

Molecular and Materials Research: Sharing Data Easily

“Science Data Center for Molecular Materials Research” Develops Digitization Modules for Scientific Data – From Acquisition to Processing to Public Archiving



Sustainable joint use of data from many laboratories is the goal of the MoMaF Science Data Center. (Photo: Laila Tkotz/KIT)

24 hours a day, the Internet offers direct access to the world’s knowledge. In this way, projects profit from the know-how of many bright minds and can be shared with interested persons. Especially researchers handling data strive for free information flow. Exchange of raw data produced in laboratories, however, is prevented by several obstacles. The “Science Data Center for Molecular Materials Research” of Karlsruhe Institute of Technology (KIT) now plans to change this situation in cooperation with the Karlsruhe University of Applied Sciences and FIZ Karlsruhe. For this, funds in the amount of EUR 2.5 million are granted by the Baden-Württemberg Ministry of Science, Research, and the Arts (MWK).

“In Baden-Württemberg, we are establishing an e-science infrastructure that offers to our researchers’ best conditions for applying novel scientific approaches. Data science combines methods of mathematics and computer science with the knowledge on various applications and opens the gate towards new findings,” says Minister Theresia

Monika Landgraf
Chief Press Officer,
Head of Corp. Communications

Kaiserstraße 12
76131 Karlsruhe, Germany
Phone: +49 721 608-21105
Email: presse@kit.edu

Press contact:

Kosta Schinarakis
Press Officer
Phone: +49 721 608-21165
Email: schinarakis@kit.edu

Holger Gust
Karlsruhe University of
Applied Sciences
Phone: 0721 925-1016
Email: pr@hs-karlsruhe.de

Bauer. "For the state of Baden-Württemberg, science-driven data centers, such as that of KIT, are of major importance."

"With the newly established Science Data Center, we ensure that knowledge crosses the borders of institutions and can be used universally," says Professor Holger Hanselka, President of KIT. "The Science Data Center will help us make quicker progress on our joint way to finding solutions for global challenges."

"Research lives on cooperation and exchange. Sharing data, however, is difficult to implement technically," adds Professor Oliver Kraft, KIT Vice President for Research. "Establishment of the Science Data Center for Molecular Materials Research is an important step on our way towards the joint and sustainable use of research data."

Accelerating Research

Data are expensive. Huge numbers of working and computing hours as well as expensive devices and materials are the price for scientific findings. Usually, this price is paid with public funding. Frequently, the value of the data obtained is everlasting and also older data can be used for current studies. New analysis methods can maximize knowledge gain from these data and, hence, their long-term use.

Systematic data protection and sustainable supply of data are major success criteria in science. It is not always easy to guarantee both. Often, efficient tools to exchange data, to structure them in a reproducible way, and to provide them with meta data are lacking. Occasionally, legal framework conditions are not clear. Sometimes, processes fail due to the large data volume that has accumulated over many project years and is to remain available for a long time. The "Science Data Center for Molecular Materials Research," MoMaF for short, now plans to reduce inhibition thresholds and to develop adequate processes and tools for chemists and materials scientists to solve current problems in research data management.

"Joint use of data in materials sciences accelerates national and international research and, hence, innovation in central research areas, such as energy and healthcare," says Professor Britta Nestler, who teaches at KIT and Karlsruhe University of Applied Sciences. Since 2016, she has also worked at KIT's Material Research Center for Energy Systems (MZE). So far, groups working in the areas of molecular chemistry and materials sciences have mostly used individual data management solutions, which leads to reduced availability and visibility of research results. The results of other research areas can

hardly be used for quicker and more comprehensive studies. “A universal, standardized kit for storing, processing, and curating research data or for AI-assisted analysis and interdisciplinary reuse has been lacking so far. The same applies to an institution that pools competencies from various disciplines and makes them usable by everybody.”

Efficient Research Data Management

“With MoMaF, we will develop modules for digitization, which will cover all phases from the generation of data to their sustainable archiving,” says Professor Stefan Bräse of KIT’s Institute of Organic Chemistry, who also works at MZE. This contribution to digitization ensures that data about molecules and their interactions for the description of materials can be stored, such that they are findable, accessible, interoperable, and reusable. This is referred to as the FAIR data principle. Not only discipline-specific and interdisciplinary research results are stored, but also the processing and analysis methods applied for a better understanding. The goal consists in supplying a software infrastructure to meet general and specific requirements relating to scientific data protection and efficient reuse. “MoMaF will provide key elements for research data management, which have not yet been available on national or international levels.”

MoMaF will be based on a concept already established at KIT, an electronic laboratory journal (ELN, Electronic Lab Notebook) with the connected research data repository (publicly accessible data archive) Chemotion for organic chemistry. The ELN offers functions for subject-specific acquisition, organization, processing, and interconnection of research data. These functions are the basis of a structured storage and use of the data obtained or for reuse by other researchers. Direct transfer of the scientific data obtained to the Chemotion research data repository enables meta data generation and automatic registration of unambiguous persistent identifiers (PID) for links to external, subject-specific databases. Chemotion is an exemplary software worldwide and was granted the 2017 SPARC-Europe Open Data Champion Award. Now, the source code of ELN and the repository will be extended by appropriate modules for use in the neighboring disciplines of molecular chemistry, macromolecular chemistry, surface chemistry, and virtual development of materials. In addition, a recommendation service will be implemented as a software system. With the help of machine learning methods, work on the organization and analysis levels will be supported and recommendations will be given with respect to data collection, storage, curation, and reuse.

Top-level Research

Participation of the computing centers of KIT and Karlsruhe University of Applied Sciences and of the KIT Library in MoMaF ensures integration of the Science Data Center in the partners' research and teaching structures and bridges the gap to already established research data services of KIT. The Steinbuch Centre for Computing of KIT operates the computing center and contributes long-standing expertise in handling big scientific data volumes as gained in the KIT projects of GridKa, LSDMA, LSDF, and bwDataArchiv. This expertise is complemented by the know-how of KIT's Institute of Applied Informatics and Formal Description Methods.

MoMaF plans to contribute to national and international initiatives, e.g. to the establishment and support of research data infrastructures, such as the German National Research Data Infrastructure (NFDI) and the European Research Data Alliance. KIT's Clusters of Excellence "3D Matter Made to Order" and "Energy Storage beyond Lithium" will be among the first users of the tools of MoMaF. Moreover, MoMaF is to cover the needs of other research alliances, e.g. of CRC1176 and CRC/TRR88, in which MoMaF researchers are involved. In the long term, MoMaF is to be used as a research instrument in the Helmholtz programs.

The "Science Data Center for Molecular Materials Research" will be operated at and by KIT. The software developed will be made available to the broad scientific community as an open source. The project is coordinated by KIT, the partners are Karlsruhe University of Applied Sciences and FIZ Karlsruhe – Leibniz Institute for Information Infrastructure. For evaluating use at various university locations, the infrastructure for operating the ELNs will also be established at Karlsruhe University of Applied Sciences. FIZ Karlsruhe will analyze legal aspects of the development and implementation of digital modules.

MoMaF will enable top-level research in Baden-Württemberg and, hence, guarantee competitiveness. The setup phase will take four years. The budget of about EUR 3.5 million consists of a KIT share of about 1 million euros and funds in the amount of EUR 2.5 million granted by the Baden-Württemberg Ministry of Science, Research, and the Arts.

Further Information:

KIT research data management: <https://www.rdm.kit.edu>
(in German only)

MZE: https://www.kit.edu/kit/english/pi_2016_161_kit-material-research-center-opened.php

Research Data Alliance: https://www.sek.kit.edu/kit_ex-press_4016.php (in German only)

ELN Chemotion: <https://chemotion.net/>

<https://openscholarchampions.eu/opendata/champion/datajournalsandrepositoriesgotogether/>

NFDI4Ing and NFDI4Chem: <https://www.tib.eu/de/service/aktuelles/detail/nationale-forschungsdateninfrastruktur-fuer-die-chemie-nfdi4chem/>



Being “The Research University in the Helmholtz Association“, KIT creates and imparts knowledge for the society and the environment. It is the objective to make significant contributions to the global challenges in the fields of energy, mobility and information. For this, about 9,300 employees cooperate in a broad range of disciplines in natural sciences, engineering sciences, economics, and the humanities and social sciences. KIT prepares its 25,100 students for responsible tasks in society, industry, and science by offering research-based study programs. Innovation efforts at KIT build a bridge between important scientific findings and their application for the benefit of society, economic prosperity, and the preservation of our natural basis of life.

This press release is available on the internet at http://www.sek.kit.edu/english/press_office.php.

The photo in the best quality available to us may be downloaded under www.kit.edu or requested by mail to presse@kit.edu or phone +49 721 608-21105. The photo may be used in the context given above exclusively.

This year's **anniversary logo** recalls the milestones reached by KIT and its long tradition in research, teaching, and innovation. On October 1, 2009, KIT was established by the merger of its two predecessor institutions: the Polytechnic School and later University of Karlsruhe was founded in 1825, the Nuclear Reactor Construction and Operation Company and later Karlsruhe Research Center in 1956.