ANNUAL REPORT 2025

ZIB
ZUSE INSTITUTE BERLIN

Preface

We hope you enjoy reading!

Your Members of ZIB

EXECUTIVE SUMMARY

2025 marked a year of strategic consolidation and renewed momentum for the Zuse Institute Berlin (ZIB), underscoring its central mission to bridge application-oriented mathematics, artificial intelligence, data-driven methods, and high-performance computing (HPC). Guided by the question of how AI and learning can be combined with mathematical and computational methods to drive innovation and enable optimal, trustworthy decisions in the real world, ZIB sharpened its profile as an institute that connects modeling and optimization, learning and AI, and scalable computing to address complex challenges in science, industry, and society. Amid rapid technological transformation, ZIB demonstrated scientific depth, cooperative strength, and strategic foresight.

A defining achievement of the year was the successful outcome of several proposals that are central to ZIB's institutional framework and long-term development. The continuation proposals for the Cluster of Excellence MATH+, the Research Campus MODAL, and the National High-Performance Computing Center (NHR) at ZIB were all successful, and funding for the next periods was granted. These successes reflect the strength of ZIB's scientific contributions, the quality of its infrastructures, and the reliability of its partnerships with universities, research institutions, industry, and public-sector actors. Moreover, two further proposals of major strategic importance—the extension proposals of the Mathematical Research Data Initiative (MaRDI) and the Berlin University Alliance—were also successful. Together, these decisions secure important pillars of ZIB's research, service, and transfer activities, and they provide a strong foundation for the coming years.

Significant progress was also made in further developing ZIB's high-performance computing, AI, and data infrastructure. Building on the successful integration of the Lise HPC system into the NHR network as a Tier-2 CPU cluster, ZIB expanded its GPU computing capabilities with a new GPU system that provides dedicated resources for AI and machine learning workloads. The ITDS platform was further enhanced with new services, including LLM and analytics services, while the Z1 and Z3 storage systems continue to provide reliable data management for the institute's research community. A unified platform concept for HPC, AI, and data analytics is being developed to streamline access to computing resources and improve the user experience across all ZIB services. This infrastructure modernization effort also includes the implementation of advanced data management tools and workflows that support the growing demand for data-intensive research, as well as the establishment of tiered storage solutions that balance performance, capacity, and data lifecycle management. Beyond hosting and serving existing models, ZIB is now actively deploying large language models on its own infrastructure to support research workflows, internal services, and collaboration partners – reinforcing the institute's commitment to technological sovereignty in AI. Furthermore, ZIB is taking the necessary steps to train its own foundation models, in particular in the context of remote sensing and scene generation for autonomous driving, where domain-specific pre-training on proprietary and sensitive datasets is essential and cannot rely on off-the-shelf commercial solutions.

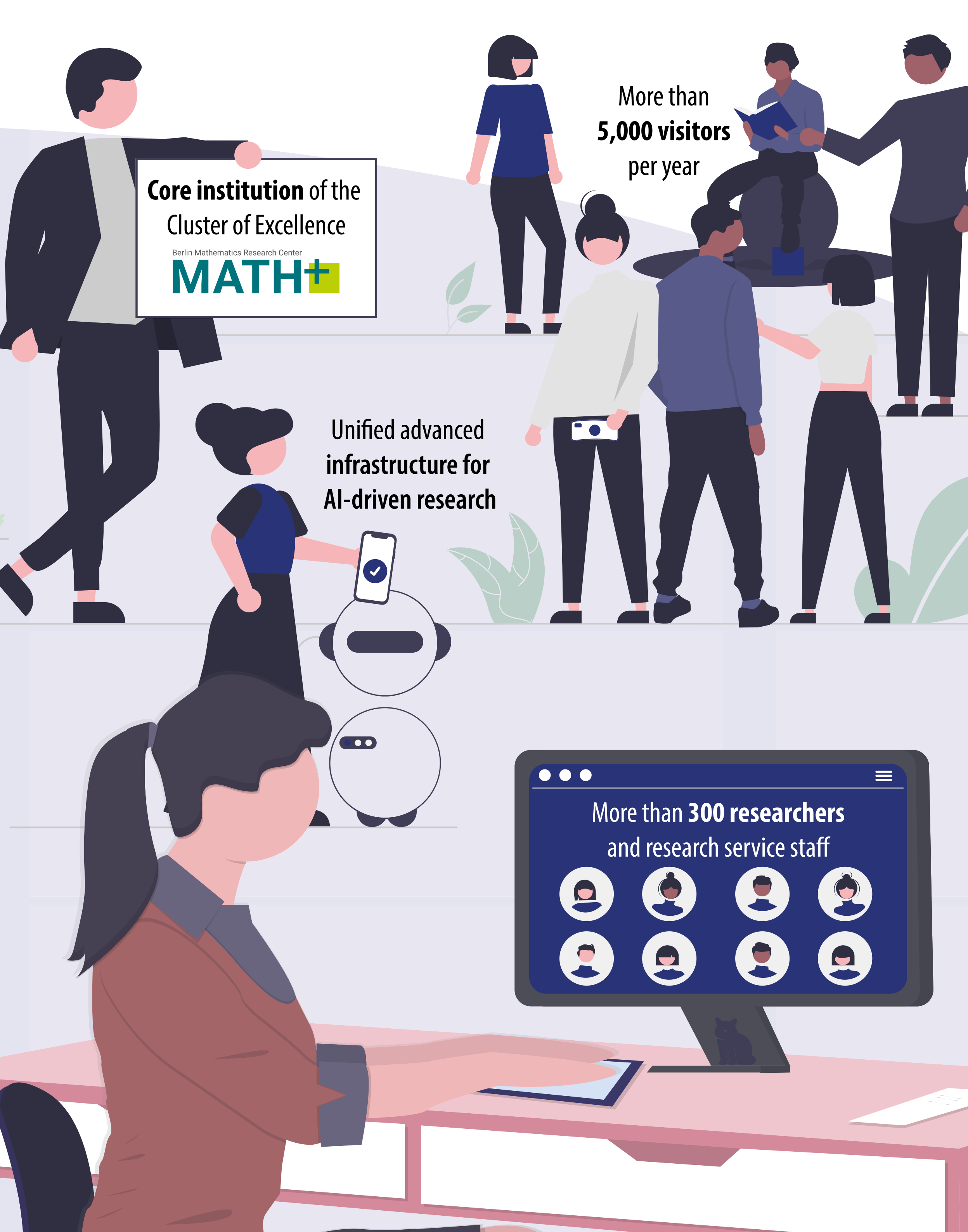
Scientific progress continued at a high pace, with a particular emphasis on research at the intersection of AI, mathematical modeling, and application domains. This year's highlights were manifold; the following three examples illustrate the wide thematic spectrum: As a first example take our wide-spread activities regarding the emergence of AI agents in research, where autonomous and semi-autonomous systems can support scientific exploration, workflow automation, and reproducible computation while raising new questions of reliability, verification, and human oversight. Simultaneously, ZIB also advanced work on opinion dynamics, developing agent-based and mean-field models that connect individual interactions with societal-scale patterns such as consensus, polarization, and echo chambers. By coupling opinions, social networks, uncertainty, and external actors such as media and influencers, this research helps illuminate mechanisms that shape public discourse and provides mathematical tools for analyzing interventions that may promote more stable and inclusive debate. In health and life sciences, AI for drug design exemplifies ZIB's ability to combine machine learning, simulation, optimization, and HPC to accelerate discovery and support data-driven decisions in molecular research.

The report is further contextualized by nine reports on different scientific developments at ZIB, ranging from topics like “opinion dynamics”, and “the mathematics of crime” via “bone dynamics”, “unravelling patterns of life”, “benchmarking AI and quantum methods” and “rolling stock scheduling”, to “AI for differentiated learning”, “taming gigantic data flows” and “efficient use of smarter hardware”. These reports are complemented by in-depth interviews with Tim Conrad and Marco Reidelbach on MaRDI and the future of mathematical research data; with Sebastian Pokutta, Andrea Walther, and Christof Schütte on MATH+ and its role in Berlin's research ecosystem; and with Ralf Borndörfer and Sebastian Pokutta on the third phase of MODAL as a long-standing bridge between mathematical research, AI, and practical transfer.

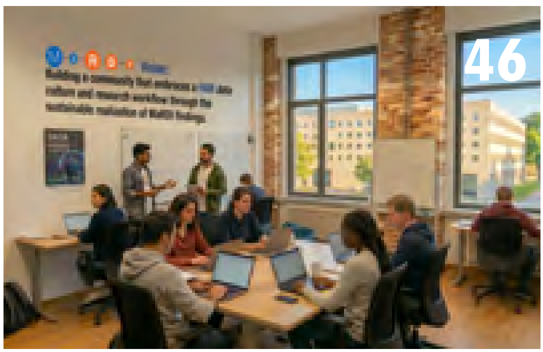
As reflected in this report, 2025 was a year in which ZIB secured essential funding lines, expanded the foundations of its computing and data infrastructure, and sharpened its scientific profile around the interplay of modeling, learning, and high-performance computing. Looking ahead, ZIB remains dedicated to advancing interdisciplinary research, empowering scientific communities in Berlin and across Germany, and translating mathematical and computational innovation into lasting impact for society.

Christof Schütte (President)	Sebastian Pokutta (Vice President)	Andrea Walther (Vice President)

ZIB IN NUMBERS


Core institution of the Cluster of Excellence
Berlin Mathematics Research Center
MATH+
More than **5,000 visitors** per year
Unified advanced **infrastructure for AI-driven research**
More than **300 researchers** and research service staff

CONTENTS



PREFACE

SPECIAL@ZIB – INTERVIEWS

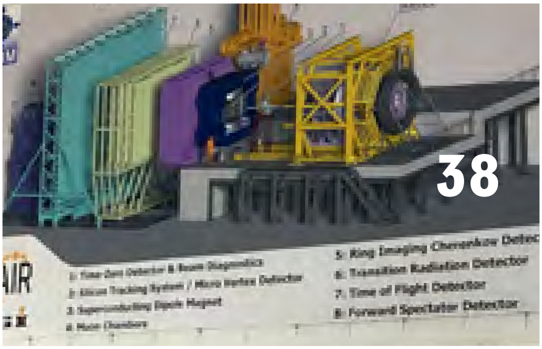
APPENDIX

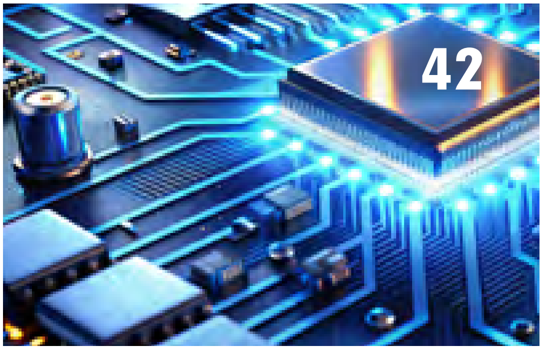

14


18

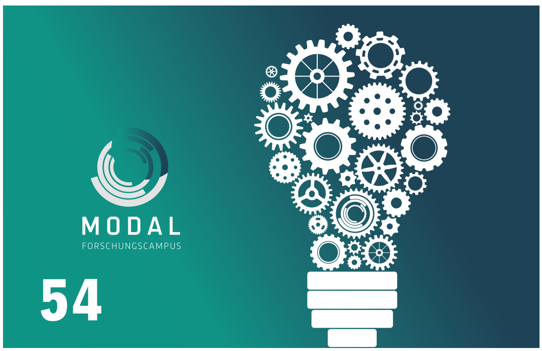

26


30


38
6: Transition Radiation Detector
7: Time of Flight Detector
8: Forward Spectator Detector


42


50
Berlin Mathematics Research Center
MATH+


MODAL
FORSCHUNGSCAMPUS
54

The Math of Crime

If you've ever watched CSI, Bones, or Silent Witness, you've seen the familiar moment where someone points at a thermometer and confidently narrows down the timeline: "time of death: around..." In real homicide investigations, that question is just as central – because pinning down when a person died can tighten timelines, test alibis, and make or break the direction of a case. One of the most trusted day-to-day tools is still surprisingly down-to-earth: estimating time since death from how the body's core temperature has cooled by the time it's found. The catch is that, outside TV studios, cooling is messy – clothing, body mass, the environment, and changing conditions can all push estimates around and leave frustratingly large uncertainties. In collaboration with forensic experts at Universitätsklinikum Jena, researchers at ZIB are therefore developing mathematical methods that make temperature-based time-of-death estimation more accurate and that work reliably across a wider range of real-world scenarios.

Cooling models

Established phenomenological methods postulate a core temperature cooling curve $T(t;p)$ parameterized with parameters p and estimate the time since death t_* from a rectal measurement T_m via $T(t_*;p) = T_m$; see Fig. 1. The dominant cooling curve in use is a double-exponential function that is a simple yet surprisingly accurate approximation of an eigenmode representation of corpse cooling. The parameters p are defined by expressions depending on body mass, clothing, and environmental temperature, which have been fitted from phantom experiments as well as actual measurements.

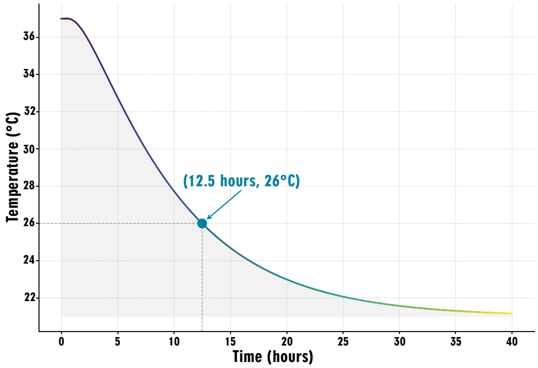


Figure 1: Basic estimation of time since death using a cooling curve. The intersection of the curve with the measured rectal temperature (y-axis) yields the time since death (x-axis).

In contrast, mechanistic models describe cadaver cooling in terms of physical processes, in particular the heat equation. Their parameters are well-understood physical quantities that can, in principle, be measured and allow the cooling model to be adapted to any situation at the crime scene. Benefiting from broader applicability and better interpretability, mechanistic cooling models promise higher accuracy and relevance, but require more complex usage involving segmentation, meshing, and finite element simulation; see Fig. 2.

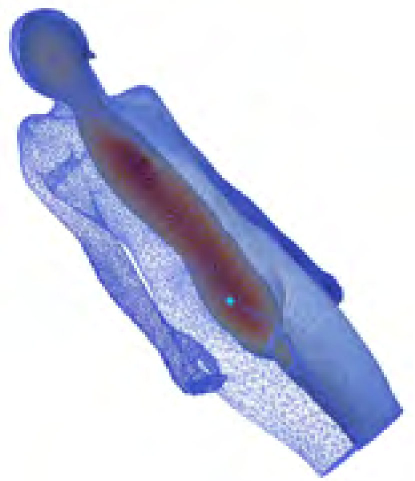


Figure 2: Finite element simulation of temperature distribution inside an actual corpse geometry. The rectal measurement location is marked in light blue.

Sensitivity analysis and design of experiments

Though in principle measurable, most parameters of mechanistic models are not available in practice, but have to be postulated or estimated simultaneously. We conducted sensitivity studies identifying the most important parameters, in particular heat capacity of water-dominant tissue, heat transfer coefficient to the environment, and time course of environmental temperature. Variations in individual anatomy and sensor position in the rectum each induce a variation in estimated t_* of around 20%. Overall, both phenomenological and mechanistic cooling models yield a marginal variance in the estimated time since death of 25% or more.

Thus, acquiring additional measurement data is necessary. We performed a design of experiment showing that multiple sequential temperature measurements, or at different positions, can reduce the uncertainty in the time estimate by a factor of five or more, assuming a compatible model. A Monte Carlo sampling showed that the posterior density is unimodal and very close to a normal distribution, enabling the use of simple maximum posterior estimates for uncertainty quantification and experimental design.

The time course of the environmental temperature is difficult to extract from corpse cooling, as it interferes with the time since death. We investigated the use of temperature measurements in different objects such as cupboards or piles of books for identifying time and amplitude of a sudden temperature decrease in the practically relevant case of a window being opened. Our results show that both can be identified with satisfactory accuracy within a four-hour time period before measurement.

Towards practical applicability

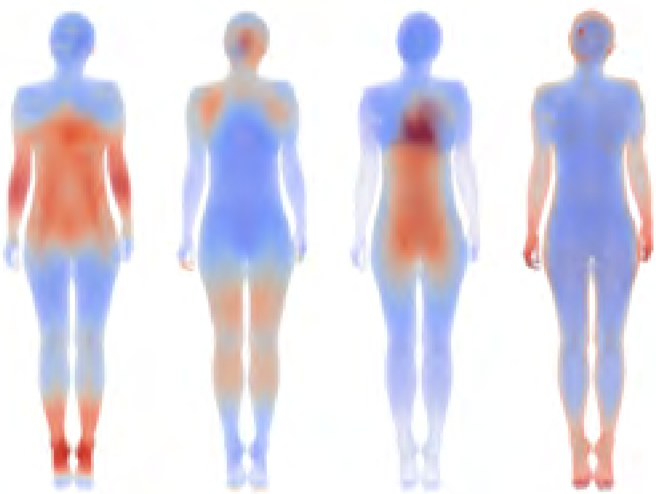


Figure 3: Orthogonal basis functions replacing the finite element basis in the reduced model.

The complexity of the mechanistic models' workflow, including segmentation, meshing, and finite element simulation, prevents their widespread adoption in practical casework. We therefore work on simplified models that still capture the characteristics of the full finite element cooling model.

Reduced basis approaches turn the high-dimensional discretized heat equation into a low-dimensional system of ordinary differential equations with excellent approximation of the temperature distribution thanks to a problem-adapted basis; see Fig. 3. Linearity of the cooling process and affine dependency on many parameters allow for precomputing of system matrices for a wide range of parameters. Such models are flexible enough to represent time-dependent parameters as well as multiple measurement locations.

Even simpler are machine-learning models mapping body features such as mass, height, waist circumference, sex, and clothing to parameters of a double-exponential cooling model. For the required explainability, we use Gaussian process regression (GPR) and restrict the input to the most relevant features. Good agreement with a dataset of real-world cooling cases has been achieved by selecting the four most important features.

Martin Weiser

OPINION

DYNAMICS

Understanding how opinions form in a connected world

Opinions play an important role in every society, shaping elections, public health responses, climate policy and social cohesion. However, with the rise of online social media, influencers and algorithmic content curation, the way opinions form and spread has become complex. Understanding these processes is one of the central challenges of modern society.

While large-scale data can reveal what people share across social networks, they often fail to explain how group behavior develops, such as collective agreement, strong divisions or "echo chambers", where people primarily interact with like-minded opinions. Agent-based models (ABMs) of opinion dynamics offer a complementary perspective by representing societies at the level of individual agents, whose behavior and interactions follow formal rules. These models reveal how local interactions between people can add up to large-scale societal trends and help us identify key mechanisms that shape public discourse and collective opinion change. Most traditional ABMs typically focus only on how opinions are exchanged, often assuming fixed social networks or deterministic rules governing how opinions are updated. In reality, however, opinions do not evolve in isolation – they co-evolve together with social dynamics. People move through social environments, cluster with like-minded others, react to uncertainty, and are influenced by powerful external actors such as media and influencers. Our research at ZIB addresses this phenomenon by devising models in which opinions, social interactions, uncertainty and external influences are coupled together.

Our work highlights that how people interact is as important as the content of their opinions. People continuously adjust whom they follow, trust or engage with, often due to differences of opinion. To capture this in our models, we have introduced a co-evolution mechanism: opinions influence social connections and social connections influence opinions. We find that this can lead to formation of opinion-homogeneous groups that reinforce shared views through repeated internal interactions, making global consensus unlikely. This helps explain a familiar real-world phenomenon: despite widespread access to the same information, polarization can persist or even grow stronger. Our model also reproduces patterns seen in data from the General Social Survey in the U.S., which focuses on political identity and opinions about government policies. It shows how co-evolution of people's identities

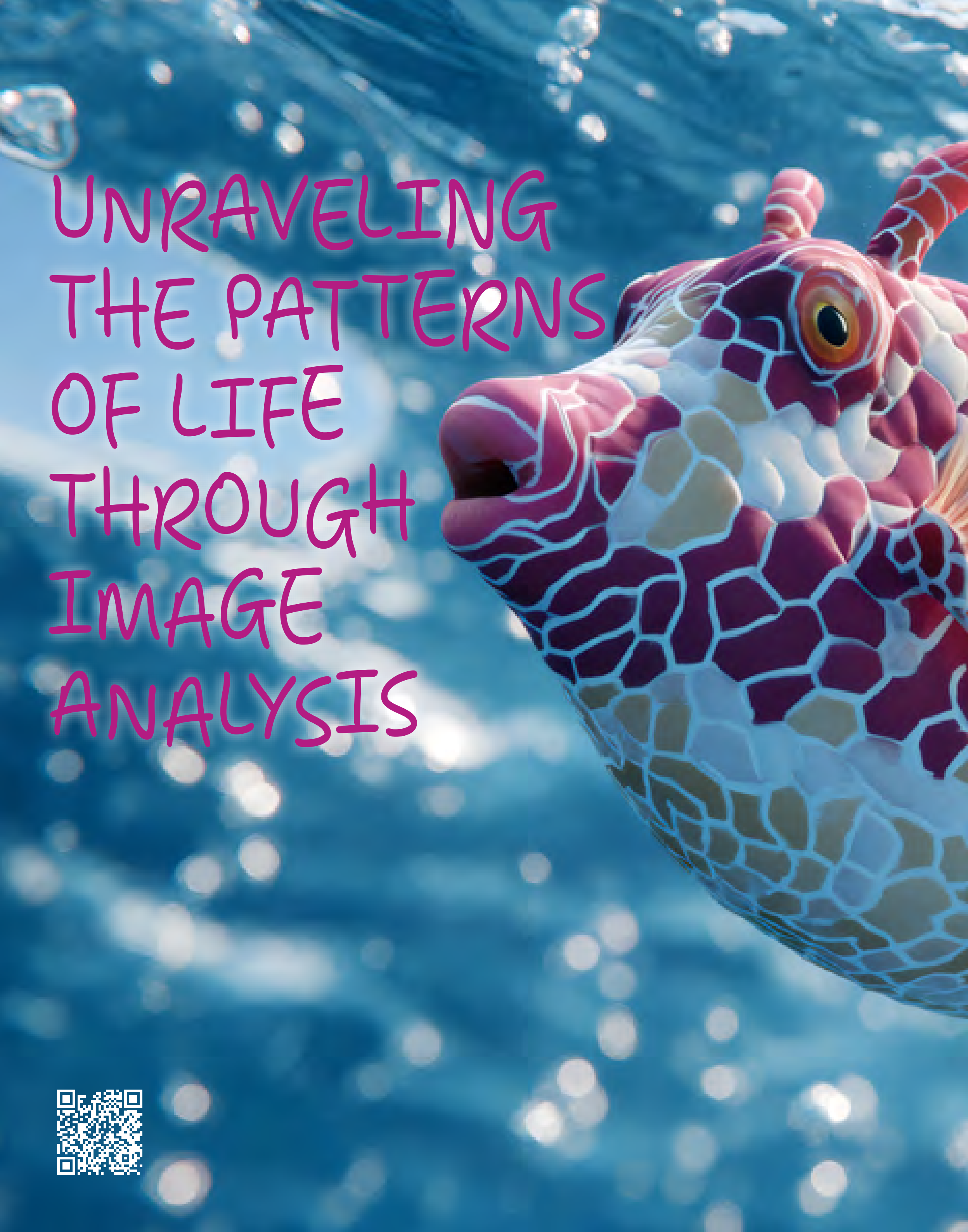
and opinions can influence each other and lead to the formation of separate groups of like-minded people, reflecting the kind of polarization observed in the survey. To study such dynamics in large populations, we derive reduced, so-called mean-field models, which describe how opinion clusters and social structures evolve at the collective level. These tools make it possible to connect individual interactions with large-scale social patterns.

In collaboration with experts in the field, such as colleagues from the Max Planck Institute for Human Development, we also study online opinion landscapes. These are shaped not only by interactions between individuals, but also by external actors, like media and influencers. We therefore extend classical opinion models to include such actors, who can reach large audiences and who adapt their actions strategically to maximize their impact and followership. As observed in real online data, our results show that such dynamics can give rise to fragmented, like-minded opinion clusters that change over time, often driven by the actions of influential agents. To better understand these processes, we develop methods to identify clusters, track how they change in time and detect when critical changes in collective behavior occur. We also formulate and study optimal strategies for media or influencers to guide these dynamics, for example, to reduce societal divisions, limit the growth of extreme groups, and promote a more stable and balanced public debate.

At ZIB, this research contributes to a broader effort to use mathematical modeling and large-scale computation to understand socially relevant processes. Our aim is to support deeper insight into the mechanisms shaping public discourse and to inform long-term thinking about how to foster more resilient and inclusive societies.

Nataša Djurdjevac Conrad

UNRAVELING THE PATTERNS OF LIFE THROUGH IMAGE ANALYSIS



The advent of the microscope in the 17th century marked the beginning of significant advances in imaging technologies. Since then, we have witnessed transformative progress, from the pioneering techniques of X-ray imaging and computed tomography to more sophisticated methods like magnetic resonance imaging, modern light microscopy, and electron microscopy. These advances now allow scientists to visualize the intricate three-dimensional organization of cells, tissues, and organs with unprecedented precision. This capability has greatly expanded our knowledge of biological systems, revealing how components interact, function, and adapt in response to various stimuli. By deciphering the composite structures that underlie life, imaging techniques have shed new light on fundamental questions about the organization and evolution of biological systems.

However, extracting valuable insights from these images remains a major challenge, as modern imaging produces vast, complex datasets that cannot be reliably analyzed by visual inspection alone. To meet this challenge, researchers at ZIB are developing sophisticated image analysis tools and methodologies that enable experts from biology and other scientific fields to extract detailed information from their images. By leveraging cutting-edge computational, mathematical and artificial intelligence (AI) techniques and collaborating closely with domain experts, we develop intuitive and user-friendly software solutions that facilitate efficient data analysis, enabling our collaborators to unlock new knowledge about the sophisticated structure and function of living organisms.

Nature is replete with intricate patterns, one of which is tessellation – a ubiquitous motif in which organisms deploy arrays of rigid tiles or units to achieve a range of purposes. These patterns can be found in diverse environments, from structures that enable optics and water transport in certain species, to the protective armor plating of certain invertebrates. A prime example of this phenomenon is seen in cartilaginous fishes, such as sharks and rays, whose skeletons are characterized by a distinctive tessellated arrangement. To tackle the challenge of segmenting such large-scale tessellations, we introduced a novel method called random-walk distance transform that uses random walks on binary image data to distinguish individual tiles of the tessellation. This has shown remarkable efficacy in distinguishing between different elements. Originally developed for cartilaginous fish skeletons, in further collaborations we were able to show its versatility and effectiveness across a range of datasets such as the internal skeletons of starfish, the mineralized sclerites of gorgonian corals, and the armor plating of boxfish.

In a project with geologists from the Center for Marine Environmental Sciences (MARUM, Bremen), we analyzed cold-water corals, which build vast coral mounds representing huge carbon reservoirs that, when resolved, may add substantially to the amount of carbon dioxide in the atmosphere. The long-term goal of the project is the automatic taxonomy of coral species from fragments found in these mounds, which were constructed over hundreds of thousands of years. Being able to efficiently identify coral species from imaging data will enable a greatly improved understanding of the environmental conditions in the past – up to 200,000 years ago. This further adds to our understanding of climate change and its risks. In the first part of this ongoing project, we have developed the Coral Analyzer, a tool that supports the analysis of visual data relating to coral colonies at an unprecedented level of detail, allowing for the full reconstruction of the growth pattern of the colony. Crucially, this tool is not restricted to corals. It can be applied to other dendroidal structures including distributed sensory networks of chitons, which are analyzed in another project currently carried out in the Visual Data Analysis group.

Advances in imaging technologies necessitate innovative image analysis solutions that harness the power of math and AI. By merging these disciplines, we will unlock novel methods for automated insight, pattern identification, and informed decision-making – catalyzing breakthroughs that reshape our understanding of life.

Daniel Baum

UNDERSTANDING THE DYNAMICS OF BONES

Bones may appear rigid and unchangeable, but in reality they are living systems that continuously adapt to their environment. Nutrition, mechanical load, disease, trauma, and even gravitational conditions influence how bone tissue grows, remodels itself, and heals. At ZIB, researchers study human bones not only as static structures, but also as dynamic systems that respond to external influences across time and scale. In close collaboration with domain experts in medicine, biomechanics, and biochemistry, ZIB researchers address the question of how bone remodeling and healing unfold across scales – from microscopic cellular processes to changes in macroscopic structure.

To address these questions, a wide range of imaging techniques such as microscopy, spectroscopy, X-ray imaging, and magnetic resonance imaging are used, each dedicated to particular tissue types and properties, providing information at different spatial and temporal scales. The resulting data form the basis for mathematical and computational models that are validated against experiments and clinical observations with the aim of predicting bone growth, bone healing, or bone remodeling under specific conditions. At ZIB, part of our research focuses on improving the imaging itself. This includes not only three-dimensional reconstruction of bony anatomy from conventional two-dimensional X-ray images, but also advanced MRI reconstruction methods that significantly improve acquisition times.

Another research focus lies on image analysis, where we develop computational methods to extract quantitative and clinically meaningful information from imaging data. Beyond building efficient software tools for visualization and measurement, our work increasingly targets new statistical and artificial intelligence methods that take into account the mathematical structure of image-derived objects. For example, the relative position, orientation, and motion of bones or bone fragments are naturally described by geometric transformations that follow specific algebraic and symmetry properties. Standard AI methods typically ignore this inherent structure, treating such data as generic numerical inputs. In contrast, our approaches are designed to adhere to these symmetries and geometric constraints, leading to models that are more data-efficient, physically consistent, and better suited for capturing biologically meaningful variation. This integration of geometry, statistics, image processing, and machine learning provides improved analytical power for studying skeletal configurations, detecting subtle structural changes, and ultimately supporting predictive modeling in clinical applications.

Besides transformations, mechanical stresses play a central role in bone adaptation and fracture healing. Researchers distinguish between static loads, such as standing or maintaining posture, and dynamic loads caused by movement or sudden impacts.

When a person stumbles, anatomical structures can undergo forces of up to eight times body weight; in high-impact sports or high-speed accidents, even greater loads occur. Measuring these forces and understanding how bones, joints, and surrounding tissues respond to them are active areas of research, providing insight into injury mechanisms, overload-related damage, and the conditions that promote healthy adaptation. At the macroscopic level, such observations also support hypotheses regarding microscopic processes, including mechano-sensitive cellular responses that regulate bone growth, healing, and remodeling.

Mechanical stresses and strains are equally crucial in medical therapy planning for fracture treatment. While slight compression promotes healing, shear motion can delay it, making implant selection and positioning critical. However, the movements occurring in different fracture patterns remain poorly understood. Dynamic X-ray imaging (video fluoroscopy) enables the recording and evaluation of motion in fracture zones under various treatment strategies. At ZIB, software tools are being developed for automated analysis and precise quantification of three-dimensional strains in fracture areas. These findings aim to improve therapy by guiding implant choice, positioning, and screw configurations to encourage beneficial compression while avoiding harmful shear forces.

Our research at ZIB is carried out in close collaboration with clinical partners, including physicians from major medical centers such as Charité – Universitätsmedizin Berlin.

Clinical questions and patient data guide the research directions, while at ZIB the emphasis lies on developing robust mathematical models, algorithms, and data-driven analysis methods. Together, these efforts aim to transform complex imaging and measurement data into reliable predictions that can support diagnosis, treatment planning, treatment, and long-term patient care.

Christoph v. Tycowicz, Stefan Zachow

Automated Rolling Stock Scheduling with Predictive Maintenance

Reliability is one of the key requirements for a functioning public transport system. This does not only mean that passengers can trust the communicated timetable, but also encompasses the availability of rolling stock itself. Facing cost increases, staff shortages, and infrastructure bottlenecks, railway undertakings have to use their resources as efficiently as possible. In particular, it is desirable that the rolling stock be scheduled in such a way that the fleet size can be kept small, "deadhead" trips without passengers are avoided, and small disturbances can be absorbed by sufficient buffer times. While all these considerations have been extensively discussed in the research landscape, they become pointless when vehicles break down.

Maintenance Strategies

To prevent technical problems during operations and to guarantee safety standards, rolling stock has to undergo regular maintenance. For railway vehicles, the standard procedure is to examine technical components after a certain time has passed or a given distance has been traveled. Such a preventive strategy can lead to unnecessary idle time of vehicles due to scheduled but unneeded maintenance when the actual condition is still satisfactory. Yet risks remain, as severe wear and tear might not be recognized in time.

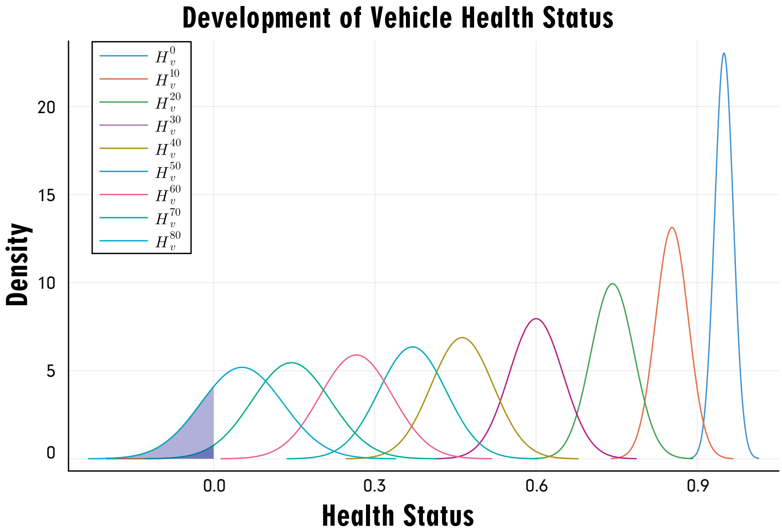


Figure 1: Probability density function of the predicted health status over time for selected normal distributions.

This can be overcome by a condition-based approach, where the health status of components is regularly tracked, for example by using sensor data. However, this might lead to maintenance slots being scheduled at very short notice, seriously impacting operational stability. We therefore propose a predictive maintenance strategy that integrates estimates of the time of failure ("remaining useful life") into rolling stock scheduling so that, with high probability, vehicles are brought to workshops exactly when needed.

BerDiBa – Berliner Digitaler Bahnbetrieb

In cooperation with Siemens Mobility, TU Berlin, and ten further partners, the primary goal of the BerDiBa project is to develop technologies for automated and remote train operation using digital twins with special emphasis on artificial intelligence methods. The task of ZIB is to implement a tool for automated rolling stock scheduling that includes predictive maintenance, thereby integrating statistical machine-learning methods with combinatorial optimization. The resulting framework is then evaluated on a test library based on the network of a German railway undertaking that operates regional trains, exploiting sensor data in this case to predict the health status of doors.

Mathematical Modeling

Rolling stock scheduling problems are classically modeled on a space-time graph. We enhance this graph by maintenance arcs, and add one or more dimensions to capture the current health status. As the resulting state-expanded event graph is discrete, even though we are dealing with continuous probability distributions for the remaining useful life, we can only solve an approximation. However, by defining a suitable rounding function, we can derive consistent bounds on failure probabilities, errors, and the optimal objective value. We complement this approach by an iterative refinement procedure and a multi-swap heuristic. To be able to cope with a realistic dynamic setting, where health status updates are revealed over time, we embed our method into a rolling horizon approach based on Bayesian inference and Kálmán filters.

Felix Prause, Dr. Niels Lindner,
Prof. Dr. Ralf Borndörfer

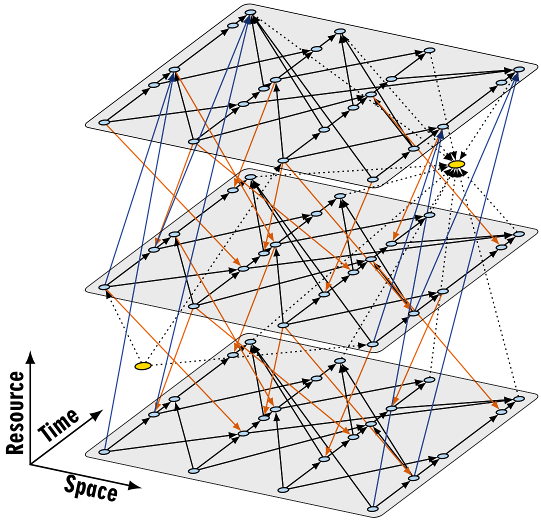


Figure 2: State-Expanded Event Graph. In this example, trip arcs (blue) degrade the health status, which is unaffected by waiting and deadhead arcs (black), and again restored by maintenance arcs (orange).

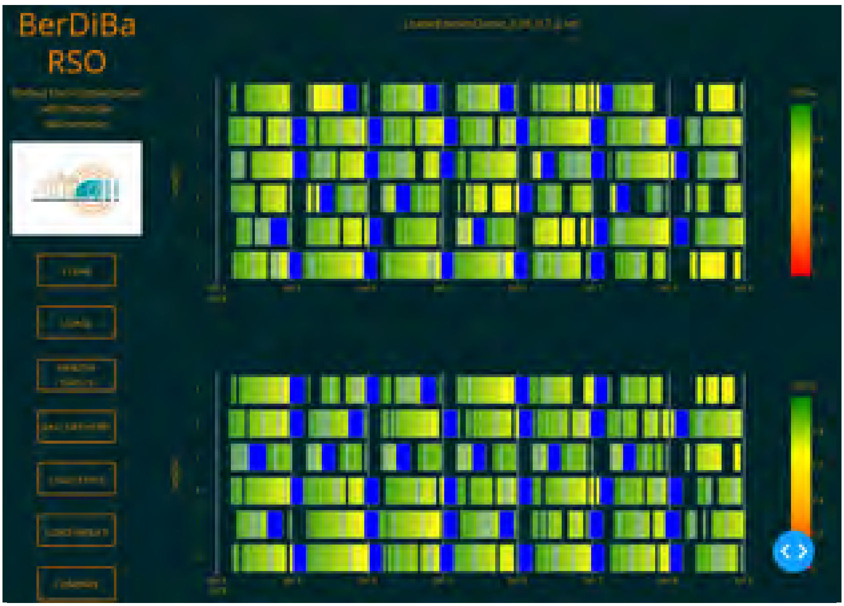
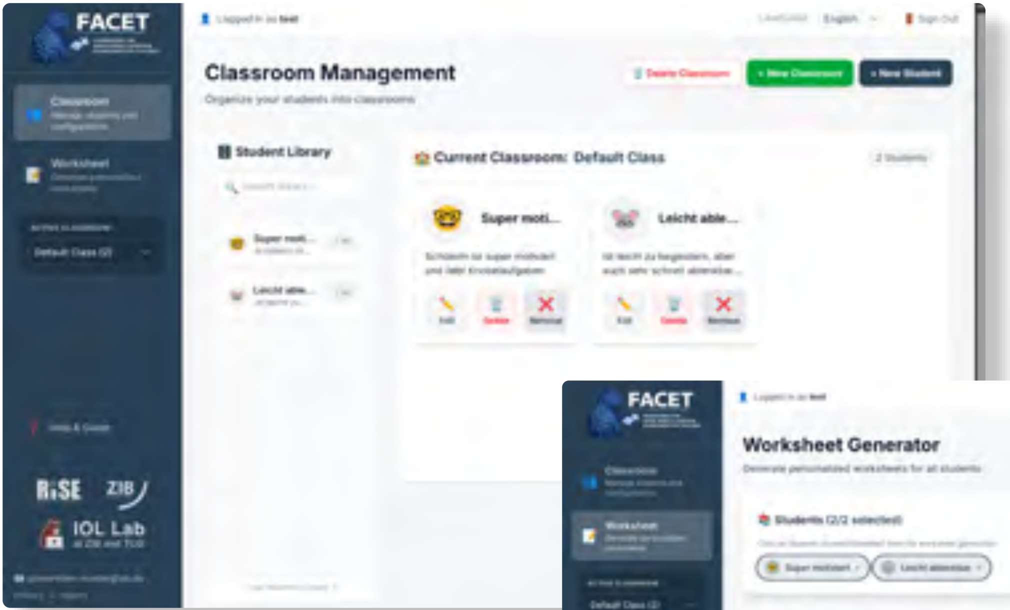


Figure 3: Interactive Gantt chart visualization of our predict-and-optimize framework, with rows corresponding to vehicles. In the original schedule (top), the health status is worse (yellow to red) than in our computed schedule (bottom). After having visited a maintenance facility (blue), the vehicle's health status is restored (green).

ONE SIZE FITS NONE – AI FOR DIFFERENTIATED LEARNING



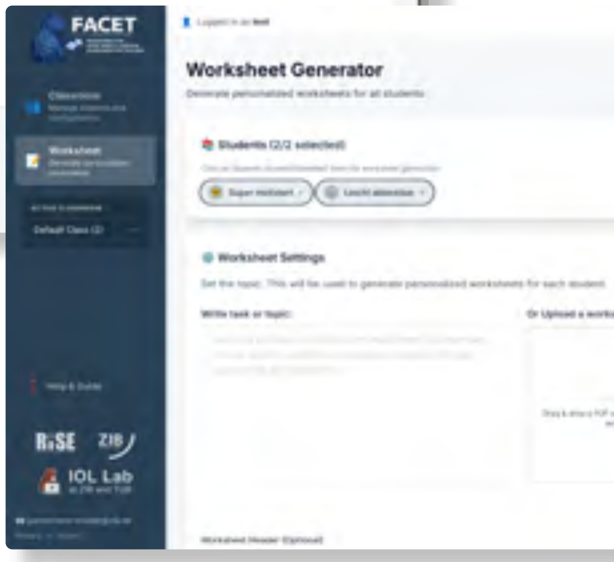
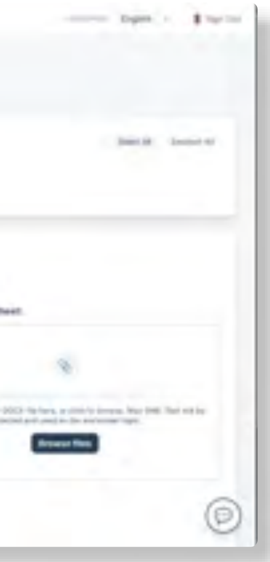


Figure 1: Current FACET interface. Teachers define student personas (learner agents) that reflect characteristics they observe in their classes, upload materials or select curriculum-based tasks, and generate differentiated materials, such as worksheets.

In the classroom, teachers face a widening gap between what pedagogy demands and what is feasible given limited time, large class sizes, and growing workloads. Students vary not only in prior knowledge but also in their motivations and individual needs (IQB Bildungstrend, PISA), including learning difficulties such as dyslexia or ADHD. Differentiated instruction addresses these differences through tasks at varying complexity, scaffolded support using hints and example solutions, and individualized feedback. OECD studies confirm that schools are facing increasing heterogeneity and complex demands, combined with teacher shortages that puts additional pressure on teachers across all member countries. Although there are already promising approaches to differentiation, the practical implementation is often challenging.

To address this need, our team developed a multi-agent tool called FACET (Framework for agent-based classroom enhancement for teachers), in close collaboration with Jennifer Haase from the Weizenbaum Institute and partner schools in Berlin, that supports teachers in developing differentiated materials for their classes. FACET coordinates four Large Language Model (LLM)-based agents built on open source LLM, hosted locally at ZIB via Ollama. Teachers create different student personas (learner agents) that reflect characteristics they observe in their classrooms, and either upload their own materials or select curriculum-based tasks. The learner agent generates natural-language thought traces simulating how a student approaches a task. A diagnostic agent then processes these into structured JSON assessments (JavaScript Object Notation, a data format for structured information exchange), analyzing cognitive processes and affective responses. The generator agent draws on both inputs to produce differentiated materials, embedding curriculum standards and evidence-based didactic principles throughout the generation process. To ensure systematic alignment with school requirements, the generator operates within a Retrieval-Augmented Generation (RAG) pipeline. This pipeline incorporates school-specific curricula, subject-specific tasks, and didactic standards, thereby grounding the output in educationally sound content. The agents generate different outputs, for example, materials following dyslexia guidelines for students with reading difficulties, those with puzzle-based tasks for high-achieving students, or materials with hints and worked examples for those needing additional support. The evaluator agent independently reviews outputs against the original learner profile and scores materials on pedagogy-based dimensions. The full cycle, consisting of simulation, diagnosis, generation, and evaluation, is completed in around 30 seconds per worksheet. Fig. 1 illustrates the current FACET interface.

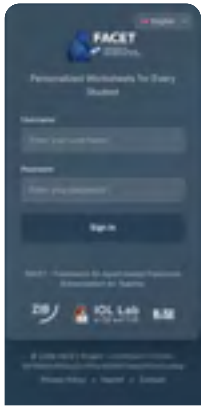


What distinguishes FACET from existing tools is its focus on differentiation beyond performance. Existing adaptive platforms differentiate based on performance data, such as test scores and homework completion. FACET extends differentiation to non-performance factors: teachers define learner profiles that include motivational orientations, learning difficulties, and linguistic backgrounds, while keeping the teachers in the loop.

We developed FACET together with teachers, who provided input on curriculum alignment, validated the didactic structure of the worksheets generated, and ensured that the tool's usability met practical classroom needs. We piloted FACET with 50 teachers across 20 schools in Berlin over the 2024-25 school year. Feedback consistently highlighted the urgent need for differentiation-oriented support and expressed appreciation for the time savings the tool provides, giving teachers more time to engage with their students. Teachers rated the quality of FACET's outputs as more thoughtfully constructed than materials that teachers were able to produce under time pressure. An additional survey with 70 K-12 teachers from the UK, the U.S., Canada, and Australia showed that the worksheets received good to very good ratings for didactic structure and task clarity, and improvements in suitability for learners with reading difficulties.

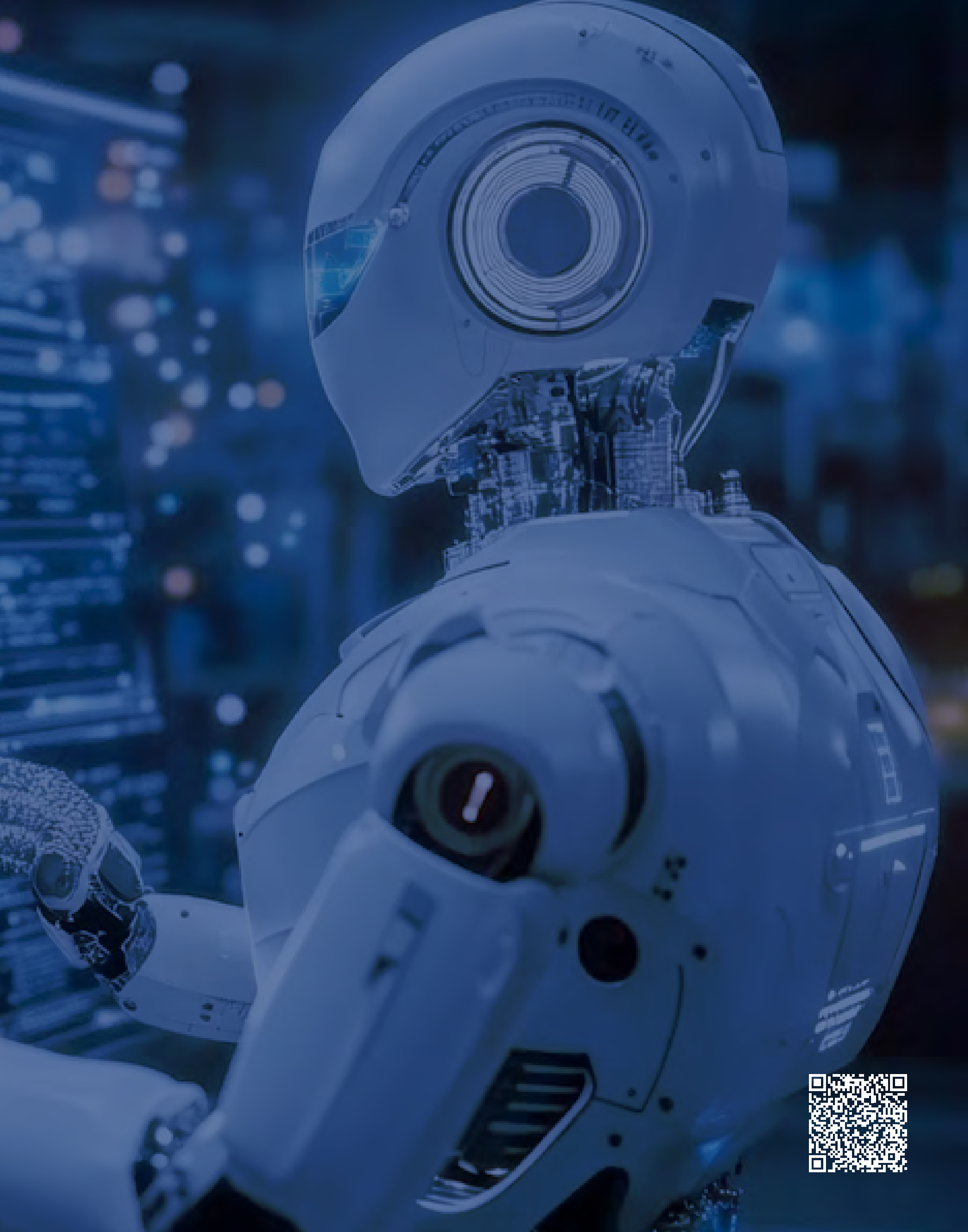
Current questions that we face are the extent to which accurately simulated learner behavior matches real learner behavior and how students perceive the differentiated materials. To address these questions, we have begun a longitudinal deployment of FACET across 10 schools in Berlin, building on our established partnerships with teachers. This ongoing study systematically measures teachers' and students' feedback, time savings, and learning outcomes, comparing results between FACET-generated and standard materials.

FACET can be explored, and teachers can request pilot access, on the following website: https://facet.iol.zib.de/login.



Jana Gonnermann-Müller, Konstantin Fackeldey, Nicolas Leins and Sebastian Pokutta

THE INTRACTABLE DECATHLON: BENCHMARKING NEW AI AND QUANTUM METHODS



Regularly, new hardware and software approaches emerge, with AI and quantum computers at the forefront, along with other systems such as data-flow machines, mem-computing, and bifurcation chips. All share the claim to "solve" challenging, i.e., NP-hard, combinatorial optimization problems much more effectively than traditional methods.

NP-hard problems are referred to as "intractable" in theory and, therefore, considered challenging to solve. As we summarized in *Challenges and opportunities in quantum optimization*, published in Nature Reviews Physics, this does not change when using a quantum computer.

However, stating that finding a solution is exceedingly difficult usually pertains to the decision version problem. In contrast, identifying some feasible solution to the optimization version of the problem is often straightforward. For instance, any permutation of cities constitutes a valid tour for the famous Traveling Salesperson Problem (TSP). The challenge lies in discovering the optimal tour and especially proving its optimality. In a blog post, NVIDIA used 896 GPUs to simulate a quantum system to compute a good – though not optimal – solution to a MaxCut problem, stating:

> *"In the math world, MaxCut is often cited as an example of an optimization problem no known computer can solve efficiently. MaxCut algorithms are used to design large computer networks, find the optimal layout of chips with billions of silicon pathways and explore the field of statistical physics. MaxCut is a key problem in the quantum community because it's one of the leading candidates for demonstrating an advantage from using a quantum algorithm."*

With our solver QuBowl, we need 0.2 seconds on a normal workstation to compute a solution of this quality. However, it takes us 2 days to find the optimal solution and another 3 days to prove its optimality.

The new approaches primarily can provide "good" solutions but fall short of proving optimality. The theoretical debate extends to whether and to what extent these problems can be approximated in polynomial time. However, the assurance an approximation algorithm offers is again merely a lower bound on the solution's quality. Since the worst-case time complexity of an entire class of problems says little about how difficult it is to solve a particular instance of that class, numerous questions remain unanswered, and ultimately, the only way to evaluate the practical performance of heuristic algorithms is to benchmark them on relevant instances.

But how to do this in a meaningful way? One cannot expect new methods to immediately beat everything developed over the last 70 years on problems designed for classical algorithms. Therefore, we selected model-independent instances from a diverse set of problem classes (see Table 1), which are known to be challenging for classic exact and heuristic methods. The problem sizes range from very small and easily solvable to currently unsolvable, even on supercomputers. This allows us to evaluate the performance of different approaches and track their progress. This is important because classical methods are not standing still. As we found, general solvers for integer programs improved by 43% each year over the last 40 years.

The *Quantum Optimization Benchmarking Library*, published in Nature Computational Science and available at github.com/ZIB-AOPT/QOBLIB, contains a more detailed description of our project, the rationale behind it, and the problem classes and instances. As part of our initial efforts, we provided reference solutions on classical hardware and ran three of the problem classes on existing quantum devices.

The call to the community is now out to provide better results on these challenging problems, by whatever means. There is no limitation on the hardware used or the algorithms employed. There is no preset formulation of the problems; we just provided the necessary data, examples, and solution checkers. These checkers can verify that a computed solution is indeed correct and feasible for the problem. In an ongoing effort, we are collecting new results and recording the progress made. The future will tell which methods prevail.

Thorsten Koch

Table and Picture

Table 1: Disciplines of the Decathlon

Problem	Description
Market Split	Multi-dimensional subset-sum
LABS	Low autocorrelation binary sequences
Birkhoff	Minimal Birkhoff matrix decomposition
Steiner	Steiner tree packing in graphs (VLSI design/wire routing)
Sports	Sports tournament scheduling (STS)
Portfolio	Multi-period portfolio optimization with transaction costs
Independent	Set unweighted maximum independent set (MIS)
Network	Communications network design problem
Routing	Capacitated vehicle routing problem (CVRP)
Topology	Graph topology design (node-degree-diameter problem)

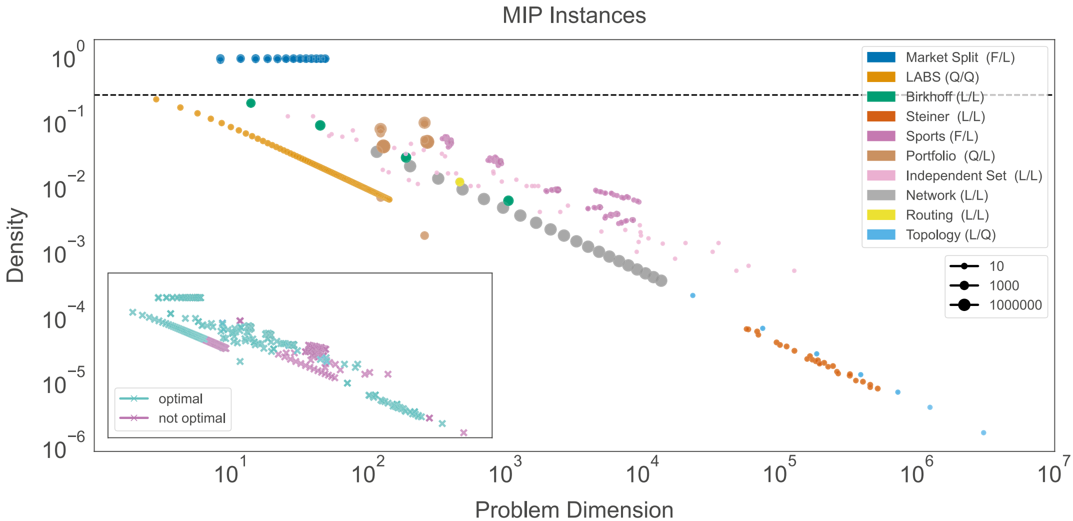


Figure 1: Sizes and density of the instances in the benchmark set. We start with small sizes that can be solved by classical methods on digital computers to optimality and progress to so-far unsolvable instances.

TAMING A PARTICLE ACCELERATOR'S DATA STREAMS WITH MODERN PROGRAMMING CONCEPTS


ed Baryonic Matter
HADES
1
2
3
4
5
6
Beam Diagnostics
em / Micro Vertex Detector
le Magnet
5: Ring Imaging Cherenkov Detector
6: Transition Radiation Detector
7: Time of Flight Detector
8: Forward Spectator Detector

The FAIR particle accelerator facility in Darmstadt provides physicists worldwide with a place to conduct research experiments with heavy ions. Among several experiments at FAIR, the Compressed Baryonic Matter (CBM) project with ZIB as collaboration partner aims to probe the interior physics of neutron stars, where extreme pressure causes conventional atomic shells to vanish and the nuclei themselves to form a novel state of matter. The experiment, which is currently under construction, is expected to generate data streams at a rate of up to a few terabytes per second, which must be aggregated and then analyzed for interesting events in near real-time.

Small fragments of the data streams from all detector subsystems are transmitted and assembled into analyzable data structures called timeslices across a dedicated compute cluster by the custom software system Flesnet, which leverages highly efficient remote direct memory access (RDMA) over Infiniband. Afterwards, the timeslices are forwarded to the data center's analysis cluster, routed through roughly 1-kilometer fiber-optic links.

At ZIB, and in collaboration with the CBM project funded by BMFTR, based on the experiences and challenges from the Flesnet and data forwarding use case, we are developing a new framework that provides a more seamless and highly flexible way to express and program data distribution systems using high-performance networks.

A new framework for efficient, flexible data distribution

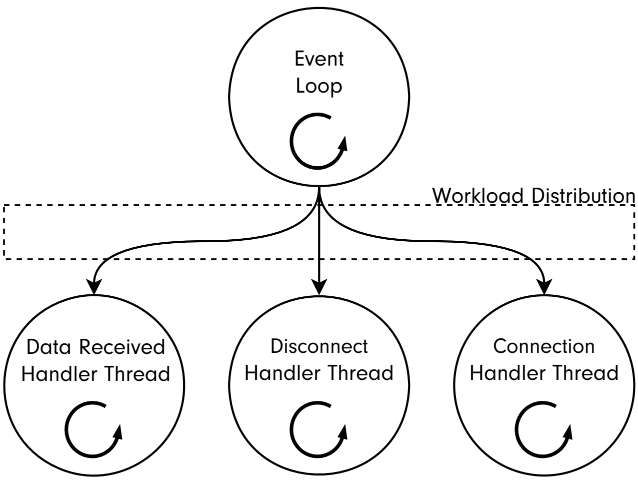


Figure 1: Workload distribution of the event loop to multiple worker threads.

In the CBM experiment, data distribution patterns and the need for high-rate data transmission streams arise in multiple contexts, each with distinct requirements. Guided by the "Don't repeat yourself" principle, we have developed a generalized toolset that enables the seamless implementation of the experiment's diverse data-processing pipelines within a single framework.

To maximize performance, we adopted a non-blocking, asynchronous software architecture. Asynchronous event handlers allow data transmissions and connection events to be processed concurrently in dedicated worker threads (see Fig. 1), enabling parallel execution and optimal use of the available hardware.

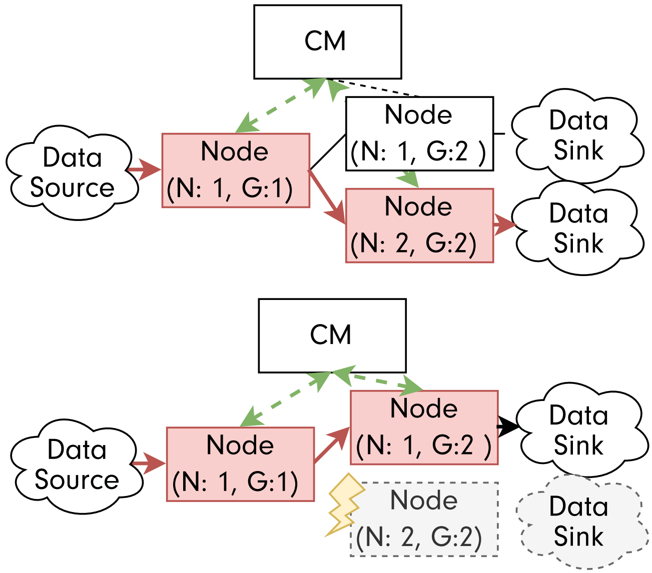


Figure 2: On a node failure, the chosen best transmission route changes dynamically.

Although asynchronous programming can be complex, our framework allows developers to express their business logic through interchangeable callback chains by modern C++ lambda expressions and to respond to events such as “new data available”, “connection established”, or “disconnected”. In this way, the framework adheres to the “Design for Change” principle, enabling runtime adjustments based on varying workloads.

An efficient and flexible buffer-management system, integrated into our framework, is essential for achieving high aggregate bandwidths. Consequently, each node can query the buffer status of every other node, transmit data, or send messages that are interpreted as commands. This flexibility allows for the construction of diverse network hierarchies and load-balancing schemas. Nodes may operate autonomously, or a central unit observes the network and computes optimal transmission routes (see Fig. 2). With inter-node messaging, the framework even supports the implementation of distributed fault-tolerant consensus algorithms such as Paxos.

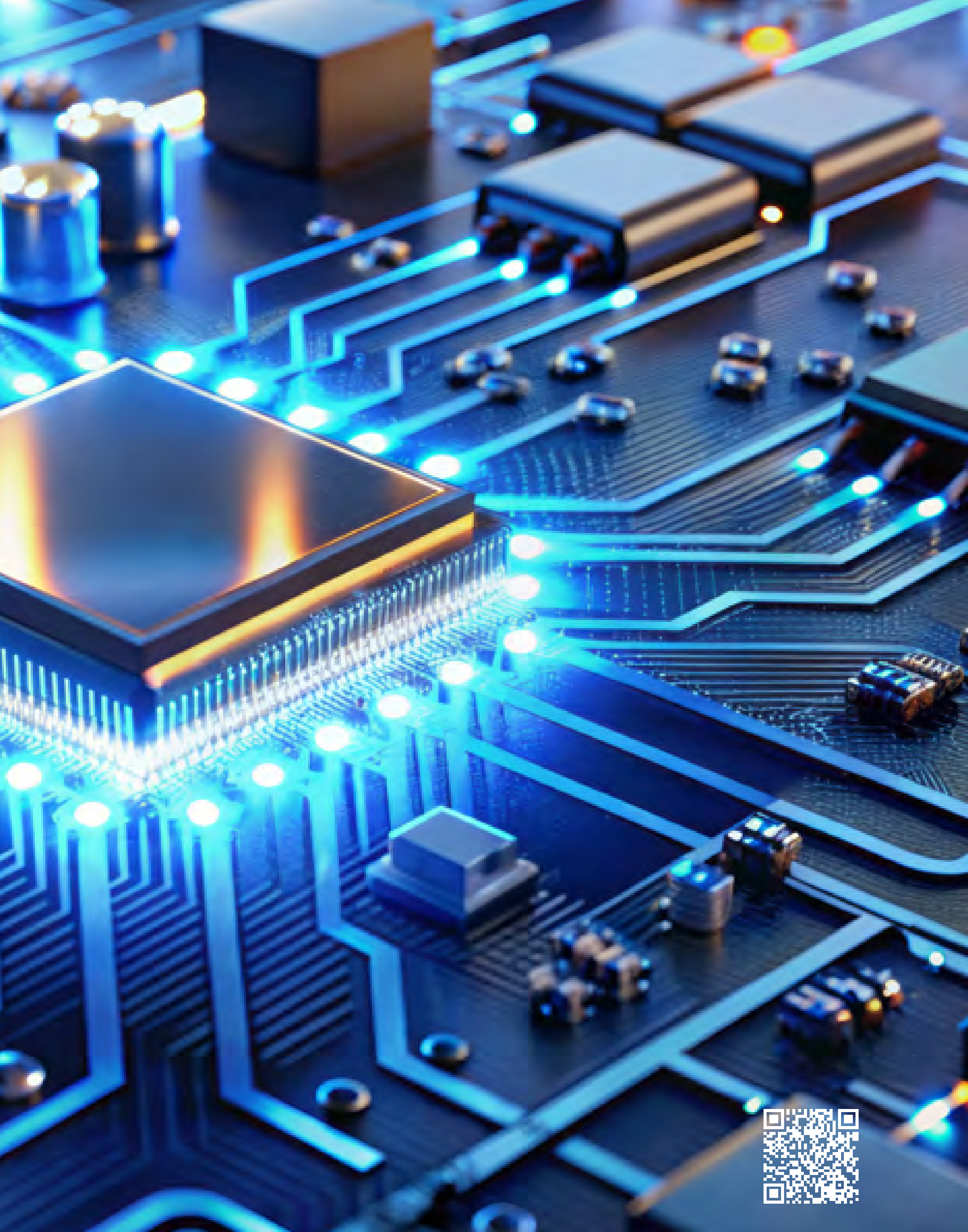
Because data may be consumed and freed in any order, the framework manages fragmented data buffers by means of a buffer map. The available data are represented by two doubly linked lists with entries for each data chunk in the buffer. One list is ordered by the entries’ relative age and the other is ordered by their physical arrangement in the buffer. Multiple buffer fragments can also be marked and treated as a single cohesive unit.

Although still under development, the framework already supports the forwarding of CBM timeslices to multiple nodes using variable forwarding patterns. It employs a network of nodes managed by a central orchestrator and utilizes RDMA over Infiniband for high-speed data transmission, while allowing nodes to be added dynamically during runtime.

Nico Greve, Florian Schintke

SMARTER HARDWARE, LOWER POWER CONSUMPTION:

HOW RECONFIGURABLE ARCHITECTURES BOOST ENERGY-EFFICIENT HIGH-PERFORMANCE COMPUTING



The Power Wall for Data Centers

Today, high-performance computing (HPC) and AI data centers are increasingly defined by their power consumption – often measured in megawatts – rather than by raw compute performance alone. Providers face mounting pressure to keep energy usage within strict limits while simultaneously meeting growing demands for faster processing of ever more complex data. This creates an inherent trade-off between achievable performance and sustainable energy consumption. Historically, improvements in processor technology and performance-per-watt metrics allowed increased compute capabilities within acceptable power envelopes. However, incremental advances are no longer sufficient. Achieving substantial gains in energy efficiency now requires disruptive approaches, including reconfigurable architectures combined with innovative algorithms.

Reconfigurable (RC) devices typically employ dataflow-oriented architectures and require a fundamentally different programming style than conventional CPUs. Instead of adapting software to fixed hardware, RC devices allow hardware itself to be tailored to the specific algorithm. Field-Programmable Gate Arrays (FPGAs), long used in other application domains, are a prominent example of such architectures. An FPGA consists of simple on-chip building blocks – such as lookup tables and integer multiply-add units – that are interconnected via an on-chip network according to the application's requirements. This principle of adapting hardware to the algorithm enables highly energy-efficient solutions, as only the functionalities required by a given algorithm are implemented. In addition, today's FPGAs provide high-performance network interfaces, which can eliminate the need for host systems to connect accelerators to a data center's fabric. In recent years, toolchains have emerged that allow FPGAs to be programmed using widely adopted high-level languages, such as SYCL (a C++-based standard) within frameworks like Intel oneAPI (see Fig. 1).

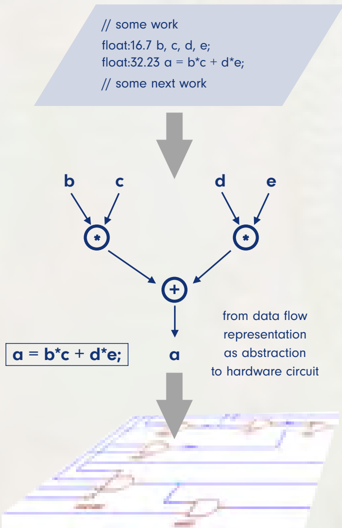


Figure 1: Schematic view of the High-Level Synthesis approach applied in the SeqAn@FPGA project. The algorithm is implemented in the SYCL language, a C++ extension, and the Intel oneAPI toolchain for FPGAs is used to generate a hardware circuit design based on the data-dependency graph.

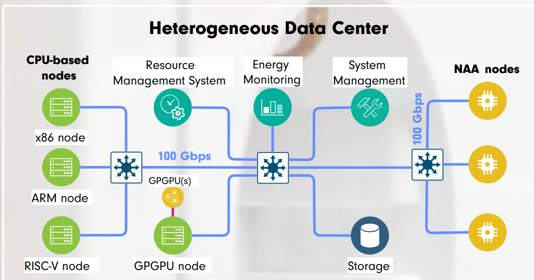


Figure 2: Illustration of a heterogeneous data center infrastructure with Network-Attached Accelerators (NAAs, right) which are decoupled from the increasingly diverse compute nodes (left) and connected via a 100 gigabit-per-second high-speed interconnect.

GreenHPC Projects at ZIB

Within the BMBF-funded projects NAAICE and SeqAn@FPGA, ZIB is developing software prototypes for energy-efficient computing based on FPGA technology.

The **SeqAn@FPGA** project addresses the rapidly growing demands of genomic data analysis in biomedicine, driven by next-generation sequencing and personalized medication. To achieve significant energy savings, we implemented selected computational kernels on an FPGA platform and integrated them into the widely used SeqAn sequence analysis library, developed at Freie Universität Berlin. Using the "Raptor" demonstrator application, we demonstrate a five- to eight-fold reduction in energy consumption on an Altera Agilex 7 FPGA compared to a contemporary CPU at comparable performance levels.

In the **NAAICE** project we are going one step further by exploring network-attached accelerators (NAAs): FPGA-based accelerators that are directly connected to an HPC fabric without requiring a conventional host system (see Fig. 2). This concept is motivated by the observation – also confirmed in the SeqAn@FPGA project – that even energy-optimized host systems consume significantly more power than an FPGA accelerator equipped with local memory. ZIB's contribution to NAAICE focuses on developing a data center-friendly infrastructure that enables host-less FPGA accelerators to be flexibly managed and allocated according to user requirements via resource management systems such as Slurm. In addition, a monitoring framework is being developed to make the energy consumption of FPGA platforms accessible to standard data center monitoring tools. Using a demonstrator application from the geosciences, the project has successfully validated this concept across the full software and hardware stack: from infrastructure supporting high-speed communication with NAAs (see Fig. 3), through application-level middleware for offloading kernels written in high-level languages, to data center operation, scheduling, and energy assessment.

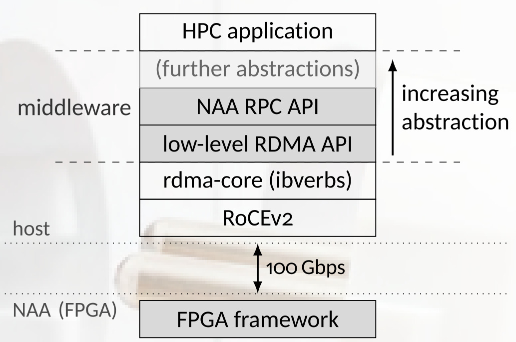


Figure 3: Communication software stack of the NAAICE project. Remote Direct Memory Access lays the foundation for a lightweight Remote Procedure Call API which is exposed to the application and hides the low-level details from the developer. Gray boxes indicate contributions made over the course of the project.

Outlook

Recently, NextSilicon introduced their dataflow-based accelerator architecture that further advances software-defined hardware. Existing C, C++, or FORTRAN applications already parallelized with OpenMP can be adapted to this accelerator. The compiler and runtime environment assist developers in identifying suitable computational hotspots and help to generate dataflow-oriented mappings that make use of the accelerator's on-chip functional units, such as floating-point, integer, and memory-access units. ZIB is currently evaluating this architecture in collaboration with NextSilicon using selected application workloads, with the goal of assessing its potential for future energy-efficient HPC systems.

Steffen Christgau and Thomas Steinke


MaRDI Vision:
Building a community that embraces a FAIR data culture and research workflow through the sustainable realization of MaRDI findings.
DATA
SCIENCE

INTERVIEW

Scientific knowledge today is increasingly digital. Research results are no longer just written down in papers, but are accompanied by software, data, models, and computational workflows. If these digital components are poorly documented or disappear, results become difficult – or even impossible – to verify and reuse. This is why research data management matters: it helps ensure that scientific work remains understandable, reusable, and trustworthy over time. When all relevant components are available and well described, research becomes reproducible, allowing results to be checked, analyses to be repeated, and new work to be built reliably on existing findings.

These challenges are addressed by the National Research Data Infrastructure (NFDI) project, which brings together institutions from many disciplines to coordinate and support sustainable research data management, guided by the vision of research data as a shared resource for excellent science. One of the initiatives within NFDI is MaRDI – the Mathematical Research Data Initiative – which focuses on research data in mathematics and on developing innovative ways to represent, connect, and reuse mathematical knowledge.

To explore how this works in practice, we spoke with Marco Reidelbach and Tim Conrad from the Zuse Institute Berlin, who are closely involved in key MaRDI services: MaRDMO, which supports researchers in planning their data management, and the MaRDI Portal and Knowledge Graph, which together provide a central, semantically connected access point to mathematical research data.

Interviewer: Marco, could you briefly introduce MaRDI?

Marco Reidelbach: MaRDI stands for the Mathematical Research Data Initiative. Within the NFDI, it represents the mathematical community and is supported by a broad consortium of research institutes, universities, and scientific organizations from the field of mathematics and its applications. Together, we develop services that support users in finding, documenting, and reusing mathematical research data in a reliable and reproducible way.

Interviewer: “Research data in mathematics? Never heard of it!” What would you say to that?

Marco Reidelbach: That reaction is very common. Many people associate research data with experiments or measurements, but in mathematics the picture is different. Research data can be formulas, proofs, software, algorithms, models, or benchmark datasets. These are central to modern mathematical research, and MaRDI works to ensure that they’re properly described, connected, and reusable – both within mathematics and in other disciplines that rely on mathematical methods.

Interviewer: What role does the Zuse Institute Berlin play in MaRDI?

Marco Reidelbach: At the Zuse Institute Berlin, we’re responsible for several MaRDI services and provide much of the project’s IT infrastructure. One key service is MaRDMO, which extends an existing data management planning tool with mathematical content and guidance. Together with FIZ Karlsruhe and other MaRDI partners, we also develop and operate the MaRDI Portal and the MaRDI Knowledge Graph. These components form the backbone for discovering, linking, and exploring mathematical research data across many sources.

Interviewer: Tim, could you explain the MaRDI Portal and the Knowledge Graph in simple terms?

Tim Conrad: The MaRDI Portal is the central entry point to MaRDI's ecosystem of services. It helps users discover what MaRDI offers and navigate to the right tools – some services provide access to mathematical research data and knowledge, others support documentation, workflows, interoperability, or good research data practices. One of these services is the MaRDI Knowledge Graph, which connects information about publications, software, models, algorithms, datasets, and people through semantic relationships. This allows users not just to find individual items, but to understand how different pieces of mathematical knowledge relate to one another.

Interviewer: Why is this knowledge graph approach important?

Tim Conrad: Mathematics is highly interconnected. A single formula might appear in many publications, be implemented in different software tools, and be used in models across several disciplines. The knowledge graph captures these relationships explicitly. This makes it easier to discover relevant work, to see how ideas are reused, and to place new results into a broader context.

Interviewer: Marco, how does MaRDMO fit into this picture?

Marco Reidelbach: MaRDMO addresses the early stages of research. It helps researchers think systematically about what data they will create, how it should be documented, and how it can later be shared. By aligning MaRDMO with the portal and the knowledge graph, good planning directly supports visibility and reuse later on.

Interviewer: How does all of this help researchers in practice?

Marco Reidelbach: It saves time and increases reliability. Researchers expend less effort searching for existing resources or recreating results that already exist.

Tim Conrad: And the benefits go beyond mathematics. Many scientific fields depend on mathematical methods. If mathematical knowledge is well structured and accessible, it becomes easier for researchers in other disciplines to apply and build on it correctly.

Interviewer: And what's next for MaRDI?

Marco Reidelbach: We're preparing for the second phase of MaRDI. If approved, it will allow us to consolidate and further develop our existing services, expand the knowledge graph, and strengthen collaboration with the mathematical community. At the same time, we want to make our tools more robust, better integrated, and easier to use in everyday research. The long-term goal is a stable and trusted ecosystem in which mathematical research data, software, and workflows can be reliably shared, connected, and reused.

Berlin Mathematics Research Center

MATH+

RENEWED:

SEVEN MORE YEARS FOR MATHEMATICS WITH IMPACT!

Michael Hintermüller, Sebastian Pokutta, Martin Skutella, Andrea Walther, Christof Schütte, Claudia Schillings

Interviewer: Early summer 2025 brought great news: MATH+ has been renewed for another seven years. What does this mean for ZIB?

Andrea Walther: It's a strong vote of confidence in what the Berlin mathematics ecosystem has built together. The renewed funding recognizes the achievements of the first phase and, even more importantly, it gives us the runway to deepen what MATH+ stands for: applications-oriented mathematics that engages with society's big challenges. For ZIB, that's very tangible – we're strongest when methods and real-world questions meet, and MATH+ is exactly the environment where that happens.

Sebastian Pokutta: I agree – and it's also a signal internationally. The MATH+ renewal underscores Berlin's position as a hub where mathematical depth meets AI, data, and decision-making at scale, especially when it's tied to data, models, and decision-making. At ZIB, we translate this into projects that reach beyond academia, into industry and public stakeholders, without losing mathematical depth.

Interviewer: MATH+ has been shaped for long-term and sustainable impact. How is the Cluster managing this generational transition in a way that keeps momentum and continuity?

Christof Schütte: We're handling it with a lot of respect for what has been built – and with real confidence in what comes next. I'm genuinely proud that a new chair team has taken over, and that it includes Sebastian and Andrea. They bring fresh energy, clear strategic focus, and a strong sense for what MATH+ needs in its next phase. In short: the ship is in excellent hands.

Interviewer: MATH+ is often described as "mathematics for the great challenges of our time". What are the themes behind that tagline?

Andrea Walther: At the heart of MATH+ are data-driven modeling, simulation, and optimization. These are the tools you need when you want to understand complex systems and also make decisions under conditions of uncertainty – whether that's in energy systems, mobility, health, or technology. The renewal period continues this trajectory, but also expands conceptually in response to global crises.

Interviewer: "Conceptual expansion" sounds ambitious. What changes in the new phase?

Sebastian Pokutta: One key addition is a sharpened focus on the mathematical foundations of AI. AI is everywhere – yet many core questions about reliability, explainability, and what's really happening "inside" these systems are still open. For MATH+ and ZIB, that's not just a buzzword topic. It's a mathematics agenda: understanding generalization, robustness, uncertainty, and how to build AI that we can trust in high-stakes settings. Another key aspect is the a strategic move to complement traditional modeling with mathematical models of human behavior – because innovation doesn't happen in a vacuum.

Interviewer: ZIB is known for method development that lands in practice. Where do you see the strongest "ZIB signature" inside MATH+?

Sebastian Pokutta: Mainly two places. First, the direct application pipeline: we take real data and real constraints seriously, so methods don't remain idealized. Second, public–private partnerships and transfer. MATH+ explicitly wants to strengthen engagement with stakeholders and expand PPP formats. A key partner here is the MODAL Research Campus – Germany's largest public-private partnership in applied mathematics – which is closely linked to MATH+ in knowledge transfer and societal impact. ZIB plays a central role in making those collaborations productive: turning a partner's problem into a mathematical model, an algorithm, and ultimately a tool someone can use.

Interviewer: Speaking of people, what about early-career support and visibility beyond science?

Andrea Walther: Training and mentoring remain at the heart of what we do. Through the Berlin Mathematical School, MATH+ runs a globally recognized graduate program and emphasizes mentoring, diversity, and equal opportunity. And we're also investing in communication and dialogue. New formats like the "Decision Theatre" on sustainable mobility, and a dedicated science communication effort in collaborations with the public, are designed to make complex topics accessible – and to build trust through transparency.

Interviewer: Final question: what's your personal takeaway from the renewal?

Sebastian Pokutta: It's both recognition and responsibility. We now have seven years to push AI and data-driven mathematics forward in a way that is rigorous *and* useful.

Andrea Walther: And to keep the spirit of MATH+: excellent mathematics, done together, with a clear line of sight to society. +

EnergyLab

MedLab

MobilityLab

SMART SOLUTIONS FOR OUR DIGITAL SOCIETY: RESEARCH CAMPUS MODAL PROCEEDS TO 3RD FUNDING PHASE

NanoLab

HPCLab

SynLab

On April 1, 2025, MODAL started its 3rd funding phase, which will be supported by the Ministry for Research, Technology and Space Travel (BMFTR) with up to €8 million in the first four of the next five years. Supported by three research partners, Freie Universität Berlin (FUB), Hochschule für Technik und Wirtschaft Berlin (HTW), and Zuse Institut Berlin (ZIB), which bring to the table complementary competencies and sophisticated infrastructure, as well as 17 industry partners, MODAL addresses major societal challenges.

The unique strength of the Campus lies in the advancement of the mathematics of artificial intelligence and optimization, enabling transferable methods that can be applied across diverse application domains. MODAL research activities are organized in six laboratories:

The **EnergyLab** focuses on the design and operation of sustainable energy systems. Its decision support tools are used by Open Grid Europe to predict gas flows across large parts of Germany (OAD forecasting), for gas dispatching (KOMPASS), and to design CO_2 and hydrogen transport networks (COCOS), and by Berliner Energie und Wärme to develop a proof of concept for holistic portfolio planning of Berlin's heat network. Open-source tools such as Fyling SCIP-Jack contribute to the optimal placement of wind parks. Phase III will address heat networks and a novel hydrogen chloride-based energy storage technology.

EnergyLab head Janina Zittel presents the research agenda for phase III.

The **MedLab** advances data-driven methods for personalized medicine. Its AI-based approaches have been integrated into industrial platforms, including ECG classification for home monitoring (Biotronik), image analysis methods (1000shapes), workflows for large-scale proteomics data (KNIME), and biomedical image analysis tools (Thermo Fisher Scientific). In the coming phase, the focus will be on secure and privacy-preserving biomedical applications.

The **MobilityLab** develops optimization methods for smart mobility and logistics. Several are used in industrial systems such as the flight route optimizer VOLAR (in Lido Flight 4D by Lufthansa Systems), the electric bus optimizer ES-OPT (in ivu.plan by IVU Traffic Technologies), and a demonstrator for air cargo load planning (with Ab Ovo). Railway dispatching will be a key topic in Phase III.

The **NanoLab** engineers nanophotonic devices. The software package RPExpand enables efficient, mode-based design of integrated optics and materials for photochemical processes. Quantum-photonic components are to be considered next.

Berlin's senator for Science, Health, and Healthcare, Ina Czyborra, FUB president Günter Ziegler, and BMFTR department head Karsten Hess, at the opening of the MODAL FUHUB premises on June 6, 2025.

The **HPCLab** supports high-performance computing by advancing vendor-independent programming standards such as SYCL and OpenMP on GPUs (NVIDIA, Intel, AMD), vector CPUs (NEC), and many-core SIMD CPUs (Intel), and by improving the energy efficiency of ARM processor architectures (Ampere, Huawei) with the Intel Cascade Lake AP architecture (the HPC system "Lise"). Design and operation of high-speed networks are the upcoming research object.

The **SynLab** bundles all MODAL developments in optimization, learning, and artificial intelligence methods in leading open source frameworks including SCIP, (mixed-integer linear and nonlinear optimization problems), Ubiquity Generator (parallelized branch-and-bound on supercomputers), FrankWolfe.jl and Boscia.jl (mixed-integer convex optimization), and SCIP-Jack (Steiner Tree Problems in graphs and 14 related problems). Methods for the globally optimal solution of MINLPs have been transferred from SCIP to the current release of the Xpress Global solver by Fair Isaac. The software information service swMATH (zbmath.org/software/) is part of Zentralblatt MATH. First order methods and science communication are the topics of Phase III.

MODAL has reached a major milestone with the opening of new premises at the FUHUB building within the FU Business and Innovation Center (FUBIC), a large innovation park close to FUB and ZIB. With its modern co-working space, MODAL is set to take collaboration among partners to the next level, in particular with the neighboring Scale-Up Lab of FUB Chemistry. The opening was celebrated in mid-June with guests from politics, academia, business, and society, who recognized the campus's importance for Berlin's research landscape, most notably in collaboration with the MATH+ Center of Excellence and within the Berlin University Alliance.

Ralf Borndörfer & Sebastian Pokutta

IMPRINT

PUBLISHED BY
Zuse Institute Berlin
Takustrasse 7 | 14195 Berlin | Germany

PHONE +49-30-841-850
FAX +49-30-8418-5125
EMAIL info@zib.de
www.zib.de

ISSN
0934-5892

LAYOUT
Tunermaxx Media GmbH
www.tunermaxx.com

PRINTING
Lochmann grafische Produktion GmbH