NHR - Graduate School: Research Fields 2026

Table of Contents

NHR@FAU – Friedrich Alexander University Erlangen-Nürnberg

1#FAU: Improving Scalable Solvers by AI 3

2#FAU: Performance Engineering from Single Core to Highly Parallel Architecture 4

3#FAU: Molecular Simulations of Bio-Molecular Complexes 5

4#FAU: Development of new electronic structure methods combining high accuracy with computational efficiency 6

5#FAU: High-throughput Simulations for a Mechanistic Understanding of Crystallization Processes 7

6#FAU: Learning simple models on sparse data for materials design 8

7#FAU: Foundation Models in Medicine: Advancing Multimodal Learning for Medical Big Data 9

8#FAU: Cognitive Computational Neuroscience for Enhanced Speech and language understanding 10

9#FAU: Multimodal Models for Co-Speech Gesture Analysis and for the Understanding of Multimodal Communication 11

10#FAU: Gestural Patterns in Media Broadcasts – Development of Big Data Methods 12

GWDG – NHR@GÖTTINGEN – Georg August University of Göttingen

1#Göttingen: High-Performance Storage 13

2#Göttingen: Large Language Models in HPC 14

3#Göttingen: Data-Center Twin 14

4#Göttingen: Scalable Big Data Analytics 14

NHR@KIT – Karlsruhe Institute of Technology

1#KIT: High-performance nanophotonic scattering simulations 16

2#KIT: Machine learning in Earth system models 17

3#KIT: Accelerating Supercomputing Simulations of Complex Flows 18

NHR@TUD – TUD Dresden University of Technology

1#TUD: Software Tools for Parallel Performance Analysis and Optimization 19

2#TUD: HPC Workflows for Big Life Science Data 20

3#TUD: High Performance Storage Systems 21

4#TUD: Machine Learning and Training Data Preparation 22

NHR@SW – Goethe University Frankfurt, Johannes Gutenberg University of Mainz, University of Kaiserslautern-Landau, Saarland University

1#SW: GPU-Bio: Accelerated Algorithms and Machine Learning for Modern Bioinformatics on GPU-accelerated HPC Systems 23

2#SW: Concurrent Data Structures & Algorithms on Modern GPU Systems for AI 25

3#SW: Algorithmic Differentiation in HPC Environments 26

4#SW: LLM-based Compiler Generation for Quantum Computers 27

5#SW: In-network Processing for HPC Applications 28

6#SW: Modular and Automated Workload Analysis for HPC Systems 29

NHR@ZIB – Zuse Institute Berlin

1#ZIB: Sparsity-Driven Optimization for Large Language Models under Resources Constraints 30

2#ZIB: AI-enhanced Molecular Dynamics of Fluid-Solid Interfaces 31

3#ZIB: Portable and Energy-Efficient Computing on Data-Parallel Platforms 33

4#ZIB: Semantic I/O with Object Stores in HPC 34

5#ZIB: Modelling Excitations in Novel Materials 35

6#ZIB: HPC-based Simulation of Agent-Based Models and Complex Networks 36

7#ZIB: AI-guided high -throughput drug design 37

PC2 – Paderborn University

1#PC2: Empirical research on HPC Systems for Optimization of Operation 38

2#PC2: FPGA-Acceleration for High-Performance Computing and AI 40

3#PC2: GPU Performance and Portability 41

4#PC2: Compilation for Parallel Computer Architectures and Accelerators 42

NHR4CES – RWTH Aachen University and Technical University Darmstadt

1#NHR4CES: Building Foundation Models for Operational Processes 44

2#NHR4CES: Modeling metastable phase formation for physical vapor deposition thin films 46

3#NHR4CES: Computational Neuroscience 47

4#NHR4CES: HPC-Driven Low-Cost Magnetic Resonance Imaging 48

5#NHR4CES: Computational 4D Metabolic Imaging through Kinetic Modelling 49

6#NHR4CES: Digital Twins 51

7#NHR4CES: Efficient Scientific Visualization on HPC Infrastructure 52

8#NHR4CES: LLM-based Data Wrangling for HPC Applications 53

9#NHR4CES: Integrating Vector Search into HPC Infrastructure 54

NHR@FAU – Friedrich Alexander University Erlangen-Nürnberg

The Erlangen National High-Performance Computing Center (NHR@FAU) is a central institution of the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU). Original focus topics in applications and HPC research of NHR@FAU are:

- performance engineering
- scalable solvers and atomistic structure simulations in chemistry life science, materials science
- physics
- artificial intelligence
- machine learning techniques.

1#FAU: Improving Scalable Solvers by AI

Description:

In many HPC applications, solving sparse linear systems is the fundamental performance bottleneck. Only a few algorithmic principles, such as, e.g., multigrid, are available to construct scalable solvers. Further acceleration is possible with techniques that avoid storing the system matrix. Such matrix-free techniques are one focus of research at NHR@FAU. These methods require that the solver is co-designed with the discretization method. But since the optimal choice of components and parameters for discretization and solvers are highly problem-dependent, it is difficult to find them manually for more complex matrices. Therefore, surrogate models e.g. based on physically inspired neural nets (PINNs) or genetic programming are used to improve the numerical methods. However, these data-driven methods can have high computational costs in the training phase.

To address this issue, another central component in building highly efficient and easily usable implementations in this co-design process are modern code-generation (meta-programming) techniques which form another research focus of NHR@FAU.

Link to working group:

- For further information, please have a look here: https://hpc.fau.de.

Professional prerequisites for applicants:

- Very good programming skills (e.g. C++, CUDA, Python).
- Basic AI knowledge, numerical methods.

Working language/required language skills:

- English
- German.

Description of supervisor and scientific institution:

- Prof. Harald Köstler, FAU University, Chair of Computer Science 10 (System Simulation) & Head of Research at NHR@FAU.

2#FAU: Performance Engineering from Single Core to Highly Parallel Architectures

Description:

Performance Engineering (PE) is a structured, performance-model-based process for the structured optimization and parallelization of basic operations, algorithms and application codes for modern compute architectures. Performance models and performance tools are the core components of the PE process. The group uses and further develops the execution cache memory (ECM) model, which is a generalization of the well-known Roofline model and allows predictions of single-core performance and scaling within a multi-core chip. The interaction of node-level bottlenecks with highly parallel code execution is another recent area of research. Furthermore, simple-to-use performance analysis and modeling tools such as LIKWID or OSACA are developed at NHR@FAU. Hardware-efficient building blocks for sparse linear algebra and stencil solvers are strong topical focus of the PE activities of the center.

Working group and link(s) to working group:

- For further information, please have a look here: ttps://hpc.fau.de/person/gerhard-wellein/.

Professional prerequisites for applicants:

- no information.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Gerhard Wellein, FAU University, Department of Computer Science, Professorship for High-Performance Computing & Head of NHR@FAU.

Potential NHR Center for Ph.D. thesis and/or secondment collaboration:

- PC2 (Paderborn)
- NHR4CES (Aachen).

3#FAU: Molecular Simulations of Bio-Molecular Complexes

Description:

Bio-molecular complexes such as protein-protein, protein-DNA, or (other) protein ligand complexes are flexible entities whose conformational dynamics are important for their function. By molecular dynamics simulations, using high-performance computing architectures, these dynamics can be explored in atomic detail. However, depending on the process to be investigated, advanced methods such as enhanced sampling techniques or approaches with hybrid quantum mechanical and classical (QM/MM) potentials are needed, often also in combination.

We are using a catalogue of molecular simulation techniques and data analysis methods to study bio-molecular complexes involved in important processes such as DNA repair, enzymatic reaction pathways, ligand-induced signaling, or immune processes at the plasma membrane.

Working group(s) and link(s) to working group(s):

- For further information, please have a look here: https://www.biomemphys.nat.fau.de/.

Professional prerequisites for applicants:

- Experience in molecular simulations.
- Solid background in statistical mechanics.
- Coding skills in python or another high-level programming language, familiar with linux.

Working language/required language skills:

- English.

Description of supervisors and scientific institutions + links:

- Prof. Petra Imhof, FAU University, Computer Chemistry Center, Technical Director.
- Prof. Heinrich Sticht, FAU University, Professorship of Bioinformatics.
- Prof. Rainer Böckmann, FAU University, Professorship for Computational Biology.

Potential NHR Center for Ph.D. thesis and/or secondment collaboration:

- NHR@ZIB (Berlin)
- PC2 (Paderborn)
- NHR@TUD (Dresden).

4#FAU: Development of new electronic structure methods combining high accuracy with computational efficiency

Description:

The workhorse for electronic structure calculations in chemistry, physics, and materials science are methods based on density-functional theory (DFT). Despite their enormous success in practical application, presently available DFT methods suffer from several shortcomings. Their accuracy does not reliably reach what is called chemical accuracy, i.e., an accuracy of 1 kcal/mol in reaction energies, and non-covalent interaction cannot be treated. A promising new type of DFT method relies on the adiabatic-connection fluctuation-dissipation (ACFD) theorem. ACFD methods exhibit unprecedented accuracy and applicability. These methods need to be implemented in highly efficient codes making full use of the opportunities high-performance computing offers.

Working group and link(s) to working group:

- Group of Prof. Andreas Görling, FAU University, Department of Chemistry and Pharmacy, Chair of Theoretical Chemistry.

Professional prerequisites for applicants:

- Good basic knowledge of quantum mechanics.
- programming skills.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Andreas Görling, FAU University, Department of Chemistry and Pharmacy, Chair of Theoretical Chemistry.

Potential NHR Center for Ph.D. thesis and/or secondment collaboration:

- PC2 (Paderborn)
- NHR@TUD (Dresden).

5#FAU: High-throughput Simulations for a Mechanistic Understanding of Crystallization Processes

Description:

The formation of materials structure can be rationalized, to first approximation, from geometric considerations. But while crystal structure is generally well understood, dynamic and kinetic aspects of crystallization are much less accessible. The prevailing strategy in computational materials science is to simulate with ever higher accuracy. Here, we propose to investigate crystallization processes systematically across parameter spaces and aim at a mechanistic understanding that correlates the geometry of matter (structure) and interaction (chemistry) to details of ordering pathways at full atomic resolution. This goal is accomplished by establishing a crystallographic structure analysis and simulation infrastructure that rivals experimental synthesis and structure solution from scattering data. Analysis of trajectory data addresses long-standing open questions in the context of multi-step nucleation and directional growth. Applications are phase diagrams of intermetallics, industrial crystal growth, nucleation and growth of nanocrystals, defect formation and propagation, and fractional crystallization in geology.

Working group and link(s) to working group:

- For further information, please have a look here: https://engellab.de.

Professional prerequisites for applicants:

- Master's degree in physics, Chemical Engineering, Materials Science or a related field.
- Experience in simulation and computer programming (C/C++, Python) is desirable.
- An interest in mathematics and physical sciences is required.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Michael Engel, FAU University, Department of Chemical and Biological Engineering, Professorship for Modeling of Self-Organization Processes.

6#FAU: Learning simple models on sparse data for materials design

Description:

The research of the Data-based-Materials-Design group orbits around the development and application of AI methods for the modeling, design, and discovery of materials. Methods span compressed sensing, symbolic regression, subgroup discovery, and interpretable deep learning, while applications cover catalysis, characterization of surfaces, and several mechanical and transport properties of nano- as well as macro-scaled materials. Focus is on methods that yield interpretable models and can cope with "small data" for training. The group participates in the NFDI FAIRmat consortium, in coordinating the creation of metadata and ontologies for the Interoperability and Reusability of FAIR (materials science) data. Currently, the group is particularly interested in developing information-theory based, reliable methods for out-of-distribution detection, design workflows for materials exploration, discovery, and design that make the most efficient use of the few accurate data points available from experiments and high-level computations.

Working group and link(s) to working group:

- Group of Data-based materials design. For further information, please have a look here: https://www.iznf.fau.de/institute/data-based-materials-design/.

Professional prerequisites for applicants:

- Master's in physics or chemistry or materials science or similar.
- Documented activity in data science (e.g., active github page).

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Luca M. Ghiringhelli, FAU University, Department of Materials Science and Engineering, Chair for Materials Simulation.

Potential NHR Center for Ph.D. thesis and/or secondment collaboration:

- NHR@ZIB (Berlin).

7#FAU: Foundation Models in Medicine: Advancing Multimodal Learning for Medical Big Data

Description:

The exponential growth of medical data—from MRI scans and ophthalmic imaging to patient speech recordings and time series data—presents both an opportunity and a challenge. Traditional machine learning models often struggle to integrate and interpret this diverse, multimodal information. We propose developing advanced foundation models tailored for medicine that leverage generative AI, image fusion, and known operator learning. By systematically integrating data across various modalities, these models aim to enhance image analysis and understand medical data in clinical settings. This project addresses critical questions in learning for medical big data and seeks to revolutionize how we interpret and utilize complex medical information for better diagnostics and patient care.

Working group and link(s) to working group:

- For further information, please have a look here: https://lme.tf.fau.de/.

Professional prerequisites for applicants:

- Experience in Computer Science with focus on Machine Learning and Pattern Recognition.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Andreas Maier, FAU University, Department of Computer Sciences, Chair for Pattern Recognition.

Potential NHR Center for Ph.D. thesis and/or secondment collaboration:

- NHR@TUD (Dresden).

8#FAU: Cognitive Computational Neuroscience for Enhanced Speech and Language Understanding

Description:

Despite significant advancements in speech processing and language understanding, machines still lag humans in comprehending and interpreting nuanced language patterns. This project aims to bridge that gap by integrating cognitive computational neuroscience with cutting-edge machine learning techniques. Utilizing MRI data and ECG time series data, we will model the neural mechanisms underlying speech and language processing. By applying knowledge about the human brain, we strive to develop systems that not only process speech and language but also understand context and intent at a cognitive level. This interdisciplinary approach promises breakthroughs in neuroscience as well as natural language and speech processing, enabling applications ranging from more intuitive virtual assistants to advanced neurological diagnostics.

Working group and link(s) to working group:

- For further information, please have a look here: https://lme.tf.fau.de/.

Professional prerequisites for applicants:

- Experience in Computer Science with focus on Machine Learning and Pattern Recognition.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Andreas Maier, FAU University, Department of Computer Sciences, Chair for Pattern Recognition.

Potential NHR Center for Ph.D. thesis and/or secondment collaboration:

- NHR@TUD (Dresden).

NHR GRADUATE SCHOOL

9#FAU: Multimodal Models for Co-Speech Gesture Analysis and for the Understanding of Multimodal Communication

Description:

Human communication is inherently multimodal, i.e. auditory and visual information support the understanding of a message. While significant progress has been made in multimodal gesture analysis using laboratory data, the availability of annotated datasets remains limited due to complex annotation schemes. This limitation is particularly pronounced for co-speech gesture, which requires tedious and meticulous annotation, resulting in very small datasets. This project aims to develop a pre-trained model that can be fine-tuned on limited data for specific downstream tasks, helping practitioners to automate gesture analysis. To achieve this, we aim to pre-train on a very large dataset of media recordings. This approach requires non-trivial data filtering and de-duplication methods. In addition, co-speech gestures and the corresponding speech signals are not always synchronized, which requires advanced modelling of inter-modality interaction.

Working group and link(s) to working group:

- For further information, please have a look here: https://www.angam.phil.fau.de/fields/engling/big-data-linguistics/.

Professional prerequisites for applicants:

- Master's in computer science, Artificial Intelligence, or a related field.
- Expertise in machine learning, in particular transformer-based architecture.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Peter Uhrig, FAU University, Department of English and American Studies, Chair of Big Data Linguistics.

Potential NHR Center for Ph.D. thesis and/or secondment collaboration:

- NHR@TUD (Dresden).

10#FAU: Gestural Patterns in Media Broadcasts – Development of Big Data Methods

Description:

The Chair of Big Data Linguistics at FAU holds more than 300,000 hours of video data and the corresponding annotations out of OpenPose at NHR@FAU. The aim of this project is to develop methods to systematically extract patterns of co-speech gesture from the dataset. This includes identifying gesture similarities across a wide range of camera perspectives and of gestural variation, including combinations of hand and finger positions, trajectories, handshapes, movement directions, speed, any many other factors relevant to the description of co-speech gesture. The applicant can focus on gesture retrieval (e.g. through formal specification of gesture features or query-by-example) or the data-driven extraction of recurrent patterns.

Working group and link(s) to working group:

- For further information, please have a look here: https://www.angam.phil.fau.de/fields/engling/big-data-linguistics/.

Professional prerequisites for applicants:

- Master's in data science, Computational Linguistics Computer Science, or a related field.
- Ideally familiarity with statistical methods such as time series analysis.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Peter Uhrig, FAU University, Department of English and American Studies, Chair of Big Data Linguistics.

Potential NHR Center for Ph.D. thesis and/or secondment collaboration:

- NHR@TUD (Dresden).

GWDG – NHR@GÖTTINGEN – Georg August University of Göttingen

The Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen (GWDG) is a joint computing center and research facility of the Max Planck Society and the Göttingen University. It provides infrastructures for data management and analysis, collaboration, and computing for various national projects.

This national high-performance data center with the Emmy supercomputer is a tier 2 center and offers advice and support to a diverse portfolio of users and is itself involved in research in application-related projects.

The NHR centre NHR@Göttingen is part of the National High-Performance Computing (NHR) and a member of the NHR network.

1#Göttingen: High-Performance Storage

Description:

Efficient data storage solutions are critical for advancing data-intensive scientific research and large-scale data analysis workflows, such as those used in machine learning. Traditional storage systems, including file systems and object storage, rely on a deep understanding of both system architecture and application-specific workloads to optimize performance.

In this thesis, we conduct a systematic evaluation of storage system performance characteristics, develop predictive models to capture workload behavior, and design novel optimization strategies to enhance storage efficiency. These strategies span multiple layers of the I/O stack—from interface-level protocols to kernel modules—enabling targeted improvements across the storage hierarchy.

By addressing bottlenecks at various abstraction levels, this work aims to bridge the gap between theoretical performance models and real-world system implementation. Candidates with a strong foundation in operating systems, distributed systems, or performance analysis are encouraged to contribute to this research, as these skills will be instrumental in advancing the project's technical goals.

Supervisor & Location:

- Prof. Dr. Julian Kunkel, Göttingen University, Professor in High-Performance Computing & Deputy Head High-Performance Computing at GWDG.

2#Göttingen: Large Language Models in HPC

Description:

In recent years, LLMs such as ChatGPT-4 have revolutionized the field of natural language processing (NLP) by demonstrating exceptional performance in various tasks, including text generation, sentiment analysis, and machine translation. They bear chances to be utilized for various tasks in HPC as well, for instance, for logfile analysis, for analysing system faults, debugging and performance analysis of parallel applications, and for providing a rich variety of user support. Recognizing the untapped potential of LLMs in High-Performance Computing (HPC) environments, this PhD project aims to explore their application in enhancing HPC operations and user experience through innovative problem-solving techniques. Additionally, platforms to deploy inference models efficiently in an HPC center are explored.

Supervisor & Location:

- Prof. Dr. Julian Kunkel, Göttingen University, Professor in High-Performance Computing & Deputy Head High-Performance Computing at GWDG.

3#Göttingen: Data-Center Twin

Description:

Data centers contain various infrastructures to, e.g., steer airflow and cooling of housed systems. Understanding and optimizing the infrastructure but also the whole life cycle of IT components is important to improve efficiency and make computing greener. The modelling and simulation of the technical processes and relevant characteristics allow investigation of the impact of various tuning options and enable what-if analysis to optimize planning in each data center but also in future data centers.

Supervisor & Location:

- Prof. Dr. Julian Kunkel, Göttingen University, Professor in High-Performance Computing & Deputy Head High-Performance Computing at GWDG.

4#Göttingen: Scalable Big Data Analytics

Description:

Machine learning methods are omnipresent in our daily lives and are also increasingly used in scientific domains such as life science and the digital humanities. However, in the data center, HPC workflows could still benefit more from opportunities for machine learning and deep learning methods. As part of this topic, we research:

- How scalable data analytics workflows can be established for relevant scientific domains.
- How existing workflows can be augmented using deep learning, machine

learning, and High-Performance Data Analytics, in general.

Supervisor & Location:

- Prof. Dr. Julian Kunkel, Göttingen University, Professor in High-Performance Computing & Deputy Head High-Performance Computing at GWDG.

NHR@KIT – Karlsruhe Institute of Technology

Karlsruhe Institute of Technology (KIT) is a university of the state of Baden-Württemberg as well as a research center in the Helmholtz Association and creates synergies from serving both purposes. The NHR centre NHR@KIT addresses four user communities: (1) earth system science, (2), materials science, (3) engineering for energy and mobility research, and (4) particle physics. Core HPC method competencies are data-intensive computing, numerical algorithms, and sustainable software development. In HPC operations, NHR@KIT focuses on secure authentication and authorization.

1#KIT: High-performance nanophotonic scattering simulations

Description:

The scattering of light in nanophotonic systems can be conveniently formulated on algebraic grounds. The optical properties of the scatterers are expressed through a T-matrix, and the interaction with light is represented as a matrix-vector product. Building upon this theoretical framework, we have developed an open-source code called teams that can flexibly handle various nanophotonic systems. NHR@KIT aims to advance this computational framework further. Currently, we seek to enable data-driven research approaches by combining the code with a published repository of T-matrices, enhancing the ability to handle the interaction of an extremely large number of scatterers, extending the code to allow for the solution of inverse problems using automatic differentiation, and integrating machine learning methods to improve the capability of the framework. To work on one or more of these tasks, we are looking for a dedicated doctoral student.

Working group and link(s) to working group:

- For further information, please have a look here: https://www.tfp.kit.edu/rockstuhl.php.

Professional prerequisites for applicants:

- Master's in physics or Optics/Photonics.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Carsten Rockstuhl, INT and TFP at Karlsruhe Institute of Technology, Theoretical Nanooptics.
- Dr. Ivan Kondov, Karlsruhe Institute of Technology, Department of Scientific Computing & Mathematics, Scientist in the Simulation and Data Lab Materials Science.

Potential NHR Center for Ph.D. thesis and/or secondment collaboration:

- PC2 (Paderborn), Group of Prof. Jens Förstner.

2#KIT: Machine learning in Earth system models

Description:

Classical modelling has long been the standard method for long- and short-term predictions in climate science, and models such as the KIT-developed ICON-ART rank among the power users of large HPC systems.

With the rapid developments in hardware and the rise of the Exascale era, scientists can process ever-growing amounts of data and refine their simulations to improve results. The downsides of this evolution are long computing times and significant power consumption. However, in recent years scientists pursued a new strategy by adapting large neural networks and using the improvements in machine learning and artificial intelligence. This ansatz yielded encouraging results while reducing both compute time and power consumption. The goal of this research work is to combine both classical modelling and cutting-edge AI techniques.

Description of supervisor and scientific institution (depending on the topic of the doctoral thesis):

- Prof. Dr. Martin Frank, Karlsruhe Institute of Technology, Director Scientific Computing Center (SCC) & Professor for Computational Science and Mathematical Methods, Department of Mathematics.
- Dr. Ole Kirner, Karlsruhe Institute of Technology, Department of Scientific Computing & Mathematics, Scientist in the Simulation and Data Lab Earth System Science.
- Dr. Jasmin Hörter, Karlsruhe Institute of Technology, Head of the Department of Scientific Computing & Mathematics.
- Dr. Jörg Meyer, Karlsruhe Institute of Technology, Head of the Department of Data Analytics, Access, and Applications.

3#KIT: Accelerating Supercomputing Simulations of Complex Flows

Description:

Environmental pollution and climate change related to energy conversion technology are among the most pressing challenges of mankind. A substantial element of many processes in this field is turbulent reacting multiphase flow of solid particles or liquid droplets in a gas phase environment. Two recent research examples are liquid spray atomization processes for gasification and the conversion of metal fuel particles in carbon-free energy cycles.

Such processes can be optimized by means of Computational Fluid Dynamics (CFD) and require High-Performance Computing (HPC) systems to tackle flows of industrial- or fundamental research relevance.

However, the inherent inhomogeneities in turbulent reacting multiphase flows pose substantial challenges to common domain-decomposition (MPI-based) HPC approaches, where large areas of the computational domain do not require much computational power, while certain regions with many droplets/particles or major chemical activity are responsible for the bulk of the computational cost.

The target is to develop efficient algorithms for improved HPC performance, which can be achieved by enhancing various levels of the numerical process. This includes optimizing dynamic load balancing and integrating GPU-capable libraries to accelerate particle tracking and linear solver computations. Applying performance engineering tools to analyze hot spots, developing new solvers, and incorporating artificial intelligence into existing CFD solvers are some of the challenges in this complex yet exciting multi-disciplinary work.

Description of supervisor and scientific institution:

- Prof. Oliver T. Stein, Karlsruhe Institute of Technology, Professor in Simulation of Reacting Thermo-Fluid Systems, Sub-Division of Combustion Technology at the Engler-Bunte Institute.
- Dr.-Ing. Jordan Denev, Karlsruhe Institute of Technology, Department of Scientific Computing & Mathematics, Scientist in the Simulation and Data Lab Engineering for Energy and Mobility.

Professional prerequisites for applicants:

- MSc in Mechanical Engineering, Chemical and Process Engineering, Physics or similar study programs.

GRADUATE SCHOOL

NHR@TUD – TUD Dresden University of Technology

The TU Dresden is one of the largest technical universities in Germany. It has a long history of running HPC installations and is specialized on data intensive HPC as well as data analytics and machine learning. Its NHR center's focus topics are: (1) Methods for Big Data and data analysis as well as management, (2) Machine Learning, (3) Tiered storage architectures and I/O optimization, and (4) Performance and energy efficiency analysis and optimization. The application area focus topics are: (5) life sciences and (6) earth system science.

The NHR center at TU Dresden cooperates closely with ScaDS.AI Dresden/Leipzig (Center for Scalable Data Analytics and Artificial Intelligence).

1#TUD: Software Tools for Parallel Performance Analysis and Optimization

Description:

Parallel performance, efficiency, and scalability are primary concerns for the HPC community. Dedicated software tools for performance analysis and optimization are indispensable to cope with several challenges for HPC programmers and cluster operators. There are several performance aspects waiting for new software tool approaches for measuring, analyzing and steering parallel performance, among others: new heterogeneous hardware architectures; new and evolving parallel programming models; more and more complex application codes; Big Data and Machine learning frameworks in HPC environments; AI analysis of monitoring data; as well as high scalability and low overhead of the tools themselves. This topic offers the opportunity to join an established performance tools research group working on well-known software tools renowned in the HPC community.

Working group and link(s) to working group:

- Performance Engineering Group, Dr. Holger Brunst.

Professional prerequisites for applicants:

- MSc in Computer Science or closely related fields with very good grades.
- Experience with HPC, parallel programming, and performance optimization is beneficial.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Wolfgang E. Nagel, Head of Information Services and High-Performance Computing (ZIH) and NHR@TUD.

2#TUD: HPC Workflows for Big Life Science Data

Description:

Single cell omics and live imaging of developing embryos and organoid cultures generate big data in the life sciences. We need to analyze these big life science data to crack the code of biological self-organization and morphogenesis, as started by Alan Turing with his work on "The Chemical Basis of Morphogenesis". This approach integrates HPC workflows, mathematical modeling and multi-scale simulations with interdisciplinary thinking to develop and apply novel algorithms. For example, the hybrid simulation of a developing organism combines a mechanistic model of a subsystem of interest, like cell- to-cell signaling and gene regulatory networks, with a processed stream of measurement data for the other subsystems, like the segmentation, tracking and interpolation of 3D time lapse microscopy data of the cell positions, shapes and contacts in a growing and deforming embryo. The results of such hybrid simulations can then be compared to measurement data for the entire system to perform robust parameter estimation and model selection for the subsystem of interest. Such novel HPC workflows will be widely applicable across life sciences and beyond.

Working group and link(s) to working group:

- Innovative Methods of Computing.

Professional prerequisites for applicants:

- MSc in Computer Science, Mathematics, Physics, or closely related fields with very good grades.
- Experience in HPC and life sciences is beneficial.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Dr. Andreas Deutsch, Head of Unit Innovative Methods of Computing (IMC) at Center for Information Services and High-Performance Computing (ZIH).

3#TUD: High Performance Storage Systems

Description:

Work on new concepts and implementations for highly efficient data workflows and storage systems that connect HPC and AI. This will combine methods from different scopes which offer mutual advantages. For example, combine HPC technologies like Infiniband, RDMA, NVMEoF, GPU direct and others with different storage options. It involves managing workflow approaches compatible with the respective HPC and AI environment. It may also focus on data and metadata management integration in data- driven workflows. Finally, it may concern performance and scalability improvements for the most demanding processing stages. Prototyping high end storage systems for data processing and realizing demanding workflows might be part of the assignment.

Working group and link(s) to working group:

- Data and Storage Architectures Group, Dr. Michael Kluge.

Professional prerequisites for applicants:

- MSc in Computer Science or closely related fields with very good grades.
- Experience with HPC and (parallel) file systems is beneficial.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Wolfgang E. Nagel, Head of Information Services and High-Performance Computing (ZIH) and NHR@TUD.

4#TUD: Machine Learning and Training Data Preparation

Description:

In many cases, rapidly increasing data generation rates require constant updates of trained machine and deep learning models at a faster rate, making it vital to pre-process the incoming data and update the model in close time. This requires adapting systems and models to handle the data in the form of data streams using online/active learning. Often the data is generated in real-time at external experiments, requiring immediate analysis, especially for mission critical applications, e.g., in medical scenarios or safety domain. Modern HPC infrastructures can play a vital role in achieving this requirement. Here, the techniques for data reduction, data transfer and data pre-processing with efficient use of resources are an ongoing area of research. Furthermore, current research involves investigation into scaling of these methods using distributed computing frameworks w.r.t. different loads, storage systems, computing systems, IO, energy consumption, etc., while aiming for low latency and high throughput computation. In addition to exploring different data quality and data integration techniques, the current work could also involve exploring secure and privacy-aware processing techniques for highly sensitive data. Within the national AI competence center ScaDS.AI Dresden/Leipzig we are involved in research on efficient distributed learning approaches and performance optimization of machine learning applications on scalable HPC infrastructures and on alternative computing architectures for the future.

Working group and link(s) to working group:

- ScaDS.AI Dresden/Leipzig - Center for Scalable Data Analytics and Artificial Intelligence, Dresden site.

Professional prerequisites for applicants:

- MSc in Computer Science, Data Science, or closely related fields with very good grades.
- Experience with HPC and machine learning is beneficial.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Wolfgang E. Nagel, Head of Information Services and High-Performance Computing (ZIH), NHR@TUD, and Director ScaDS.AI.
- Dr. Rene Jäkel, Administrative Director ScadS.AI.

GRADUATE SCHOOL

NHR@SW – Goethe University Frankfurt, Johannes Gutenberg University Mainz, University of Kaiserslautern-Landau, Saarland University

NHR South-West (NHR@SW) is a collaboration of Goethe University Frankfurt, Johannes Gutenberg University Mainz (JGU Mainz), University of Kaiserslautern-Landau (RPTU) and Saarland University. Research activities at NHR@SW focus on performance engineering, artificial intelligence, machine learning, quantum computing, life sciences, and physics.

1#SW: GPU-Bio: Accelerated Algorithms and Machine Learning for Modern Bioinformatics on GPU-accelerated HPC Systems

Description:

Recent years have seen a tremendous increase in the volume of data generated in life sciences. The abundance of generated data allows us to address a plethora of questions that could not be answered previously with numerous applications in precision and personalized medicine as well as drug discovery. However, the analysis of large-scale data sets poses difficult computational challenges and is thus a highly active field of research. For example, as protein sequence database growth outpaces Moore's law, and increased use of advanced machine learning (ML) applications such as AlphaFold3 require increasingly more multiple sequence alignments (MSAs) to be generated. Thus, corresponding MSA generating tools need to be innovative quickly enough to cope with increased complexity.

The explosive growth of life science datasets and the rapid evolution of sequencing technologies thus motivates the development of algorithms and tools that can scale towards large-scale input datasets. This in turn requires the efficient exploitation of modern high-performance computing (HPC) systems that are moving towards increased levels of parallelism by means of many-core architectures such as GPUs with thousands of cores per chip.

The goal of this project is the design and implementation of highly efficient massively parallel algorithms optimized for CUDA-enabled GPUs with associated software tools for relevant analysis tasks in life sciences. This project will comprehensively tackle efficiency, sensitivity, and precision bottlenecks in a variety of applications such as protein structure prediction and metagenomics.

Our approach will be based on developing libraries of key algorithms in a massively parallel way. They will then be used in several applications to gain significant speedups compared to current approaches while scaling large data sets. Thus, this project will boost the downstream analysis of sequencing and structural datasets for many important tasks in ML-based bioinformatics.

Working group and link(s) to working group:

- High Performance Computing Research Group, Johannes Gutenberg University Mainz.

Professional prerequisites for applicants:

- Very good programming skills (e.g. C++, CUDA, Python).
- Basic knowledge of GPUs, HPC, algorithms & data structures and bioinformatics.

Working language/required language skills:

- English.
- German.

Description of supervisor and scientific institution:

- Prof. Dr. Bertil Schmidt, JGU Mainz, Chair of High-Performance Computing Group.

2#SW: Concurrent Data Structures & Algorithms on Modern GPU Systems for AI

Description:

Data structures are fundamental to many algorithmic problems in computer science such as answering various types of database queries and performing fast string processing. The continually increasing volume of data and usage of AI systems has resulted in a growing demand for the design of fast concurrent data structures that can be supported by multiple computing threads (or processes). Due to their massive amount of parallelism, building efficient concurrent data structures that scale to the level of (multiple) GPUs is a particularly challenging problem. Recent examples include the design of GPU-enabled hash tables, B-Trees, and suffix arrays. In this project we want to design even more massively parallel CUDA-based concurrent data structures for big data and AI applications (such as bioinformatics and life sciences). Those should take advantage of the features of modern GPUs such as cache-hierarchy, distributed shared memory, NV-link interconnects, atomic operations, and Tensor Cores.

Working group and link(s) to working group:

- High Performance Computing Research Group, Johannes Gutenberg University Mainz.

Professional prerequisites for applicants:

- Very good programming skills (e.g. C++, CUDA, Python).
- Basic knowledge of GPUs, HPC and algorithms & data structures.

Working language/required language skills:

- English.
- German.

Description of supervisor and scientific institution:

- Prof. Dr. Bertil Schmidt, JGU Mainz, Chair of High-Performance Computing Research Group.

3#SW: Algorithmic Differentiation in HPC Environments

Description:

Modern HPC architecture faces several rapid changes, for example, with respect to domain-specific languages, accelerators and communications models that have evolved over the last decades. On the other hand, the calculation of exact derivative information based on the theory of Algorithmic Differentiation is indispensable for many tasks like backpropagation in the machine learning context or in the field of derivative-based numerical optimization.

AD techniques can be used to enhance DSL frameworks or multi-language applications, for example with python as a wrapper language with the capability to additionally generate the derivative code of the problem formulation. The generated derivative code is then again optimized for different hardware architectures and run-time environments. Here, important examples are the PGAS programming model, CUDA, ROC, FPGAs, TPUs or GPUs. Each library needs to be analyzed for its specific challenges and extended such that a differentiated version can be used in the context of AD.

Working group and link(s) to working group:

- Scientific Computing Group, University of Kaiserslautern-Landau (RPTU).

Professional prerequisites for applicants:

- Very good programming skills (e.g. C++, CUDA, Python).
- Basic knowledge of algorithmic differentiation.

Working language/required language skills:

- English.
- German.

Description of supervisor and scientific institution:

- Prof. Dr. Nicolas Gauger, University of Kaiserslautern-Landau (RPTU), Chair of Scientific Computing Research Group & Director of Computing Center (RHRZ).
- Co-supervised by Prof. Dr. Andrea Walter, Humboldt-Universität zu Berlin, Chair of Mathematical Optimization & Member of the Department AI in Society, Science, and Technology (AIS2T) at NHR@ZIB.

Potential NHR Center for Ph.D. thesis and/or secondment collaboration:

- Prof. Dr. Andrea Walter, NHR@ZIB.

4#SW: LLM-based Compiler Generation for Quantum Computers

Description:

Compilers for quantum computers (QC) transform high-level quantum algorithms into low-level instructions that can be executed on a specific quantum computer architecture. The compiler therefore decomposes high-level quantum gates into hardware-compatible elementary gates, maps logical qubits from the algorithm to physical qubits on the hardware, and swaps or moves qubits between different parts of the quantum processor.

Optimizing compilers for quantum computers is therefore highly dependent on the underlying QC architecture. Large-language models (LLMs) can help to automatically generate and optimize quantum compilers, so that changes to the architecture do not require to develop and implement new algorithms in case of architecture improvements. The use of LLM-based QC compilers should be investigated according to the needs of trapped-ion quantum computers developed at Johannes Gutenberg University. Eventually, novel compilers will be applied on this hardware.

Working group and link(s) to working group:

- Efficient Computing and Storage Research Group, Johannes Gutenberg University Mainz.
- Cold Ions and Experimental Quantum Information Processing Research Group, Johannes Gutenberg University Mainz.

Professional prerequisites for applicants:

- Good programming skills (e.g. C++, CUDA, Python).
- Knowledge of quantum computing and compilers.
- Basic knowledge of AI/ML.

Working language/required language skills:

- English.
- German.

Description of supervisor and scientific institution:

- Prof. Dr.-Ing. André Brinkmann, JGU Mainz, Chair of Efficient Computing and Storage.
- Prof. Dr. Ferdinand Schmidt-Kaler, JGU Mainz, Chair of Cold Ions and Experimental Quantum Information Processing.

5#SW: In-network Processing for HPC Applications

Description:

Data Processing Units (DPUs) represent a new generation of intelligent network interface cards, capable of offloading key applications and management tasks traditionally handled by HPC compute nodes. By processing packets directly within the network, DPUs reduce latency, alleviate CPU workloads, and minimize application and network jitter.

However, their constrained processing and memory capabilities require rethinking fundamental data structures, such as indexes and filters, to ensure efficiency. Moreover, applications and storage management systems must be adapted to fully leverage the benefits of outsourcing functionalities to DPUs. This research explores the relationship between DPUs, operating systems, and HPC applications, addressing challenges in performance optimization, system integration, and architectural design.

Working group and link(s) to working group:

- Efficient Computing and Storage Research Group, Johannes Gutenberg University Mainz.

Professional prerequisites for applicants:

- Very good programming skills (e.g. C++, CUDA, Python).
- Basic knowledge of HPC, operating system and computer networks.

Working language/required language skills:

- English.
- German.

Description of supervisor and scientific institution:

- Prof. Dr.-Ing. André Brinkmann, JGU Mainz, Chair of Efficient Computing and Storage.

GRADUATE SCHOOL

6#SW: Modular and Automated Workload Analysis for HPC Systems

Description:

Modern HPC systems are highly complex, and achieving theoretical peak performance depends on numerous parameters and configurations. To optimize performance and resource efficiency, methods such as simulation, benchmarking, and monitoring are commonly used. However, these tools often lack compatibility and address only specific performance aspects—such as network, I/O, resource allocation, or parallel execution. Each approach generates valuable insights that could benefit similar use cases or system configurations. To ensure this knowledge is not lost or limited to one-off applications, it must be made accessible to the wider community.

The MAWA-HPC (Modular and Automated Workload Analysis for HPC Systems) framework aims to develop a universal, modular workflow and tool infrastructure - from execution environment to evaluation dashboard - applicable across diverse workloads and scientific domains. Its architecture enables integration of community tools, improving interoperability and supporting a broader range of use cases.

By leveraging established monitoring, profiling, and benchmarking tools, including node-level performance engineering, network benchmarks, and microbenchmarks—MAWA-HPC will integrate multiple performance models, such as a multi-dimensional Roofline model and a holistic performance scoring system.

This score will consider compute, network, parallel I/O, and time-based metrics. Previous work introduced a platform-agnostic I/O Roofline model; by incorporating time as an additional dimension, it allows performance analysis over time and helps identify anomalies. Potential PhD research topics include generating user hints, detecting bottlenecks, pattern analysis, performance prediction, and optimization—potentially enhanced by ML and AI techniques.

Working group and link(s) to working group:

- High Performance Computing and its Applications Research Group, Johannes Gutenberg University Mainz.

Professional prerequisites for applicants:

- Very good programming skills (e.g. C++, CUDA, Python).
- Basic knowledge of HPC, parallel I/O and computer networks.

Working language/required language skills:

- English.
- German.

Description of supervisor and scientific institution:

- Prof. Dr. Sarah Neuwirth, JGU Mainz, Chair of High-Performance Computing and its Applications and Co-Director of NHR@SW.

Potential NHR Center for Ph.D. thesis and/or secondment collaboration:

- NHR4CES
- NHR@Göttingen
- NHR@FAU.

NHR@ZIB – Zuse Institute Berlin

At Zuse Institute Berlin (ZIB), model-based simulation and optimization, data-driven problem solving, and advanced computing meet with high-performance research services and transfer to diverse application fields and industries. The NHR center at ZIB is part of a research institute with a focus on Scientific Computing and a broad spectrum of interdisciplinary research projects and about 300 scientists and research infrastructure specialists. The NHR center at ZIB will take the leadership in selected application domains, e.g., life sciences with a focus on model-driven simulations, advanced integration of artificial intelligence/machine learning and simulation, or heterogeneous HPC architectures jointly with other NHR centers.

1#ZIB: Sparsity-Driven Optimization for Large Language Models under Resource Constraints

Description:

The exponential growth in the size of large language models (LLMs) presents significant challenges in terms of memory and computational resources, particularly in high-performance computing (HPC) environments. At ZIB, we propose to explore innovative sparsity-driven optimization techniques that can effectively reduce the resource footprint of LLMs without compromising their performance. This research will focus on developing novel algorithms that leverage sparsity in the form of removed weights or entire neural network layers to enhance model efficiency, enabling the deployment of LLMs on resource-constrained HPC architectures. The project will involve three key components: (1) designing sparsity-inducing regularization methods tailored for LLMs, (2) implementing efficient sparse matrix operations on heterogeneous HPC platforms, and (3) evaluating the trade-offs between sparsity, model accuracy, and computational efficiency. By integrating these approaches, we aim to advance the state-of-the-art in scalable AI solutions, facilitating the broader application of LLMs in scientific computing and industry.

Working group and link(s) to working group:

- Interactive Optimization and Learning (IOL) (ZIB).

Professional prerequisites for applicants:

- Deep Learning Knowledge.
- Preferably PyTorch experience.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Dr. Sebastian Pokutta, Zuse Institute Berlin and Technische Universität Berlin.

GRADUATE SCHOOL

2#ZIB: AI-enhanced Molecular Dynamics of Fluid-Solid Interfaces

Description:

The convergence and combination of classic HPC methods with approaches from artificial intelligence (AI) is an ongoing trend in science. Molecular dynamics (MD) benefits from AI as well: machine-learned interaction potentials (MLIPs) enable the accurate modeling of many-body interactions in materials research. They allow for predictions of undiscovered chemical species, describing the chemical space of a given structure, or precisely predicting the pathways between folded and unfolded conformations of a protein. Many flavors of MLIPs are developed by a growing research community, using deep neural networks, Gaussian approximation kernels, or message-passing atomic cluster expansions. Compared to classical force fields, all these approaches are computationally still extremely costly.

To effectively support molecular dynamics research with these promising AI-based models, they need to be combined with existing HPC methods and codes. Accelerator hardware is crucial to fulfill the computational demands of such research. In this project, you will develop an MLIP framework for use with accelerators (NVIDIA GPUs and others) within the simulation software "HAL's MD package" (https://halmd.org). The focus is on efficient algorithms and implementations and their application to AI-enhanced MD simulations of fluid-solid interfaces.

Working group and link(s) to working group:

- Dynamics of Complex Materials (Freie Universität Berlin/ZIB).

Complex materials are created through the interaction of different phases and types of substances. The dynamic processes take place simultaneously on various temporal and spatial scales. A numerical simulation of complex materials therefore requires new algorithmic approaches to map this variability. To this end, we develop and implement parallel and scalable methods for future innovative HPC architectures. The range of applications includes physical, chemical and biological systems.

Professional prerequisites for applicants:

- M.Sc. in Physics, Theoretical Chemistry, Computer Science, Applied Mathematics, or related field.
- Interest in algorithm development for accelerator hardware.
- Solid knowledge of modern C++ and experience with Linux at CLI level.
- Experience with parallel programming on GPUs (CUDA, SYCL, OpenMP, MPI) is a plus.
- A background in atomistic simulations is a plus.

Working language/required language skills:

- You write and speak English fluently.
- First experience with authoring scientific texts is a plus.

Description of supervisor and scientific institution:

- Prof. Dr. Felix Höfling, ZIB and Freie Universität Berlin.

Prof. Höfling's research focuses on dynamic phenomena that emerge in complex fluids and soft matter, typically covering many scales in length and time. He is passionate about high-precision computer simulations using cutting-edge technology, which he combines with mathematical analyses of coarse-grained models.

ZIB is working closely with researchers and software developers from the MD domain to enable the described combinations of AI and HPC. In addition, a focus of ZIB's research in the Supercomputing department is on the cross-platform support of AI methods on accelerators by using modern programming languages and vendor-independent frameworks.

3#ZIB: Portable and Energy-Efficient Computing on Data-Parallel Platforms

Description:

Today, with GPUs, data-flow engines, TPUs, and others, a variety of data-parallel accelerators with native support of multiple data representations are available. The more domain-specific algorithms can be effectively mapped on these accelerators, the more energy-efficient data processing can be realized in the future. Whereas for dense and partially sparse linear algebra advantages of mixed precision and approximate computing strategies have been demonstrated on these devices, there are many questions in other areas of numerical mathematics like PDE solvers unanswered. Further, from the technical point of view some challenges need to be tackled. The current "art of coding" is under question as with the diversity of accelerators one must select the associated programming models. At ZIB, we want to continue our interdisciplinary work in the areas of mathematical numerics and modern software engineering by exploring the potential of mixed-precision algorithms, e.g., on unstructured grids on current and next-generation data-parallel processors. For that, our emphasis is on programming models being portable across vendors and processor architectures (e.g., SYCL and OpenMP).

Working group and link(s) to working group:

- Algorithms for Innovative Architectures (ZIB).
- Computational Anatomy and Physiology (ZIB).

Professional prerequisites for applicants:

- Basic knowledge of C/C++; Fortran is a plus.
- Solid experience with Linux at CLI level.
- Experience with parallel programming, i.e., MPI, OpenMP and CUDA is a plus.

Working language/required language skills:

- English.
- German.

Description of supervisor and scientific institution:

- Dr. Thomas Steinke, Zuse Institute Berlin.
- Dr. Martin Weiser, Zuse Institute Berlin.

GRADUATE SCHOOL

4#ZIB: Semantic I/O with Object Stores in HPC

Description:

Object stores like S3 are a common storage infrastructure in cloud environments. Nowadays, object stores have become increasingly popular in HPC as an alternative concept to store data as they overcome scalability issues of POSIX-complaint file systems used traditionally in HPC configurations. Examples of HPC-relevant objects store are based on distributed key/value stores. This offers the possibility to integrate application-specific metadata directly into the managed objects at a low level. Moreover, with the flexibility of a key/value store the efficiency of I/O operations can be further increased as the logical structure of the data to be stored can be mapped to an optimized schema of the key/value backend. As such, these data containers are natively supported by the storage system with best performance for storing and retrieving the data. At ZIB, we perform research with one of the current implementations of a distributed object store, the Distributed Asynchronous Object Store (DAOS). Using its infrastructure, the potential benefit of taking the semantic knowledge of data into account will be investigated.

Working group and link(s) to working group:

- Algorithms for Innovative Architectures (ZIB)

Professional prerequisites for applicants:

- Basic knowledge of C/C++; Fortran is a plus.
- Solid experiences with Linux at the CLI level.
- knowledge of the object store concept is a plus.

Working language/required language skills:

- English.
- German.

Description of supervisor and scientific institution:

- Dr. Steffen Christgau, Zuse Institute Berlin.

5#ZIB: Modelling Excitations in Novel Materials

Description:

Ab-initio calculations are a valuable tool to study materials for next-generation optoelectronic devices, including those that exploit solar energy. In this context, a new class of optoelectronic materials has emerged, comprising two-dimensional systems – such as graphene and transition metal dichalcogenides (TMDCs) – in combination with organic molecules. These hybrid systems offer promising properties for novel applications, yet their accurate theoretical description remains challenging and largely unexplored.

Standard density functional theory (DFT) approaches are insufficient for capturing their electronic and optical properties with the required accuracy. Instead, methods going beyond DFT, especially those based on the GW approximation and the Bethe-Salpeter equation (BSE), or time-dependent DFT (TDDFT) methods, are necessary. Their application, however, require larger computational resources.

The PhD candidate will leverage the powerful hardware available at NHR@ZIB to implement and apply these state-of-the-art methods on modern energy efficient architectures. A closely related research direction involves studying the dynamics of exciton creation and propagation in these hybrid inorganic/organic materials, which play a crucial role in optoelectronic devices. Depending on the research focus, these topics may also be combined with the development of computational tools that expand the functionality of a selected ab-initio package.

Working group and link(s) to working group:

- NHR@ZIB, https://www.zib.de/research-services/nhr.

Professional prerequisites for applicants:

- Good knowledge of the density functional theory or other methods for electronic structure calculations.
- basic knowledge of Fortran, C/C++, or Python.

Working language/required language skills:

- English.
- German.

Description of supervisor and scientific institution:

- Dr. Ronaldo Rodrigues Pela, Zuse Institute Berlin.

6#ZIB: HPC-based Simulation of Agent-Based Models and Complex Networks

Description:

Simulations of agent-based models (ABMs) have become an essential tool for understanding complex systems in social science and mobility research. These models represent large populations of heterogeneous, autonomous agents whose interactions give rise to emergent phenomena such as traffic flows, urban dynamics, opinion formation, or epidemic spreading. Due to the computational intensity of modeling millions of agents and their fine-grained interactions over space and time, high-performance computing (HPC) infrastructures are crucial to enabling realistic, scalable, and timely simulations.

By leveraging parallelization strategies — such as spatial domain decomposition, task scheduling, or GPU acceleration — HPC architectures can significantly reduce computation time and allow for real-time decision support, large-scale sensitivity analysis, and uncertainty quantification. In mobility research, ABMs support the design and evaluation of transportation systems, infrastructure planning, and policy-making under dynamic conditions. In social dynamics, they offer insight into collective behavior, network effects, and diffusion processes.

Furthermore, coupling ABMs with data assimilation methods and AI techniques enhances predictive power and allows data-driven scenario exploration. HPC-based agent simulations thus bridge micro-level behavior and macro-level outcomes, making them invaluable for evidence-based planning and interdisciplinary research across urban studies, epidemiology, and economics.

Working group and link(s) to working group:

- Modeling and Simulation of Complex Processes (ZIB)
- Mathematics for Sustainability Transitions (Freie Universität Berlin)

Professional prerequisites for applicants:

- M.Sc. in Physics, Computer Science, Applied Mathematics, or related field.
- Interest in algorithm development for HPC hardware.
- Some background in ABM construction and simulation
- Experience with parallel programming on GPUs (CUDA, SYCL, OpenMP,
- MPI) is a plus but not a strict requirement.

Working language/required language skills:

- English.
- German.

Description of supervisor and scientific institution:

Modeling and Simulation of Complex Processes:

- Dr. Natasa Conrad, Zuse Institute Berlin
- Prof. Dr. Christof Schütte, Zuse Institute Berlin /Freie Universität Berlin

Mathematics for Sustainability Transitions:

- Prof. Dr. Sarah Wolf, Technische Universität Berlin.

7#ZIB: AI-guided high -throughput drug design

Description:

AI-guided drug design is revolutionizing the discovery of new therapeutic compounds by enabling the exploration of vast chemical spaces far beyond human capacity. In this approach, artificial intelligence models—often based on deep learning or reinforcement learning—are trained to generate, filter, and prioritize candidate molecules based on predicted biological activity, toxicity, synthesizability, and pharmacokinetic properties. These models operate within a multi-objective optimization framework, balancing several competing criteria to identify compounds that best meet the therapeutic goals.

One of the key advantages of AI-driven methods is the ability to evaluate billions of potential drug candidates in silico, rapidly screening them against a biological target such as a protein or receptor. By integrating molecular docking, molecular dynamics simulations, and structure-based scoring functions, AI systems can efficiently identify promising binders while discarding inactive or unsafe molecules. Generative models can further propose novel scaffolds that may not exist in current databases, fostering true innovation.

High-performance computing resources are essential in this process to manage the large-scale parallel simulations and data throughput required. This enables iterative refinement of models and candidates through active learning, where the AI learns from feedback cycles. Ultimately, AI-guided drug design accelerates early-stage drug discovery, reduces experimental costs, and opens up new therapeutic avenues in precision medicine and rare disease research.

Working group and link(s) to working group:

- Modeling and Simulation of Complex Processes

Professional prerequisites for applicants:

- M.Sc. in Physics, Theoretical Chemistry, Computer Science, Applied Mathematics, or related field.
- Interest in algorithm development for HPC hardware.
- Some background in using AI methods
- Experience with parallel programming on GPUs (CUDA, SYCL, OpenMP,
- MPI) is a plus but not a strict requirement.

Working language/required language skills:

- English.
- German.

Description of supervisor and scientific institution:

- Prof. Dr. Christof Schütte, Zuse Institute Berlin /Freie Universität Berlin
- Prof. Dr. Konstantin Fackeldey, Zuse Institute Berlin/ Technische Universität Berlin
- PD Dr. Marcus Weber, Zuse Institute Berlin.

PC2 – Paderborn University

The Paderborn Center for Parallel Computing (PC2) is a central research institute and HPC center at Paderborn University with the mission to conduct interdisciplinary research at the intersection of computational science, computer science, and innovative computer architecture. Within the NHR Alliance, PC2 focuses on applications in computational physics and chemistry; numerical methods and libraries for heterogeneous computing systems; and energy-efficient computer architectures such as FPGAs and GPUs. The supervision of the proposed doctoral topics with a strong connection to natural sciences is carried out jointly with corresponding chairs from the Faculty of Natural Sciences.

1#PC2: Empirical research on HPC Systems for Optimization of Operation

Description:

In this line of research, we plan to strengthen and expand our activities on empirical studies in the operation of high-performance computing systems. The project aims to collect and analyze real-world operational data of data center infrastructure and HPC systems – such as energy consumption, processor and memory usage patterns, compute job characteristics – gathered through infrastructure, server and job-specific monitoring mechanisms. Therefore, the HPC system itself and its use become the subjects of research. By analyzing the gathered usage data, we seek to understand how resources are used, identify optimization opportunities, and develop strategies to improve both system operations and scientific workflows.

Examples of possible outcomes include:

- Identifying distinct usage patterns across scientific communities and their open-source research software, enabling targeted software optimization.
- Supporting strategic planning for future HPC system acquisitions by aligning infrastructure choices with actual user needs.
- Enhancing operational efficiency by detecting and addressing bottlenecks in computation, storage architecture, and network performance.
- Ultimately, the research aims to bridge the gap between system operation and user needs, promoting more efficient and energy-aware use of HPC resources.

Working group and link(s) to working group:

- For further information, please have a look here: https://en.cs.uni-paderborn.de/hpc/team/group/members.

Professional prerequisites for applicants:

- Solid background in computer engineering (computer architecture, operating systems, low-level programming).
- Knowledge and practical experiences with HPC (parallel programming, use of HPC systems).
- Data science skills (statistics, visualization).

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Dr. Christian Plessl, Professor for High-Performance Computing at the department of Computer Science at Paderborn University and managing director of the Paderborn Center for Parallel Computing (PC²).

2#PC2: FPGA-Acceleration for High-Performance Computing and AI

Description:

The Paderborn Center for Parallel Computing (PC²) is among the world's leading institutions in the investigation of Field Programmable Gate Arrays (FPGAs) for use in high-performance computing (HPC) and artificial intelligence (AI). In this line of research, we aim to expand our activities by exploring the integration and utilization of FPGAs in scientific computing environments.

This project is closely aligned with the key research areas of PC² and offers several promising directions for a PhD thesis, including:

- Accelerating atomistic simulations with FPGAs, particularly in the context of solid-state physics and theoretical chemistry, where compute-intensive kernels can greatly benefit from hardware specialization.
- Designing reusable FPGA libraries for scientific computing, especially for mixed-precision linear algebra operations, iterative solvers, and other performance-critical computational patterns.
- Developing massively parallel HPC applications that make intensive use of FPGA interconnects and communication infrastructures, enabling scalable and efficient custom computing pipelines.

This research line contributes to the broader goal of making reconfigurable computing a first-class citizen in HPC and AI infrastructures, with a strong emphasis on energy efficiency, customization, and scientific impact.

Working group and link(s) to working group:

- For further information, please have a look here: https://en.cs.uni- paderborn.de/hpc/team/group/members.

Professional prerequisites for applicants:

- Solid background in computer engineering or embedded systems (digital logic, hardware description languages, computer architecture).
- Experience with FPGA programming (e.g., VHDL/Verilog or high-level synthesis).
- Interest in HPC and AI applications.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Dr. Christian Plessl, Professor for High-Performance Computing at the department of Computer Science at Paderborn University and managing director of the Paderborn Center for Parallel Computing (PC²).

3#PC2: GPU Performance and Portability

Description:

This Ph.D. project focuses on developing new methods, libraries, and approaches for improving the performance and portability of scientific workloads on GPU architectures. The target applications are from the domain of computational physics and theoretical chemistry.

The project shall focus on workloads that currently still rely on CPU-based execution and have not yet been adapted for GPU acceleration. A central research objective is to explore innovative techniques for porting and optimizing such workloads, including:

- Unconventional use of tensor cores on GPUs to accelerate numerical kernels.
- Improving cross-vendor portability, targeting at least NVIDIA and AMD GPU platforms.
- Exploring modern programming models and libraries such as SYCL, Kokkos, or RAJA to develop portable and efficient GPU implementations.

The overarching goal is to lower the barrier for GPU adoption in scientific computing and provide generalizable tools and insights for the community.

Working group and link(s) to working group:

- For further information, please have a look here: https://en.cs.uni- paderborn.de/hpc/team/group/members.

Professional prerequisites for applicants:

- Background in computer science, computational physics, or a related field.
- Experience with parallel programming and GPU computing (e.g., CUDA, OpenCL, or SYCL).
- Familiarity with scientific computing workflows and numerical methods.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Dr. Christian Plessl, Professor for High-Performance Computing at the department of Computer Science at Paderborn University and managing director of the Paderborn Center for Parallel Computing (PC^2).

4#PC2: Compilation for Parallel Computer Architectures and Accelerators

Description:

This research project focuses on domain-specific compilation strategies for modern parallel computing architectures and accelerators. The aim is to leverage and extend the capabilities of the LLVM-based MLIR (Multi-Level Intermediate Representation) compiler infrastructure to support a wide range of accelerators, including FPGAs, GPUs, and AI engines.

A central aspect of the project will be the development of new MLIR dialects tailored to specific scientific application domains, or the extension of existing dialects – such as those for stencil computations or distributed computing – to better the support the needs of scientific applications in computational physics and chemistry. Possible focus areas within this topic include:

- Architecture-specific backend optimizations to fully exploit the capabilities of heterogeneous compute platforms.
- Automatic techniques for execution on distributed parallel systems.
- Integration with numerical libraries for linear algebra, solvers, spectral transformations, and other fundamental operations in scientific computing.
- Performance benchmarking and comparison of compiler-generated versus hand- optimized code.
- Emphasis on energy efficiency in addition to performance.
- Enhancing portability to future hardware architectures, while main-taining backward compatibility with current processor generations and vector instruction sets.

This project offers opportunities to contribute to cutting-edge compiler research and to make a tangible impact on how scientific software is developed and executed on modern HPC infrastructures.

Working group and link(s) to working group:

- For further information, please have a look here: https://en.cs.uni- paderborn.de/hpc/team/group/members.

Professional prerequisites for applicants:

- Strong background in compiler construction, computer architecture, or programming languages.
- Experience with LLVM or similar compiler frameworks.
- Familiarity with scientific computing and parallel programming.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Dr. Christian Plessl, Professor for High-Performance Computing at the department of Computer Science at Paderborn University and managing director of the Paderborn Center for Parallel Computing (PC²).

NHR4CES – RWTH Aachen University and Technical University Darmstadt

In NHR4CES, The Rheinisch-Westfälische Technische Hochschule Aachen (RWTH Aachen University) and The Technical University of Darmstadt (TuDa) join forces to combine their strengths in HPC applications, algorithms, and methods, and the efficient use of HPC hardware. Our goal is to create an ecosystem combining best practices of HPC and research data management to address questions that are of central importance for technical developments in the economy and society. We support the development of scalable algorithms and software for the investigation, development, design, construction, evaluation, and production of engineering applications with a particular focus on engineering and materials science as well as engineering-oriented physics, chemistry, and medicine.

1#NHR4CES: Building Foundation Models for Operational Processes

Description:

The project *Building Foundation Models for Operational Processes* aims to pioneer a new class of foundation models tailored to the unique structure and semantics of operational process data. While current foundation models and large language models (LLMs) are predominantly trained on textual corpora, enterprise systems like SAP and Salesforce generate rich, structured, event-centric data that remains underutilized by modern AI.

Leveraging the high-performance computing (HPC) resources of NHR4CES, this project will develop and train large-scale models capable of understanding, generalizing, and predicting complex business process behavior across domains. The research builds on and extends the work of the Process and Data Science (PADS) Group at RWTH Aachen University, which has
laid the theoretical and methodological groundwork in object-centric process mining. Through close collaboration with the CSG Data Science and Machine Learning in NHR4CES, the project integrates advanced deep learning and data engineering techniques to create foundation models that can drive automation, compliance, and optimization in digital
enterprises. The resulting models are expected to serve as a universal representation layer for operational processes, enabling novel applications in predictive analytics, process simulation, and generative process design.

Working group and link(s) to working group:

- For further information, please have a look at the Process and Data Science (PADS) Group at RWTH Aachen University and the Cross-Sectional Group Data Science and Machine Learning: https://www.nhr4ces.de/cross-sectional- groups/csg-date-science-and-machine-learning/.

Professional prerequisites for applicants:

- Strong data science and machine learning skills.
- Being able to develop complex software systems in an HPC setting.
- Interest in improving operational processes, as is demonstrated by basic knowledge of process mining and process modelling.

Working language/required language skills:

- English.
- German.

Description of supervisor and scientific institution:

- Prof. Dr. Wil van der Aalst, Chair of Process and Data Science, RWTH Aachen University.

2#NHR4CES: Modeling metastable phase formation for physical vapor deposition thin films

Description:

Metastable coatings of thin films prepared through physical vapor deposition (PVD) techniques, such as magnetron sputtering, are widely used in cutting and forming applications. Understanding the conditions for metastable phase formation, e.g., metastable solid solubility, is crucial to the target for determining the material's properties. The limitations of purely thermodynamic energy-based approaches become apparent, as they cannot account for the full extent of metastable solid solubility measured experimentally.
Describing the metastable phase formation during thin film growth is based on both thermodynamics and kinetics. The proposed project is designed such that the metastable phase formation of the proposed materials will be studied by ab initio calculations and machine learning. Additionally, the results will be validated through experiments.

Working group and link(s) to working group:

- For further information, on the SDL Materials Design working group, please have a look here:
  https://www.nhr4ces.de/simulation-and-data-labs/sdl-materials-design/.
- For further information, on the Materials Chemistry working group, please have a look here:
  https://www.mch.rwth-aachen.de/cms/~jeuz/MCH/lidx/1/.

Professional prerequisites for applicants:

- Experience in *ab initio* calculations and materials simulations.
- Experience in coupling modelling to PVD synthesis techniques.
- Strong background in Materials Sciences.
- Programming skills in Python, working experience in the Linux interface.

Working language/required language skills:

- English (mandatory).
- German (optional).

Description of supervisor and scientific institution:

- Prof. Dr. Jochen Schneider, RWTH Aachen University, Chairman of the Board of StrucMatLab.

GRADUATE SCHOOL

3#NHR4CES: Computational Neuroscience

Description:

The neural network of the human brain is not hardwired. Even in the mature brain, new connections between neurons are formed and existing ones are deleted, which is called
structural plasticity. This process is key to understanding how learning, memory, and healing after lesions such as stroke work. In this PhD project, the candidate will design scalable algorithms to simulate structural plasticity and apply them to study memory and learning models.

Working group and link(s) to working group:

- For further information on Laboratory for Parallel Prgramming and Cross-Sectional Group Parallelism & Performance please have a look here: https://www.nhr4ces.de/cross-sectional-groups/csg-parallelism-and-performance/.

Professional prerequisites for applicants:

- Master's degree in computer science or related discipline.
- Experience in algorithm design.
- Knowledge of graph theory.
- Proficient in C/C++.
- A background in neuroscience and/or parallel programming is an advantage.

Working language/required language skills:

- Very good command of the English language.

Description of supervisor and scientific institution:

- Prof. Dr. Felix Wolf, Professor for Parallel Programming at TU Darmstadt, Department of Computer Science and Head of NHR4CES@TU Darmstadt.

4#NHR4CES: HPC-Driven Low-Cost Magnetic Resonance Imaging

Description:

Medical imaging, especially magnetic resonance imaging (MRI), is essential for the early detection of cancer. However, the high cost of these procedures limits the frequency of examinations and the scope of preventive screening. We therefore advocate combining affordable low-cost MRI hardware with powerful high-performance computers (HPC). By offloading image accuracy and reconstruction intelligence to computers, diagnostic quality is maintained while costs are drastically reduced.

The research project involves developing comprehensive simulation models for MRI systems that integrate all components - from system engineering and signal encoding to image reconstruction and post-processing - into a single framework. The core idea is to substitute costly physical parts with advanced simulations and machine learning models, utilizing regular access to HPC resources even after optimization. This innovative solution requires tightly coupled multiscale calculations that generate large amounts of data and require GPU clusters in HPC environments for efficient processing. When traditional physical modelling reaches its limits, neural networks with residual physics help close gaps in dynamic approximations, while reinforcement learning agents optimize pulse sequences under low signal-to-noise ratio conditions. With reconstruction and system modelling outsourced to HPC, this new class of imaging systems can deliver high diagnostic quality, enabling cost-effective, simulation-based imaging for cost-efficient diagnostics.

Link to working group:

- For further information, please have a look at the website of working group Simulation and Data Lab Digital Patient: https://www.nhr4ces.de/simulation-and- data-labs/sdl-digital-patient/.

Professional prerequisites for applicants:

- Master's degree in physics, electrical engineering, or computer science.
- Solid background in Physics and numerical simulation (Maxwell/Bloch/FEM).
- Experience with machine learning (PyTorch/JAX), reinforcement learning, and high-performance computing.
- Strong programming skills in C++ and Python; familiarity with CUDA or GPU-accelerated frameworks is a plus.

Working language/required language skills:

- English (mandatory).
- German (optional).

Description of supervisor and scientific institution:

- Prof. Dr. Volkmar Schulz, RWTH Aachen, Director of Chair for Imaging and Computer Vision, Faculty of Electrical Engineering, Physics and Medicine.
- Prof. Dr. Julia Kowalski, RWTH Aachen, Director of Chair of Model-based Development (MBD), Faculty of Mechanical Engineering.

5#NHR4CES: Computational 4D Metabolic Imaging through Kinetic Modelling

Description:

Positron emission tomography (PET) – the most sensitive method for examining metabolism is on the verge of transitioning from static snapshots to continuous 3D films of the entire body, providing terabytes of raw data for each patient. To convert these large amounts of data into real-time maps of tracer kinetics, a fast simulation and reconstruction platform is required that can model, predict and instantly update the flow of the radiotracer through each organ by accessing HPC. This project aims to develop an integrated computational framework that models the entire PET imaging chain - from detector physics and annihilation events to particle transport and tracer kinetics - within a unified simulation and optimization platform. High-fidelity Monte Carlo simulations will model the physical interactions, while organ-specific tracer uptake will be described by coupled differential equations. These equations will be solved in real time using GPU-accelerated computing. To enhance efficiency, neural-network-based surrogate models will approximate computationally intensive components. Additionally, reinforcement learning will be employed to dynamically optimize scan parameters for different tracers and anatomies, with the goal of maximizing image quality and quantitative accuracy at minimal radiation dose. By offloading computationally intensive tasks to scalable computing platforms, we enable total-body PET scans with detailed kinetic modelling, paving the way for a new level of diagnostic precision.

Link to working group:

- For further information, please have a look at the website of working group *Simulation and Data Lab Digital Patient*: https://www.nhr4ces.de/simulation-and- data-labs/sdl-digital-patient/.

Professional prerequisites for applicants:

- Master's degree in physics, electrical engineering, or computer Science.
- Sound knowledge of physics, numerical solutions of PDEs and Monte Carlo simulations.
- Experience with machine learning (PyTorch/JAX) and high-performance computing.
- Strong programming skills in C++ and Python; familiarity with CUDA or
- GPU-accelerated frameworks is a plus.

Working language/required language skills:

- English (mandatory).
- German (optional).

Description of supervisor and scientific institution:

- Prof. Dr. Volkmar Schulz, RWTH Aachen, Director of Chair for Imaging and Computer Vision, Faculty of Electrical Engineering, Physics and Medicine.
- Prof. Dr. Felix Mottaghy, Director of the Clinic for Nuclear Medicine, University Clinic RWTH Aachen.

6#NHR4CES: Digital Twins

Description:

Due to the considerable heterogeneity of patients, especially in intensive care medicine, optimal individualized treatment strategy requires continuous monitoring of all relevant patient specific parameters. However, despite substantial academic and industrial research efforts with AI-based or mechanistic equation-based models, Currently, there are no models available that fulfil the clinical requirements for reliability, predictive quality, individual adaptability, explainability and response time in a clinical IT environment. Furthermore, the disease process develops dynamically in interaction with the respective therapeutic measures. Thus, the concept of a digital twin is the technology of choice.

For the simulation of patient-specific ventilation strategies in intensive care medicine, our group uses the ICSM simulator, which has around 1500 algebraic and dynamic equations that can be adapted to a single patient with the help of up to approx. 600 parameters. As part of our research, we are developing a novel mathematical strategy in which transfer learning based on very large, generic patient data sets and a hybrid approach of machine learning and mechanistic modelling is used. We aim to develop a dynamic concept that allows reliable quantification of uncertainty for patient-specific settings.

Link to working group:

For further information, please have a look here:

- https://www.ukaachen.de/kliniken-institute/joint-research-center-for-computational-biomedicine/lehre/institute-for-computational-biomedicine/.
- Simulation and Data Lab Digital Patient: https://www.nhr4ces.de/simulation-and-data-labs/sdl-digital-patient/.

Professional prerequisites for applicants:

- Strong programming skills (e.g. Python, MATLAB).
- Experience in ML tools.
- Knowledge in numerical and optimization methods is desirable.

Working language/required language skills:

- English (mandatory).
- German (optional).

Description of supervisor and scientific institution:

- Prof. Dr. Andreas Schuppert, RWTH Aachen University.

7#NHR4CES: Efficient Scientific Visualization on HPC Infrastructure

Description:

Advances in technology and algorithms have enabled researchers to use large-scale computational simulations to study complex physical phenomena such as weather, mechanics, and biology. As simulation data becomes increasingly large and complex, efficient processing and analysis have become significant challenges. Visualization provides the means for efficient visual analysis, exploration, and communication of such data but needs to be adapted to support HPC-related applications. Novel efficient techniques and In-situ visualization help address this by focusing on computational resources where needed and enabling real-time data analysis during the simulation. As part of the CSG Visualization, you develop and implement such novel strategies to support the effective and efficient visual analysis of complex data on HPC clusters.

Link to working group:

- For further information, please have a look here: https://www.nhr4ces.de/cross-sectional-groups/csg-visualization/.

Professional prerequisites for applicants:

- Basic Knowledge of Computer Graphics Concepts.
- Good programming skills (e.g., C++, CUDA).

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Dr. Tim Gerrits, RWTH Aachen University, Lead of the Visualization Team.

8#NHR4CES: LLM-based Data Wrangling for HPC Applications

Description:

The PhD position on LLM-based Data Wrangling for HPC Applications aims to explore the integration of large language models (LLMs) into the data preparation pipeline for high-
performance computing (HPC) workflows, with a particular focus on tabular data. In many scientific domains, HPC simulations and analyses rely on complex tabular datasets that
require extensive wrangling—cleaning, transformation, normalization, and semantic alignment—before they can be effectively used. This process is often manual, error-prone, and domain-specific. Recent advances in LLMs offer a promising avenue for automating and generalizing these tasks by leveraging their ability to understand and generate structured data representations. The research will investigate how LLMs can be adapted, fine-tuned, or augmented to support scalable, context-aware data wrangling for large and heterogeneous scientific tables, enabling more efficient data preparation and accelerating scientific discovery in HPC environments.

Link to working group:

- For further information, please have a look here: https://www.nhr4ces.de/cross-sectional-groups/csg-data-management/.

Professional prerequisites for applicants:

- Master's degree in computer science or a related field, with a strong background in data management and AI.
- Experience with large language models, tabular data processing, and deep learning frameworks (e.g., PyTorch, TensorFlow) is expected.
- Solid knowledge of data management and database systems is required.
- Familiarity with HPC environments or scientific workflows is a plus.
- Strong programming skills and proficiency in English are essential.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Dr. Carsten Binnig, TU Darmstadt, Professor on Data and AI Systems.

9#NHR4CES: Integrating Vector Search into HPC Infrastructure

Description:

The PhD position on Integrating Vector Search into HPC Infrastructure focuses on rethinking vector search systems, crucial for applications such as similarity search in scientific data, machine learning, and AI, by leveraging the unique capabilities of high-performance computing (HPC) environments. Traditional vector databases typically do not fully exploit HPC hardware features such as high-throughput interconnects, parallel file systems, and
large-scale GPU clusters. This research will investigate how to redesign vector search algorithms and systems to natively integrate with HPC infrastructure, optimizing for low-latency communication, efficient use of hierarchical storage, and scalable GPU acceleration. The goal is to enable fast, large-scale similarity search over high-dimensional vectors within scientific workflows, thereby expanding the role of vector search in data-driven HPC applications.

Link to working group:

- For further information, please have a look here: https://www.nhr4ces.de/cross-sectional-groups/csg-data-management/.

Professional prerequisites for applicants:

- Applicants should have a master's degree in computer science or a related field, with strong skills in systems programming and parallel computing.
- Experience with vector search, GPU programming (e.g., CUDA), and HPC technologies such as MPI, fast storage, and high-speed networks are highly desirable.
- Familiarity with machine learning and data-intensive applications is a plus.
- Proficiency in English and strong programming skills are required.

Working language/required language skills:

- English.

Description of supervisor and scientific institution:

- Prof. Dr. Carsten Binnig, TU Darmstadt, Professor on Data and AI Systems.