

Seminar im Rahmen des GRK 2078

Referee: **Dr.-Ing. Michael Selzer**
Multiscale Materials Modelling and Data Processing,
IAM – Computational Materials Science, KIT, Germany

Date: Tuesday, November 17, 2020
Time: 14:00 h
Format: IRTG Online-Seminar

Title: **Karlsruhe data infrastructure for material sciences (KaDI4Mat)**

Abstract

The handling of digital research data and research software plays an increasingly important role in all fields of application of natural and engineering sciences. The reason is especially the growing amount of Data obtained from experiments and simulations [1]. Without suitable methods of analysis, the constantly growing amount of data is no longer be usable meaningfully. This is especially true for materials science, since the research and understanding of new materials is becoming increasingly complex [2]. The extraction of knowledge from these data is considered the fourth paradigm of Science and can be summarized under the keyword data science [3]. An important aspect in order to perform appropriate data analyses smoothly, the structured storage of the data is the key to a smooth research data and the associated metadata. A uniform research data management requires a suitable infrastructure, which allows easy handling of research data and publication in repositories. With such an infrastructure we aim to solve interinstitutional hurdles, easy comparison of theoretical and experimental data and reproducible workflows for data analysis. Researchers can be supported by linking the data with persistent Identifiers to reference them directly in their work. Also the publication of such data, either alone or as a supplement to text publication, is increasingly demanded [4]. Repositories enable the internal as well as the public exchange of research data and can be reused for new investigations. Some of the underlying systems for these repositories can also be installed by the user e.g. for internal use within individual work groups. Additionally, this allows full data sovereignty over stored data.

[1] T. Hey und A. Trefethen. „The data deluge: An e-science perspective“. In: Grid computing: Making the global infrastructure a reality (2003), S. 809–824.

[2] J. Hill, G. Mulholland, K. Persson, R. Seshadri, C. Wolverton und B. Meredig. „Materials science with large-scale data and informatics: unlocking new opportunities“. In: Mrs Bulletin 41.5 (2016), S. 399–409.

[3] T. Hey, S. Tansley, K. Tolle et al. The fourth paradigm: data-intensive scientific discovery. Bd. 1. Microsoft research Redmond, WA, 2009.

[4] L. Naughton und D. Kernohan. „Making sense of journal research data policies“. In: Insights 29.1 (2016).

Alle Interessenten sind herzlich eingeladen.

Prof. Dr.-Ing. Thomas Böhlke
(Sprecher des GRK 2078)