

Project MetaMoSim: Generic metadata management for reproducible high-performance-computing simulation workflows

Thursday 6 October 2022 09:40 (15 minutes)

Modern science is to a vast extent based on simulation research. With the advances in high-performance computing (HPC) technology, the underlying mathematical models and numerical workflows are steadily growing in complexity.

This complexity gain offers a huge potential for science and society, but simultaneously constitutes a threat for the reproducibility of scientific results. A main challenge in this field is the acquisition and organization of the metadata describing the details of the numerical workflows, which are necessary to replicate numerical experiments, and to explore and compare simulation results. In the recent past, various concepts and tools for metadata handling have been developed in specific scientific domains. It remains unclear to what extent these concepts are transferable to HPC based simulation research, and how to ensure interoperability in the face of the diversity of simulation based scientific applications. This project aims at developing a generic, cross-domain metadata management framework to foster reproducibility of HPC based simulation science, and to provide workflows and tools for an efficient organization, exploration and visualization of simulation data. Within the project, we so far did a review of existing approaches from different fields. A plethora of tools around metadata handling and workflows have been developed in the past years. We identified tools and formats like the odML that are useful for our work. The metadata management framework will address all components of simulation research and the corresponding metadata types, including model description, model implementation, data exploration, data analysis, and visualization. We have now developed a general concept to track, store and organize metadata. Next, the required tools within the concept will be developed such that they are applicable both in the Computational Neuroscience and Earth and Environmental Science.

Please assign your poster to one of the following keywords.

In addition please add keywords.

Please assign yourself (presenting author) to one of the stakeholders.

Please specify "other" (stakeholder)

Primary author: THOBER, Stephan (Dept. Computational Hydrosystems, Helmholtz-Centre for Environmental Research, Leipzig, Germany)

Presenter: THOBER, Stephan (Dept. Computational Hydrosystems, Helmholtz-Centre for Environmental Research, Leipzig, Germany)

Session Classification: Session