Enabling FAIR Data in Computational Science, Engineering and Mathematics through Knowledge Graphs

Björn Schembera[1,*], Cedric Riethmüller[1], Dominik Göddeke[1,2]

[1]) Institute of Applied Analysis and Numerical Simulation, University of Stuttgart, Stuttgart, Germany
[2]) Stuttgart Center for Simulation Science, University of Stuttgart, Stuttgart, Germany
*Corresponding author; bjoern.schembera@ians.uni-stuttgart.de

In computational science, engineering and mathematics, dark data is on the rise [1, 2]: Huge amounts of data are neither findable, accessible, interoperable, nor reusable. As such, research data may be uncurated, unavailable, unannotated, biased, or incomplete. Those data are diametrically opposed to the FAIR principles. One mean to align the data with the FAIR principles with

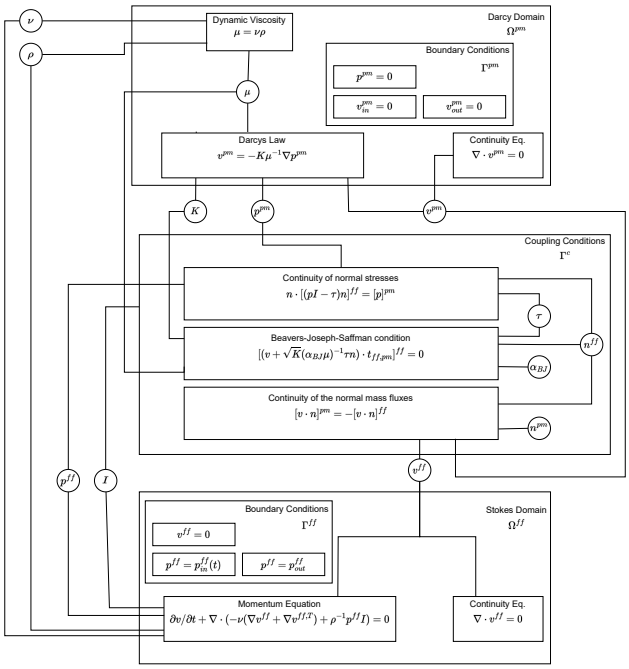

Figure 1: A Model Pathway Diagram of a coupled Stokes-Darcy model.

respect to reproducible research is the introduction of semantic technology, including metadata, ontologies and knowledge graphs. This has been recognised by the DFG and consequently, the NFDI (Nationale Forschungsdateninfrastruktur) programme has been established. It is here that NFDI consortia, especially the Mathematical Research Data Initiative (MaRDI) [3], can provide dedicated

support by encouraging academic communities to engage in community-driven metadata standardization and ontology development. Along case studies from manifold fields of the applied sciences, such as engineering, materials science, and high-performance computing, mathematical concepts and workflows are surveyed, analyzed and categorized to create a bi-directional bridge between mathematical developments and other disciplines. In the poster it will be highlighted how a case study on the comparison of partitioned coupling and monolithic block-preconditioning approaches for solving Stokes-Darcy systems [4], originally stemming from work conducted within SimTech, enables MaRDI to carve out relations, interdependencies and details of the underlying mathematical models. These are displayed on the MaRDI portal as a Wiki and in Model Pathway Diagrams [5], as depicted in figure 1. This is a preceding step towards more semantic technology, i.e. a knowledge graph of models (as depicted in figure 2), as models are at the intersection of mathematics and many other disciplines, i.e., the interdisciplinarity that is reflected by SimTech. The ontology underneath the knowledge graph will be presented.

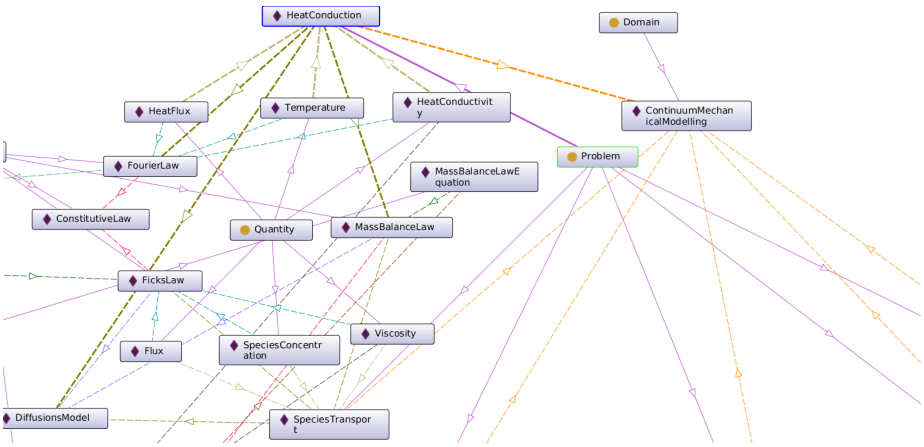

Figure 2: Heat conduction and species transportation and their relations to quantities and the application domain as data inside the knowledge graph.

References

[1] Schembera, Björn, and Juan M. Durán. "Dark data as the new challenge for big data science and the introduction of the scientific data officer." Philosophy & Technology 33 (2020): 93-115.

[2] Boege, Tobias, et al. "Data Management Planning in the German Mathematical Community." arXiv preprint arXiv:2211.12071 (2022).

[3] Benner, Peter, et al. "Die mathematische Forschungsdateninitiative in der NFDI: MaRDI (Mathematical Research Data Initiative)." GAMM Rundbrief 1 (2022): 40-43.

[4] Schmalfuss, Jenny, et al., Partitioned Coupling vs. Monolithic Block-Preconditioning Approaches for Solving Stokes-Darcy Systems, in: 9th edition of the Int. Conf. on Computational Methods for Coupled Problems in Science and Engineering (2021).

[5] Koprucki, Thomas, et al. "Model pathway diagrams for the representation of mathematical models." Optical and Quantum Electronics 50 (2018): 1-9.