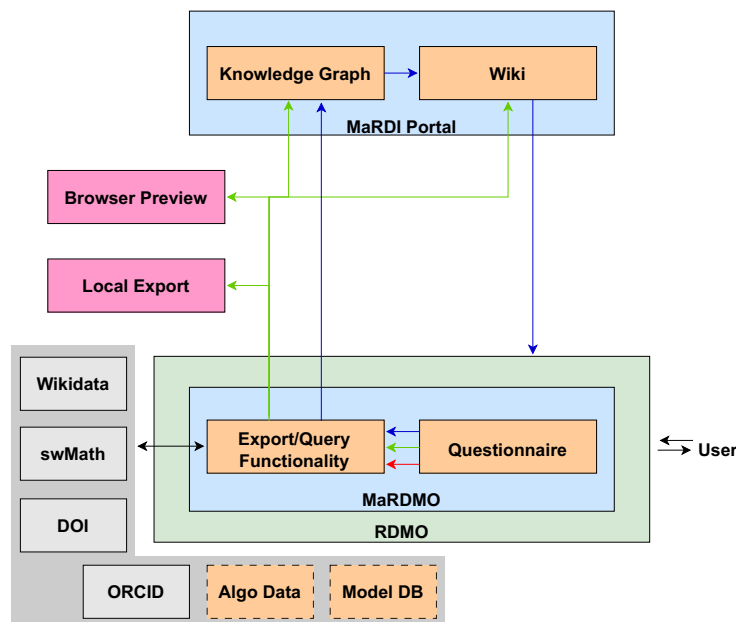


Figure 1. Overview of MaRDMO - MaRDI portal interaction. Interdisciplinary workflow documentation (green) and retrieval (blue) pathways, plugin settings (red) and interactions with other data sources (black) are indicated by arrows. AlgoData and ModelDB are MaRDI services which will be integrated in the future.

2.1 Question Catalogue

The MaRDMO questionnaire is divided in 7 sections, serving three purposes, i.e. documentation or retrieval of interdisciplinary workflows <https://de.overleaf.com/project/64e61490cb3d6285a> and plugin settings.

To do a documentation questions providing a general description, e.g. research objective, involved disciplines and data streams, describing the underlying mathematical model, e.g. model description, involved variables and parameters and discretisation, providing the relevant process information, e.g. process steps, applied algorithms, used software, hardware, experimental devices, input and output data, and describing the reproducibility need to be answered. All entries should be provided with identifiers, e.g. MaRDI identifiers, Wikidata QIDs, DOIs and swmath IDs, to properly integrate the individual documentations into the existing research data landscape.

To retrieve documentations the entities to search for, e.g. research objective, applied model, methods and software, need to be defined. Entities may be searched via keywords or identifiers.

2.2 Export/Query Functionality

The export/query functionality gathers all the information provided via the questionnaire and creates a documentation or retrieves documentations by an entity search.

For documentations, all answers are integrated into MaRDI's predefined documentation scheme. Depending on the user's needs, the completed scheme is then saved locally, previewed in the browser, e.g. to detect faulty rendering, or sent to the MaRDI portal to create a wiki page.

If the documentation is sent to the MaRDI portal, individual components, e.g. name, associated publication, research objective, model, methods, software and input data, are also integrated into the KG of MaRDI.

To exemplarily link a new interdisciplinary workflow item with an associated publication, the following steps are performed:

- Check if an item with the custom DOI already exists in MaRDI KG, and if so, link it to the new interdisciplinary workflow item.
- Check if an item with the custom DOI already exists in Wikidata KG, if so, integrate it into MaRDI KG and link it to the new interdisciplinary workflow item.
- If an item with the custom DOI cannot be found in either KG, get the full citation from DOI and ORCID, create a new item in MaRDI KG, and link it to the new interdisciplinary workflow item. Repeat the first two steps for the authors and journal. If no entries are found, create them and link them to the new publication entry.

To retrieve documentations, all answers are translated into a SPARQL query. If relevant information is found in the MaRDI KG, appropriate documentations are returned as wiki pages.

3 Conclusion

MaRDMO is a software package that enables the entire research community to document and retrieve interdisciplinary workflows using established infrastructure (RDMO). Documentations are integrated into the existing research data landscape to exploit net-

work effects and provide scientists with instant visibility, while the retrieval of documentations brings together new research impulses from various disciplines. Due to its simple structure, MaRDMO may also be used with other questionnaires (and KGs) to document and search research objects other than interdisciplinary workflows.

Data availability statement

The MaRDMO Plugin is available and documented on <https://github.com/MarcoReidelbach/MaRDMO>.

Underlying and related material

Knowledge Graph settings (items and properties) required for MaRDMO and a script to make any Wikibase Knowledge Graph compatible with MaRDMO can be found on <https://github.com/MarcoReidelbach/MaRDMO>.

Author contributions

Marco Reidelbach (Conceptualization, Software, Methodology, Writing - Original Draft, Writing - review & editing), Eloi Ferrer (Software, Methodology, Writing - review & editing) and Marcus Weber (Conceptualization, Writing - review & editing, Funding acquisition)

Competing interests

The authors declare that they have no competing interests.

Funding

Marco Reidelbach, Eloi Ferrer and Marcus Weber are supported by MaRDI, funded by the Deutsche Forschungsgemeinschaft (DFG), project number 460135501, NFDI 29/1 "MaRDI – Mathematische Forschungsdateninitiative".

Acknowledgements

We thank MaRDI's Task Areas 4 and 5 for fruitful discussions concerning workflow documentation and portal integration.

References

- [1] M. Baker, "1,500 scientists lift the lid on reproducibility," *Nature*, vol. 533, pp. 452–454, May 2016. DOI: [10.1038/533452a](https://doi.org/10.1038/533452a).
- [2] Krishna Tiwari, Sarubini Kananathan, Matthew G. Roberts, Johannes P. Meyer, Mohammad Umer Sharif Shohan, Ashley Xavier, Matthieu Maire, Ahmad Zyoud, Jinghao Men, Szeyi Ng, Tung V. N. Nguyen, Mihai Glont, Henning Hermjakob and Rahuman S. Malik-Sheriff, "Reproducibility in systems biology modelling," *Molecular Systems Biology*, vol. 17, e9982, Feb. 2021. DOI: [10.15252/msb.20209982](https://doi.org/10.15252/msb.20209982).
- [3] D. Fanelli, "How many scientists fabricate and falsify research? a systematic review and meta-analysis of survey data," *PLOS ONE*, vol. 4, no. 5, e5738, May 2009. DOI: [10.1371/journal.pone.0005738](https://doi.org/10.1371/journal.pone.0005738).

- [4] Christiane Görge and Rainer Sinn, "Mathematik in der nationalen forschungsdateninfrastruktur," *Mitteilungen der Deutschen Mathematiker-Vereinigung*, vol. 29, no. 3, pp. 122–123, Oct. 2021. DOI: [10.1515/dmvm-2021-0049](https://doi.org/10.1515/dmvm-2021-0049).
- [5] Tobias Boege, René Fritze, Christiane Görge, Jeroen Hanselman, Dorothea Iglezakis, Lars Kastner, Thomas Koprucki, Tabea Krause, Christoph Lehrenfeld, Silvia Polla, Marco Reidelbach, Christian Riedel, Jens Saak, Björn Schembera, Karsten Tabelow and Marcus Weber, "Research-data management planning in the german mathematical community," Nov. 2022. DOI: [10.48550/arXiv.2211.12071](https://doi.org/10.48550/arXiv.2211.12071).
- [6] R. Arbeitsgemeinschaft. "RDMO - Cooperations." (2023), [Online]. Available: <https://rdmorganiser.github.io/en/cooperations/> (visited on 04/16/2023).
- [7] Harry Enke, Daniela Hausen, Christin Henzen, Gerald Jagusch, Celia Krause, Sabine Schönau, Lukas Weimer and Jürgen Windeck, "Data management planning: Concept for setting up a working group in the nfdi section common infrastructures," Jan. 2023. DOI: [10.5281/zenodo.7540682](https://doi.org/10.5281/zenodo.7540682).